THE EFFECTIVENESS OF ANNUAL NATIONAL ASSESSMENT IN MONITORING MATHEMATICS EDUCATION STANDARD IN SOUTH AFRICA

by

ZWELITHINI BONGANI DHLAMINI

Thesis

Submitted in fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MATHEMATICS EDUCATION

in the

FACULTY OF HUMANITIES

(SCHOOL OF EDUCATION)

at the

UNIVERSITY OF LIMPOPO

SUPERVISOR: Professor J.K. Masha

CO-SUPERVISOR: Professor I. Kibirige

2018

DEDICATION

This work is dedicated to

My mother, the late Dinah Thokozile Dlamini,

You raised me as a single mother. You are so special to my life, your legacy still remains with us. Your spirit still lingers on. You were a true loving mother, a giant of all times and an inspiration to my life, may your soul rest in peace,
I will always love you.

M**y daughters Mapaseka and Precious Dhlamini**

For sacrificing your valuable time and allowing your father to be away and your pride in having me as your father.

My son Ciniso Mncedisi Dhlamini

For your understanding that I could not provide everything during these difficult times when you were also a University student. I hope I have showed you that education is the key to success. Pave the way and show your sisters and make the Dhlamini family proud.

DECLARATION OF ORIGINALITY

I declare that “The Effectiveness of Annual National Assessment in Monitoring Mathematics Education Standard in South Africa” is work of my own and that all sources that I used or quoted have been acknowledged by means of absolute references.

03 September 2018

Signature

Date

Mr. Z.B. Dhlamini

ACKNOWLEDGEMENTS

I thank GOD the Almighty for giving me strength and wisdom, knowledge and dedication to persevere during difficult times. I would have not reached this far if it was not your will. My pillar of strength has been that you planned this to happen for a reason. Please bless me and give light for the future.

I would like to express my sincere honour and appreciation to a number of people who dedicated their effort and time to the success of this work. Words of thanks are directed to:

- Prof Kwena Masha, my supervisor, for allowing me to do the work with the university. You welcomed me with honour and acted as a true parent to me. Your true academic leadership guided me throughout this work. Your humble approach always gave me insight into the right direction. I salute you. If everyone in this province and the university follows your steps, will succeed.
- Prof Israel Kibirige, my supervisor, for being there for me when it was difficult. You provided true academic leadership and showed me what it means to be an academic.
- My Head of Department, Prof S.K. Singh, for being the father figure and understanding at all times. For allowing me time away to do my studies when there were difficult times in my career. I appreciated everything, please do the same the all in the DMSTE family.
- The DMSTE department for making me feel at home and for the support that you gave as colleagues. You are a truly professional and loving academic family. You became my home away from home.
- The schools that supported me with the necessary data. You showed academic maturity even when the ANA results were not that good.
- Last, Bernice McNeil for your professional work done in editing this document.

ABSTRACT

The purpose of the study is to explore the effectiveness of Annual National Assessment (ANA) in monitoring the standard of mathematics education and to assess the mathematical proficiencies tested and exhibited by Grade 9 learners in South Africa. The research problem was premised on the dearth of data that justifies ANA as an evaluative assessment. As such, the study utilised five strands which were; procedural fluency, conceptual understanding, strategic competence, adaptive reasoning and productive dispositions as a theoretical framework to assess mathematics that was tested and exhibited by learners. To explore the research problem, the study used mixed methods in the context of exploratory sequential design. Document analysis was used first to capture mathematics content and cognitive levels examined by ANA. Second, learner responses were explored using four variables of achievement levels; no response, correctly answered, incorrectly answered and partially answered.

First, the results from the analysis of ANA questions indicated that ANA mostly tested questions of low complexity. Second, the results from the learners' responses revealed that the majority of learners were not proficient to ANA irrespective of low complexity testing. Third, the Porter's alignment index for ANA and TIMSS was between moderate and perfect. Subsequently, content and cognitive levels were misaligned in the three consecutive years of ANA testing. It implies that learners were most likely to show a deficit of higher order problems solving skills which are a prerequisite of courses in advanced mathematics. Additionally, the results suggest that ANA had challenges of reliability and validity as an evaluative assessment due to inconsistency in the testing. As such, it is recommended that the complexity of ANA be addressed, the content areas where learners are not proficient be addressed and the alignment of ANA must be frequently calculated to monitor the standard of mathematics education in South Africa effectively.

Keywords

Alignment index, mean deviation, strands of mathematical proficiency, systemic assessment.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AMESA	: Association for Mathematics Education of South Africa
ANA	: Annual National Assessment
DBE	: Department of Basic Education
DFID	: Department for International Development
DoET	: Department of Education and Training
IEA	: International Association for the Evaluation of Educational Quality
OBE	: Outcomes Based Education
SACMEQ	: Southern and East Africa Consortium for Monitoring Educational Quality
SAGM	: Subject Assessment Guideline for Mathematics
SMC	: Singapore Mathematics Curriculum
SMP	: Strands of Mathematical Proficiency
TIMSS	: Trend in International Mathematics and Science Study

1. CHAPTER ONE
ORIENTATION TO THE STUDY

1.1 Introduction

In the release statement of the 2013 Annual National Assessments results, the Minister of Basic Education labelled the Grade 9 (ANA) for mathematics as a 'missing link' and urged researchers to inform the Department of Basic Education (DBE) about causes of the low achievement of 13%, 14% and 11% in 2012, 2013 and 2014 respectively (DBE, 2012a, 2013a, 2013b, 2014). Such a low achievement poses serious concern to the DBE which has recently implemented the use of ANA to monitor the education system. The introduction of ANA in South Africa has seen both positives and negatives; with challenges in: epistemology, psycho-genetic and didactical obstacles (Association for Mathematics Education of South Africa AMESA, 2012; Bantwini, 2010; Graven & Venkatakrishnan, 2013).

This study is an epistemological analytical inquiry concerning the use of ANA in South Africa to monitor the education system. Firstly, the inquiry began with an analytical review of the Grade 9 mathematics ANA question papers to document strands of mathematical proficiencies (SMP) that were questioned (Kilpatrick, Swafford & Findell, 2001). Second, three questions in the 2014 Grade 9 ANA learners' responses were analysed to examine SMP that were exhibited by learners in the ANA test. Lastly, the study verified the alignment of ANA and the Trend in International Mathematics and Science Study (TIMSS) (Porter, 2002). Through this inquiry the study contributes an alternative way of reporting of ANA results in South Africa that aims to convey a meaningful message to all stakeholders now that ANA is being used to monitor the South African Education system (DBE, 2013b).

Mathematics knowledge, skills and values in this study are regarded as SMP, a theoretical framework used to examine mathematical content skills intended by the curriculum. These skills must be portrayed in the ANA questions, and experienced by learners as exhibited in their scripts. Kilpatrick et al., (2001) outlined the SMP as:

conceptual understanding, procedural fluency, strategic competence, adaptive reasoning and productive disposition (Dhlamini & Luneta, 2016). Some significant studies contextualise the existence of SMP in mathematics classroom discourse (Hiebert & Lefevre, 1986; Groves, 2012; Schoenfeld, 2007; Suh, 2007) although there is a dearth of literature on studies that use the SMP in document analysis. Through literature that informs each SMP, this study advances SMP to be examinable through document analysis.

1.2 Background Context in National Systemic Assessment Testing

There are three policy issues that national assessments must address and these are, Quality, Equity and Provision. Issues relate to: *Quality* teaching and learning referring to what mathematical knowledge learners' exhibit. *Equity* the national assessment can help to determine how the education system is responding to gender, socio-economic diversities, ethnic groups and school governance (public or private). *Provision* a national assessment must provide evidence on the provision of education such as the challenges of curriculum reform, learner retention rate and its effect on teaching and learning. (Department for International Development DFID [Sa]). This study is located in the first policy issue, issues relating to quality. Subsequently, the DBE pronounced that ANA is premised on three roles, namely: 1) as a measure for achievement for every three years of study, that is, Grade 3, 6 and 9; 2) ANA is used to assess the suitability of the curriculum at specific intervals in order to determine where to improve. ANA plays a role in providing valuable data that achieves sound levels of reliability; and 3) In addition to ANA, formative and summative assessment are used to provide validity and reliability of assessment processes (DBE, 2012. Additionally, the effective implementation of the curriculum in terms of making sure that the intended outcomes are addressed during teaching, learning and assessment is vital in the success of ANA (DBE, 2014). Hence it is vital that ANA testing that has been paused, continues to enhance quality education through assessment.

South Africa under-performed in the 2003 TIMSS, but contrary to that, South African learners indicated that they enjoyed mathematics. The poor achievement by South African learners did not imply a total revamp of the education system but rather focused on diverse cultural factors that might have negatively affected teaching, learning and learners' achievement, (Leung, 2005).

South African learners performed poorly in mathematics in all TIMSS and other Southern and Eastern Africa Consortium for Monitoring Education Quality (SACMEQ) studies (Leung, 2005; 2014; Howie, 2003; 2004; Kotze & Strauss, 2006; Reddy, 2006). These results focused on socio-economic obstacles and not epistemological, psycho-genetic and didactical obstacles, and the omission of these represents a serious knowledge gap. These studies have compared achievement amongst the participating countries and revealed mostly South Africa's socio-economical obstacles. The implementation of ANA has compared learner aggregated scores in the provinces, districts and schools, and also revealed socio-economical obstacles (DBE, 2012a & 2013b). This is not supposed to be the only focus of a national assessment. As such, it is a replica of what the international and regional systemic assessments have revealed where South Africa has participated (DFID, [Sa]; Howie, 2003, Kotze & Strauss, 2006).

Graven and Venkatakrishnan (2013) outline unfolding issues in the implementation of ANA:

> *"The introduction of the Annual National Assessments began in 2011. The ANA was explicitly focused on providing system-wide information on learner performance for both formative purposes, such as providing class teachers with information on what learners were able to do, as well as summative purposes, such as providing progress information to parents and allowing for comparisons between schools, districts and provinces…Assessments such as ANAs of course have an influence on what happens in schools and in classrooms."* (Graven & Venkatakrishnan, 2013: 12).

As some means of providing an alternative to policy makers, this study examines issues of quality in the ANA testing. There are other equally important issues, such as provision and equity which cause a dilemma in this discourse and still need to be researched to provide worthwhile data.

What follows below is an outline of the research problems that prompted this researcher to undertake this study.

1.3 The Research Problem

Assessment is a fundamental component that drives teaching and learning. In schools, teachers use different assessment instruments to gauge learners' achievements (Gonzales & Fuggan, 2012). In addition, teachers are expected to do evaluative assessment of the curriculum in use. Several countries use systemic assessments to monitor the functionality of their systems and South Africa is not an exception (Kanjee & Moloi, 2014) because it uses ANA, which started in 2012 the DBE for Grades 3, 6 and 9 to monitor the performance of the curriculum (Graven & Venkat, 2014). In the past, there has been no instrument to gauge how well the curricula in use were doing. ANA as an instrument was introduced to address that gap. ANA as an evaluative instrument was envisaged to gauge how well the curriculum is succeeding. The problem is that since the introduction of ANA, there has been no information as to whether it has been able to monitor the standard of the curriculum in schools (Kanjee & Moloi, 2014). Since 1994, post-apartheid, South Africa has changed curricula from NATED 550, to Outcomes Based Education (OBE), the National Curriculum Statement (NCS), the Revised National Curriculum Statement (RNCS) and finally, to the Curriculum and Policy Statement (CAPS) (DBE, 2012a; DoET, 2002a; 2002b). In line with such drastic changes in the curriculum it is most likely that teachers would face challenges to catch up and this was supported by Leung (2005) who contends that any curriculum change must be as a result of evidence from classroom practice and other relevant issues. The challenges that teachers face due to curriculum changes are most likely to affect learners' achievements (Bansilal, 2012). However, learner achievement is not the focus of this study, as Kanjee and Moloi (2014) point out that ANA is an evaluative assessment which focuses on providing information for purposes of improving teaching and learning (Graven & Venkat, 2014). The expected outcome of curriculum change would be high quality education if the standard is monitored and maintained (Volante & Cherubini, 2010). There are still questions concerning the desirable mathematics that need to be learned in schools (Maoto, Masha & Maphutha, 2016).

Some studies have focused on challenges on the implementation of ANA in Grade 3 and 6 (Graven & Venkat, 2014; Kanjee & Moloi, 2014) and there are no studies on Grade 9, which presents a knowledge gap. In both studies, teachers suggested that there were challenges as a result of implementation. However, teachers were positive with the usefulness of ANA to their teaching. Another study by Long and Wendt (2017) raised that systemic assessments, and ANA is not an exception, lack studies on what learners can and cannot do, a knowledge gap in this discourse. Only after three consecutive ANA testing in Grade 9 mathematics that were marred by poor performance, there were disagreements between the DBE and teacher Unions that halted ANA testing. Some issues raised by the teacher unions were that ANA was no longer serving its purpose of improving teaching and learning. Instead DBE used it to name and shame low performing schools and provinces (South African Teachers' Democratic Union SADTU, [Sa]).

An analysis of the 2012 Grade 9 mathematics ANA by AMESA (2012) points to obstacles posed in the implementation of ANA that need to be researched to inform policy-makers on the effectiveness of ANA. To address the knowledge gap in Grade 9 ANA testing, this study examines the evaluative instrument, ANA, and assesses whether it serves its purpose. Currently, the success of the existing curriculum has not yet been documented. Therefore, this study assesses the effectiveness of ANA as an instrument that monitors the standard of mathematics education and the success of the curriculum in use in South African schools.

Below, I outline the purpose and the research questions of this study.

1.4 Purpose of the Study and Research Questions

To respond to the research problem, the purpose and research questions of the study are outlined.

1.4.1 The Purpose of the Study

The purpose of the study is to explore the effectiveness of Annual National Assessment in monitoring the mathematics education standard in South Africa through assessing the SMP tested by ANA and exhibited by Grade 9 learners in South Africa. This is achieved by reviewing cognitive demands of Grade 9 mathematics ANA in relation to SMP and viewing learners' scripts to analyse levels of mathematical proficiencies exhibited. To achieve this purpose, the study responded to the following research questions.

1.4.2 Research Questions

The inquiry undertaken in this study responded to the following research questions:

- How are cognitive levels of mathematics tested by ANA reflective of SMP?
- What levels of mathematical proficiency do learners exhibit in response to the ANA tests?
- How do the content and cognitive levels tested by ANA compare with TIMSS?

The next section focuses on the importance of embarking on this study.

1.5 Significance of the Study

To address the knowledge gap outlined in the research problem, the current study examines the effectiveness of ANA in monitoring the standard of mathematics in three ways. Firstly, this study adapts SMP to be compatible for document analysis, and addresses a knowledge gap in this discourse. Second, the analysis of ANA question papers and learners' scripts, is a methodological alternative that reports ANA results on proficiency levels instead of aggregated scores (DoET, 2002a). Third, the calculation of Porter's alignment between ANA and TIMSS, content standards may be maintained to prepare South African learners for effective participation in international systemic assessment testing. This will go a long way in enhancing quality education through assessment. Although, during the course of this study, there has been a pause in ANA testing, the significance of the study noted here is of paramount importance for future consideration of ANA testing.

Below is an outline some key concepts that are used throughout this study.

1.6 Key Concepts

An early discussion of concepts serves as a reference in organising the structure of a thesis as well as providing a tool for metacognition in a doctoral study (Berman, 2013). Bordage (2009) explains that the discussion of key concepts of a study supports the explanation of discursive multiple theoretical frameworks, literature base and the organisation of the professional educational context of a study. Below are some key concepts of the study.

1.6.1 Strands of Mathematical Proficiency

Kilpatrick et al. (2001) introduced the SMP as being intertwined, interconnected and inseparable. These SMP are: procedural fluency and conceptual understanding (knowledge); strategic competence and adaptive reasoning (skills); and productive disposition (values). (Groves, 2012; Luneta & Dhlamini, 2012).

The concept of a rope that Kilpatrick et al. (2001) used to represent each strand symbolises a component of mathematical proficiency. When the strands are intertwined the resultant rope is stronger than each strand, which symbolises strong mathematical knowledge, skills and values. This study used the SMP as theoretical framework to explore the mathematics cognitive levels examined by ANA tests and subsequent learners' responses. This researcher used SMP because they include mathematical knowledge, skills and values, and hence they address mathematics holistically, (Ally & Christiansen, 2013; Schoenfeld, 2007).

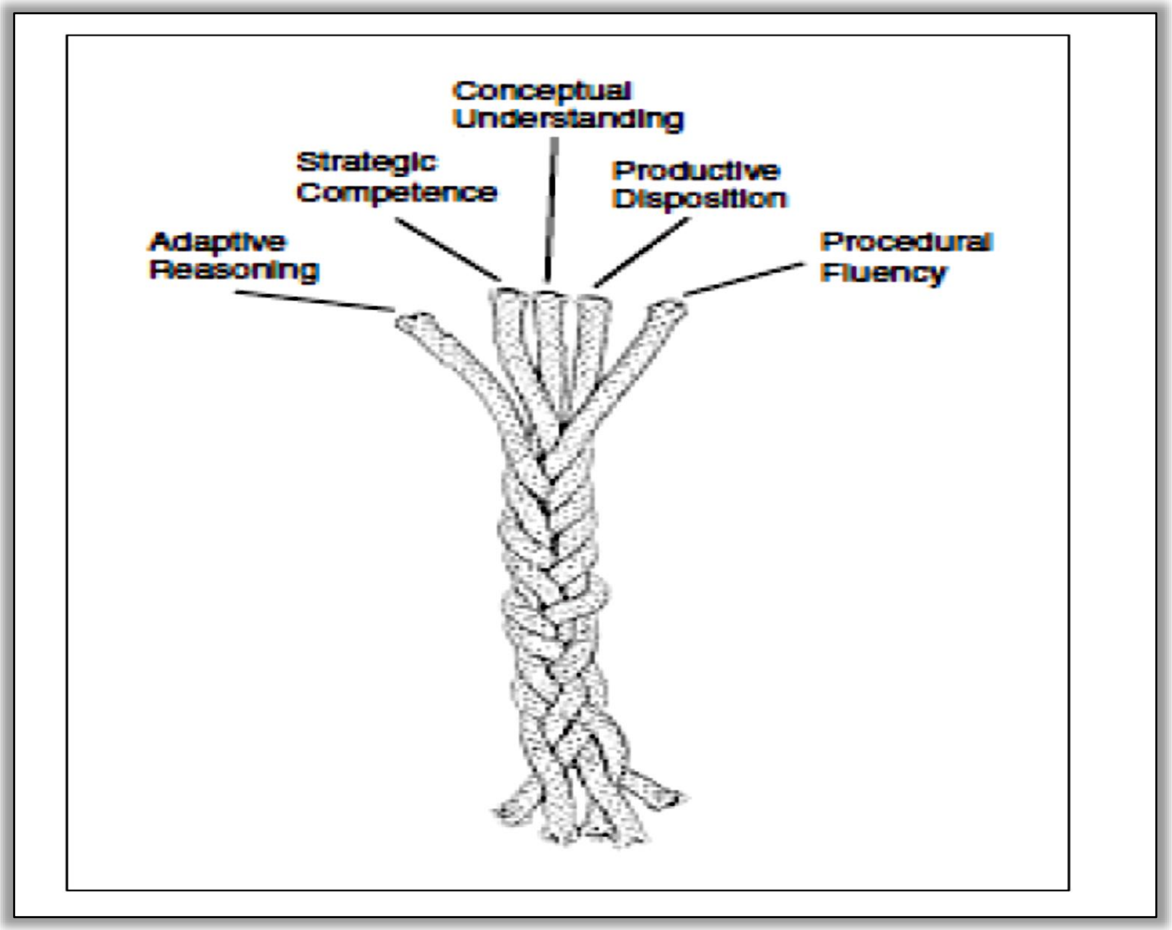


Figure 1.1: Intertwined SMP (Kilpatrick et al., 2001: 117)

1.6.2 Systemic Assessments

Systemic assessment is a way of externally evaluating the education system by a comparison of learners' performance against nationally set indicators of learner achievement (DoET, 2002a&b). Systemic assessments provide policy makers with evidence on students learning in an education system, (Kellaghan, Greaney & Murray, 2009). This implies that systemic assessment can be done internationally, regionally and nationally. Dunne, Long, Graig and Venter (2012) argue that systemic assessment must assess current performance and variability within a group of learners of a certain age using externally set standards that outline required proficiency levels that monitor progress over a certain period of time. Additionally, Dunne et al. (2012) pointed out that the results of systemic assessment must be interpreted in conjunction with other forms of assessment such as assessment for learning. Another important question is: to what extent are the ANA achieving what a national systemic assessment ought to achieve?

1.6.3 The Alignment Index

To calculate the alignment index, matrices are formed which match assessments and curriculum in cells and the alignment index is calculated per cell.

Alignment index = 1 - $\frac{\sum |X-Y|}{2}$ where x and y stand for cells in corresponding matrices. (Porter, 2002). The alignment index ranges from 0 to 1, 0-0.50 which depicts no alignment to moderate alignment, 0.51 to 1.0 range from moderate to perfect alignment (Ndlovu & Mji, 2012; Porter, 2002; Porter, McMaken, Hwang & Yang, 2011).

1.6.4 South Africa's Cognitive Levels for Mathematics

The following percentage levels cognitive levels are expected when formulating assessments in the senior phase (Grades 7-9) for mathematics in South African schools: knowledge (25%); routine procedure (45%); complex procedures (20%; and problem solving (10%) (AMESA, 2012). The cognitive levels were adopted from the TIMSS 1995 study with no alterations (Berger, Bowie & Nyaumwe, 2010; DoET, 2007; Reddy, 2006). The use of similar cognitive levels standardises the South African ANA examination with international assessments.

1.7 Research Methodology

The current study used mixed methods in the context of the transformative paradigm. First, a review of SMP that had been examined by the ANA tests was done. Second, this study used document analysis to identify SMP that learners' exhibit in their responses to the ANA test questions. Lastly, the 2012, 2013 and 2014 Grade 9 mathematics ANA tests were aligned with the TIMSS 2011 Grade 8 mathematics test items by calculating the Porter's alignment to measure the content message that these document relate.

1.7.1 Research Design

The study combined both qualitative and quantitative design features which was in the context of the QUAL-quan, normally called the exploratory sequential design. According to Creswell (2014), this design is dominated by the collection of qualitative data that is followed by quantitative analysis. The research design was divided into three phases.

Part one Phase one was the development of themes and subthemes using the SMP to analyse the ANA Grade 9 question papers with special attention on identifying patterns on the SMP tested by the papers. This was preceded by Part one. Phase two generation of descriptive statistics, means and standard deviations to explain proficiencies examined by ANA in three consecutive years.

Part two, Phase one was to assess learners' responses to the ANA tests with focus on identifying SMP exhibited by learners in response to the ANA tests. Part two phase two was the generation of descriptive statistics, means and standard deviations to explain proficiencies exhibited by learners in various schools.

Part three, Phase one was the calculation of the Porter's alignment index in matrices of content and cognitive levels preceded by part three phase two, generation of descriptive statistics, means and standard deviations compare cells of content and cognitive levels of the alignment index of the Grade 9 ANA mathematics 2012, 2013 and 2014 question papers and 2011 TIMSS Grade 8 mathematics test items.

1.7.2 The Mixed Methods Approach

The study used mixed method research. It was dominated by qualitative document analysis, as the data was from a thick description of strands of mathematical proficiencies of the Grade 9 ANA test question papers and learners' responses to the test. Calculation of the alignment index for test items was both qualitative and quantitative. The description of learner responses from the scripts used proficiency levels (qualitative) and the analytical description of the SMP displayed by learners was done using descriptive statistics, means and standard deviations (quantitative).

Mixed methods research is a process of generating theory by researchers through integrating qualitative and quantitative techniques (Creswell, 2014). The approach is determined by the purpose of the research, the research questions and the context faced by the researcher (Johnson & Christensen, 2012). Mixed methods research is expanding and gaining popularity due to the variety of topologies it employs and an elaborate discussion of many topologies (Harrits, 2011; Luyt, 2012). Hence, mixed methods are relevant for this study as it also requires various topologies to explore the research problem which has not been explored in previous studies.

1.7.3 Sampling

There are two different samples that the study deals with due to the nature of research questions pursued. In relation to learners, the population is all Grade 9 learners in South African schools who have participated in the ANA for 2014. Through the application of the purposive sample procedures (Creswell, 2014), altogether 1250 learners in the seven schools in the Capricorn District of the Limpopo Province, were selected. In order to engage rigorously with the scripts, the focus for analysis was on items 3, 6 and 10. The choice of these items was motivated by their high relative frequency or anchoring role as observed in the ANA for 2012 to 2014. In pursuing the other purpose of the study, ANA tests for Grade 9 for 2012, 2013, and 2014 were selected. In order to benchmark the tests against international practices, the 2011 TIMSS Grade 8 mathematics paper was selected.

1.7.5 Research Assumptions

The research assumptions in the current study were informed by the transformative paradigm in the context of the epistemological transformative assumptions (Mertens, 2007 & 2010a). The assumptions were: first that assessment tools such as tests examine a spectrum of knowledge, skills and attitudes. Tests used in assessment can be benchmarked with other tests to reveal their content messages that they demand.

Second, tests must be set such that they are coherently balanced in the SMP. Third, in response to tests for any form of assessment, learners exhibit only the SMP that the test examines. Such tests allow learners to exhibit a variety of mathematical proficiencies such that, as learners progress in their schooling, they simultaneously acquire a coherent knowledge, skills and attitudes base that allows them to sustain learning situations that demand high level cognitive demands. The assumptions were informed by the SMP as identified by Kilpatrick et al. (2001) as well as the Porter's alignment index. Furthermore, this philosophical assumption is confirmed by Schoenfeld (2007) who states that; *"What You Test Is What You Get, (WYTIWUG)."* (Schoenfeld, 2007: 72).

1.7.4 Data Collection Methods

The Grade 9 ANA 2012, 2013 and the 2014 mathematics question papers were accessed from the website of the DBE. The 2011 Grade 8 mathematics TIMSS set of questions and were accessed from website of the International Association for the Evaluation of Educational Achievement (IEA). Learners' scripts were accessed from seven schools in the Limpopo Province, Capricorn District.

1.7.6 Data Analysis and Interpretation

This study first analysed ANA question papers and learners answers. Question paper were analysed first by categorising items into content areas, then the theoretical framework, the SMP was used to view each question item. Subsequently patterns emerged from the analysis of ANA question and were coded to capture SMP that each question examined. Furthermore, patterns were identified from the immerging codes and themes finally emerged. Last, the quantitative data was analysed using descriptive statistics: means and mean deviations.

Second, I analysed learners' answers (scripts) that was for their responses to question 3, 6 and 9 was categorised according to four variables, *correctly answered*, *partially answered*, *incorrectly answered* and *no response* that described learners' proficiency levels. When analysing available data (McMillan & Schumacher, 2014), researchers often visualise meaning from the documents as the respondents may not be accessible. For quantitative data, learners' responses across the schools were analysed using descriptive statistics: means and standard deviations.

Third, to benchmark ANA and TIMSS, the Porter's alignment index was calculated from Grade 9 mathematics 2012, 2013 and 2014 ANA question papers and the Grade 8 mathematics from 2011 TIMSS response items. The following formula was used to calculate alignment, Porters Alignment index = 1 - $\frac{\sum |X-Y|}{2}$, where x and y stand for proportions which can be written as X_i and Y_i, respectively. For quantitative data, descriptive statistics mean and standard deviation were used to show content and cognitive levels in the cells.

1.7.7 Quality Criteria of the Study

To ensure that this study maintains quality, I ensured that the following three principles are upheld, *credibility*, *confirmability* and *dependability* (Creswell, 2014). For credibility (Gay et al., 2014), I triangulated (Torrance, 2012) qualitative data collected from document analysis during data analysis using descriptive statistics in the quantitative paradigm. For dependability (Johnson & Christensen, 2012), instruments and theoretical framework were scrutinised against the relevant theories. In relation to confirmability (McMillan & Schumacher, 2014) I was not influenced by policy makers as this study dealt with available data and the results were given to peers to check bias.

1.7.8 Ethical Considerations

For ethical considerations (McMillan & Schumacher, 2014), informed consent, confidentiality and safety in participation, trust and, risks and benefits were all addressed. To gain informed consent (Creswell, 2014), I sought permission from the Department of Basic Education to access question papers and the 2014 mathematics ANA scripts. For confidentiality, I concealed the names of the schools and the learners by using pseudonyms (Springer, 2010). The school and the learners had the right to withdraw from the study at any time (Fraenkel & Wallen, 2010). To ensure trust (Gay et al., 2014). I kept the scripts safe and returned them the schools. Finally, to address the risks and benefits, I conformed (Johnson & Christensen, 2012) to copyrights for the ANA question papers and the TIMSS response items.

1.8 Scope and Delimitations

The research problem outlined earlier is centred on quality issues within epistemological obstacles. However, there are other equally important obstacles such as psycho-genetic and didactic issues. This focus poses a serious dilemma as studies on systemic assessment report mainly on socio-economic obstacles and disregard other pertinent issues that may assist in achieving important goals of national or international systemic assessment testing. The dilemma is two-fold. Firstly, there is dearth of literature on the quality issue especially within the context of systemic testing both national and international. The second challenge is that there are other equally important issues that are not the focus of this study such as psycho-genetic and didactic issues. The study was restricted to Grade 9 ANA scripts. This number does not include other section of ANA such as Grade 3 and 6 and future studies need to address those areas.

1.9 Overview of Chapters

This dissertation is divided into five chapters. Chapter One is an introduction to the study that includes a concise background of the study. That consists of the South African context of the research problem, the purpose of the study which presents the objectives and the research questions. The significance of the study is also outlined as well as an explanation of key concepts used in the study. A summary of the research methodologies followed in this study is given, as well as the quality criteria and finally the chapter concludes by an organisational overview of the thesis.

Chapter Two provides a review of the literature relevant to this study, including studies that report on regional and international systemic assessments where South

Africa has participated. Furthermore, there is a discussion on policy issues on conducting a systemic assessment, followed by South Africa’s ANA and issues on the calculation of the alignment index of systemic assessments. The literature on South Africa’s cognitive levels in mathematics is reviewed and a model for monitoring the education system using the ANA is proposed. Lastly, the theoretical framework, SMP are adapted to make them compatible to document analysis.

In Chapter Three, the research methodology, mixed methods are described. This is followed by the research design, the sampling procedures and the assumptions of the study. The research process is outlined, followed by the research ethics, and lastly challenges and strengths of this study are presented.

Chapter Four presents the findings of the study which begins by the presentation of findings of document analysis of the 2012, 2013 and 2014 ANA question papers. Furthermore, the findings of the analysis of learners’ responses to the 2014 Grade 9 ANA mathematics test are presented. Lastly, the results on the Porter’s alignment index between ANA tests and TIMSS response items are presented.

Chapter Five provides a summary of the main findings, highlights the importance of the study and how the findings contribute to the existing body of knowledge. It also provides suggestions for further research, and presents recommendations and conclusions.

1.10 Conclusion

The main purpose of this chapter was to present the background of the study, its focus, purpose and the methods. It provided the context of the study, its significance, research questions and key concepts used in the entire study. Furthermore, the quality criteria, ethical considerations and the structure of the thesis were discussed. The next chapter provides the literature review and the theoretical framework.

2. CHAPTER TWO
LITERATURE REVIEW AND THEORETICAL FRAMEWORK

2.1 Introduction

The previous chapter presented the background context of national systemic assessments in South Africa. Subsequently, it was observed that the history of systemic assessment testing has a dearth of literature due to the fact that it is a new process that has been recently introduced. On the other hand, South Africa has substantial history in the participation in international systemic assessment, TIMSS, in the following studies, 1995, 1999, 2003 and 2011 and did not participate in the 2007 study (Leung 2005; 2014; Howie, 2003; 2004). As such, the results of the mathematics TIMSS studies reveal that South Africa performed poorly, actually positioned last in all these studies. Although the results of the 2011 study show some significant increase in achievement, South Africa still remains last in mathematics achievement amongst participating countries (DBE, 2013b).

In a quest to view South Africa's national systemic assessment, literature in TIMSS and SACMEQ is reviewed. In his comparative studies, Leung (2005; 2014) warned South Africa that the low performance does not imply a total revamp of the education system, but rather issues that have direct impact on the low performance must be addressed. Studies have identified that only few learners perform above international average, from the minority of the South African population (Howie, 2003; Kotze & Strauss, 2006). The current study does not focus on achievement which has been widely reported in studies on systemic assessment testing. The focus is on finding alternative ways of ANA testing that may justify its use as a worthwhile monitoring mechanism (DBE, 2013b). An analysis of the 2012 Grade 9 ANA mathematics question paper by AMESA (2012) identified challenges relating to epistemological, psycho-genetic and didactical obstacles. Reporting of the ANA results focuses only on socio-economical obstacles which leaves this discourse of knowledge with a wide range of gaps in the literature.

The purpose of this study centred on quality issues by Grade 9 ANA now that they are used to monitor the education system. This discourse of knowledge seems huge for this study, therefore the current study reviews literature on three issues, alignment, cognitive levels tested by ANA in relation to strands of mathematical proficiency and strands of mathematical proficiency tested by the ANA as well as those exhibited by learners in response to the ANA tests. Finally, the SMP are adapted to be compatible to document analysis which is used in data analysis.

2.2 Monitoring Mathematics Education

The use of mathematics to monitor the quality of curriculum implementation has been widely accepted in various countries and South Africa is not an exception (Kanjee & Moloi, 2014). South Africa has changed curricula without data to inform the quality in the implementation (Graven & Venkat, 2014). The expectation is that teachers grapple with curricula changes and it is problematic when such changes are not informed by data. South Africa has engaged in three systemic assessments, TIMSS, the regional assessment Southern and Eastern Africa Consortium for Monitoring Education Quality (SACMEQ) and the ANA. In all these systemic assessments, mathematics achievement is monitored to inform the implementation of the curriculum. The assumption is that data gathered through the monitoring of mathematics education, this may inform policy makers as to the state of the education system (Volante & Cherubini, 2010). Below this section looks at the three levels of systemic assessment to gather information on South Africa’s participation.

2.2.1 Levels and Importance of Systemic Assessments

There has been a substantial increase in systemic assessment testing in the past decades through testing the quality of mathematics education (Drent, Meelissen and van der Kleij, 2013). There were 12 countries participating in The First International Mathematics Study (FIMS) in 1963 and recently there were 69 countries in the TIMSS 2011, (Drent et al., 2013; Emmett & McGee, 2013). Such an increase justifies the need to monitor the effectiveness of education systems, but the use of aggregated scores that compares the achievement of learners has been widely criticised as being less informative (Dunne et al., 2012; Suter, 2000). Studies on systemic assessments, especially TIMSS, have focused on student achievement, school level factors and classroom factors with less attention on the format of the tests (Chen, 2014). More attention has been given to mere aggregated scores and rankings of education systems irrespective of the substantial time, cost and effort spent on systematic testing (Koretz, 2009).

In the current study, the 2011 Grade 8 mathematics TIMSS response items are aligned with South Africa's ANA tests for purposes of benchmarking. The need for South Africa's participation in TIMSS and for it to conduct its own national assessments is justified, however, the relevance of using these results remains questionable. South Africa has participated in TIMSS 1995, 1995, 2003, and 2011 and was withdrawn in 2007 (Howie, 2003; 2004). Pressure from stakeholders resulted in the Minister of Basic Education reversing the withdrawal of South Africa in 2011. South Africa participated in the 2011 TIMSS and some improvement was noticed. However, the achievement was still lower than the international TIMSS average. More studies have reported challenges in the TIMSS such as sampling and the formats of the test items. Research conducted by Sofroniou and Kellaghan (2004) reported problems on the TIMSS 1997 mathematics test items that were in eight booklets. However, they fail to provide substantial analytic evidence connected with the problems posed by the test items. They further noted that stratified sampling was used to sample learners while teachers in participating schools were given questionnaires without being sampled and this created problems in validating the findings. These issues justify the need for intensive research on South Africa's Annual National Assessments now that it is used to monitor the education system.

Recently, South Africa has been involved in three levels of systemic assessment. The first is the TIMSS where South Africa participates with countries internationally (Reddy, 2006). Such results benchmark South Africa at an international level. Second, there is participation in SACMEQ, a regional systemic assessment that involves countries from South and Eastern Africa (Spaull, 2010). In this systemic assessment, South Africa benchmarks achievement regionally. Third, South Africa has its national assessments which benchmark achievement in schools and provinces (DBE, 2011). In this systemic assessment schools and provinces are benchmarked to gather data with the aim of improving teaching and learning (Kanjee & Moloi, 2014). Below I view the TIMSS and South Africa participation with an aim of mapping international testing and the South African context.

2.2.2 International Systemic Assessments

International systemic testing has been justified and it is widely agreed that it provides useful information on various educational systems (DFID, [Sa]; Schmidt & McKnight, 1998). South Africa participation in all the TIMSS, especially in mathematics has shown low achievement among participating countries (Howie, 2004). Some studies have focused on South Africa's participation and related factors (Howie, 2003; 2004; Leung, 2005; Wang, Osterlind & Bergin, 2012). However, little has been done to address concerns the raised from these results as South Africa's performance still remained lower than the international average in the 2011 TIMSS Grade 8 mathematics (DBE, 2014a).

A report on the 2011 TIMSS Grade 8 mathematics revealed a significant improvement on South Africa's achievement, however South Africa was positioned forty fourth out of forty five participating countries. The situation is bad considering that South Africa used Grade 9 learners for the Grade 8 response items when a majority of countries used Grade 8 learners. South Africa scored an average of 352, way below the 500 TIMSS Centre Point Human Sciences Research Council (HSRC)

(2013). These results still indicate that South Africa continues to perform poorly in mathematics when benchmarked internationally.

Figure 2.1 is an illustration of learners' performance and domestic economic state. There is an indication that in the 2011 TIMSS Grade eight, (32%) of learners in the participating schools had more learners from affluent schools' homes and had the highest achievement. Contrary to these results, (36%) of learners in participating schools from disadvantaged homes and these learners had the lowest achievement (International Association for the Evaluation of Educational Achievement IEA, 2013). These results justify that learners from affluent homes achieved better than learners from disadvantaged homes.

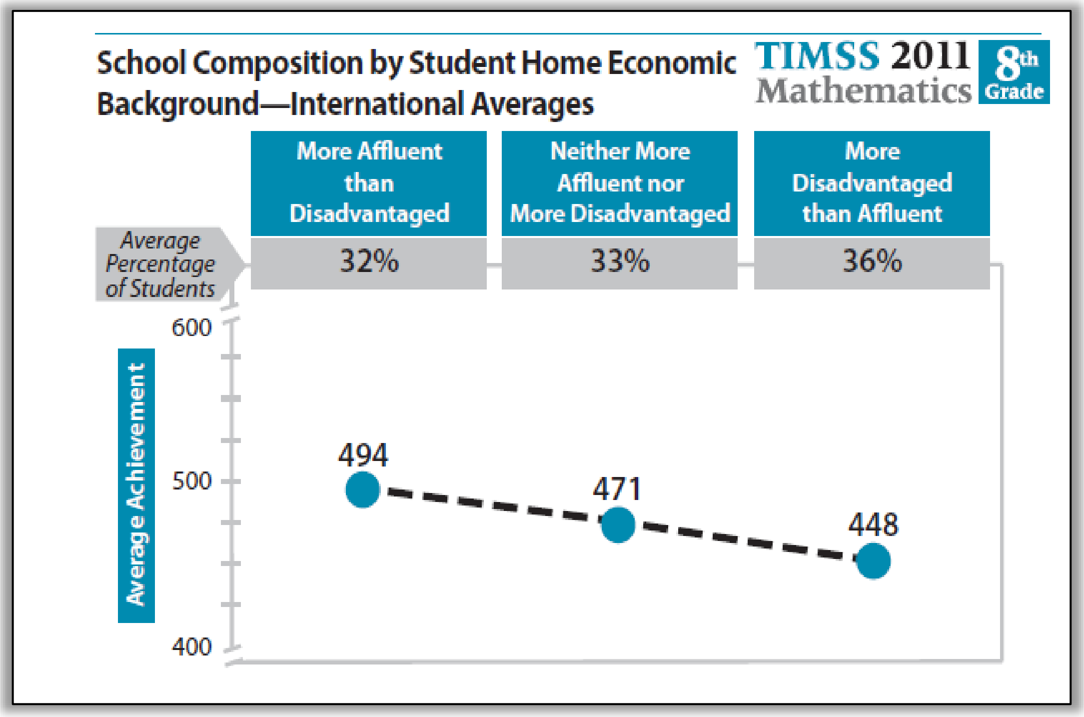


Figure 2.1: International averages for student economic background, (IEA, 2013: 14)

In their study on four countries and their mathematics achievement levels in the TIMSS 2003, Wang et al. (2012) identified two categories of social contextual factors and these are: school climate and social-familial influences. The results of this study show that in South Africa, schools that are better resourced and well-managed showed high mathematics achievement. For parents who have higher educational level, their children had high achievement. These results were consistent in the four countries, South Africa, Singapore, United States of America and Russia.

South Africa's performance in mathematics systemic assessments remain low as compared to these countries (HSRC, 2013; DBE, 2013b). Visible change in South Africa in recent years has been in the curricula, firstly C2005, then the NCS, followed by the RNCS and most recently the CAPS. This has happened irrespective of warnings from researchers on systemic assessment such as Leung (2005) that major changes must not be implemented in South Africa before identifying factors that might have negatively affected teaching and learning.

A comparative study by Reddy (2006) explains that the TIMSS study requires a minimum of 150 schools in each participating country with a minimum of one whole class participating in one school. South Africa had 225 schools that were randomly selected and stratified per province. The results of the 1995 TIMSS showed that South African learners achieved 275 points in the National average out of 800 points in the mathematics test while the International average was 487. The top Province was Western Cape with 381 points (this is below the international average), second and third were Gauteng Province and Northern Cape Province both achieving 318 points and last was Limpopo Province the lowest with 226 points.

The results of review studies by Howie (2003; 2004) on both TIMSS 1995 and 1999 revealed that South African learners performed worse than other participating countries, including developing and developed countries. The study also revealed the following results in South Africa's participation; (1) Learners Afrikaans and English as home languages performed better than African language learners. (2) Learners who believed they were strong mentally in mathematics performed better. (3) Learners in rural schools performed badly as compared to learners from urban schools. (4) Learners with teachers using traditional methods of teaching performed well as compared to teachers who used methods of the reformed curriculum. (5) Learners in large classes performed badly as compared with those in small classes (large classes are those with an average of 50 in a class).

A comparative study by Leung (2005) that focused on possibilities that mathematics achievement can be attributed to classroom practices showed that South African learners performed badly in the 2003 TIMSS study as compared to all participating countries. The study revealed the following findings: (1) In countries that had high learner achievement like East Asia, learners did not enjoy mathematics due to traditional methods of teaching by qualified teachers who argued that their teaching taught clear and simple procedures for pedagogical and efficiency reasons at the expense of rich mathematics concepts. (2) The quantitative results of the 2003 TIMSS video study showed a negative correlation between learner achievement and enjoyment of mathematics in East Asia. (3) The qualitative results showed advanced mathematics learning practices and relevant reasoning without compromise that could see more learners accessing mathematics at the expense of advanced mathematics. (4) South African learners were good in terms of enjoyment and self-confidence in mathematics which had no correlation with the achievement which was low. If change is done the positive attitude must be maintained. In his studies Leung (2014; 2005) warned that the poor achievement by South African learners did not imply a total revamp of the education system but rather a focus on diverse cultural factors that may have negatively affected teaching, learning and learners' achievements.

The observation of this researcher is that such results have been consistent in all studies on international systemic assessment. A steady increase in achievement in South Africa 2011 Grade 9 TIMSS may be a sign of lessons from previous participation. However, this still remains below the TIMSS international average (HSRC, 2013). A need for studies that focus on international systemic assessment from a different lens is eminent if one considers what these assessments were initially aimed at doing. Dunne et al. (2012) pointed out two valid reasons why systemic assessment are not effective in educational reform. The first point is;

> *"A fairly recent expectation is that the results of systemic assessment be made available to parents. This new access to information may be well intentioned, but the form of the information is problematic, precisely because the data from a single and necessarily limited instrument are so fragmentary and imprecise. Systemic assessment is generally not fine-grained enough to report to teachers, or parents, the results of individual learners, as if these single test*

performance results, ascertain from an instrument of about an hour's duration are adequate summative insight into a year's progress in the classroom." (Dunne et al., 2012: 3).

This is an indication that the structure recently used in systemic assessment is deformed and does not necessarily cover all areas for the initial intentions of systemic assessments. The second reason is illustrated below;

"*On the basis of the systemic test score alone, a learner or parent is given a qualitative description that, however well intentioned, is simply arbitrary, invalid and possibly fraudulent, until other evidence justifies the descriptions offered. It is arguable that such descriptions are generally damaging, but especially when test design has not been informed at all by any criteria for item construction and selection that might relate to either the cut-points and the preferred 10% intervals or the objectives chosen."* (Dunne et al., 2012: 3).

In this instance Dunne et al. (2012) clearly show the insufficient information that is contained in the systemic assessment which is aimed at reporting on national or international achievement which is vast, and yet is narrowed to aggregated scores. The need for coherent means of dealing with information in systemic assessment is rather obvious. Kellaghan et al*.* (2009) argue that the disadvantages of using international assessments are: (1) the test is used in more than one country; (2) its content may not be representative of the curriculum of a single country; (3) does not pay enough attention to contexts of individual participating country, and the technology used cannot adapt fully to diverse local cultural and contextual education complexities of all participating countries. This is an indication that although South Africa has participated in international systemic assessment, there is still a need to conduct national systemic assessments that respond to the South African context.

2.2.3 Regional Assessments and South Africa Participation

The (SACMEQ), where South Africa has participated in the past, undertakes research in 15 Southern and Eastern African countries with these initial aims: (1) to widen the scope of education planners and researchers in these countries; (2) provide relevant technical skills for the monitoring of the conditions of their education systems; and, (3) to engage in research that processes evidence-based information that education planners can employ in improving their education systems in their respective countries (Hungi, Makuwa, Ross, Saito, Dolata, Cappelle, Paviot & Vellien, 2010). These studies seem to be a replica of the international systemic assessments.

In their comparative study on contextual factors that affect mathematics performance of Grade 6 learners in the SACMEQ studies, Kotze and Strauss (2006) acknowledge that South African learners come from a wide spectrum of social, political, ethnic, racial, economic and cultural backgrounds and that this factor contributes to achievement in mathematics. However, they pointed out that learner diversity in mathematics achievement can be attributed to factors intrinsic or extrinsic, such as imbalances in socio-economic of schools and learners' exceptional intelligence; mentally impaired, gifted, talented, specific learning disabilities, physical challenges, and chronic health problems, and communicative disorders, emotional and behavioural disorders. Moreover, these diversities have implications for provision, one of the primary goal for national assessments.

Findings from the study by Kotze and Strauss (2006) reveal contextual issues that affect mathematics achievement which have a considerable impact on learning such as parental education, books at home, possessions at home and to general quality of learners' homes. Their findings are: (1) Western Cape Province and Gauteng Province had better general quality of Grade 6 learners' homes, whilst Limpopo and Eastern Cape Province had poor homes. (2) A significant number of Grade 6 learners do their homework in conditions that are not favourable to learning. (3) On average Grade 6 learners' parents lacked basic education. (4) Western Cape and Gauteng Province had the highest number of books in Grade 6 learner's households. (5) Western Cape Province and Gauteng Province had more Grade 6 learners' household possessions that promoted learning such as electronic media, radio, television sets and electricity. These results are consistent with those revealed earlier by Wang et al. (2012), as well as IEA (2013).

A study by Spaull (2010) on the analysis of the South African Grade 6 mathematics SACMEQ III revealed the following findings: (1) On average poor students perform low academically, (2) Wealthy, functional schools have an enhanced ability to produce numerate students. (3) Most poor and dysfunctional schools are unable to produce numerate students. (4) Equal quality education remains a myth. (5) Western Cape and Gauteng had fewer quintile 1 (quintile is poverty index range 1-5; 1 for poor schools and 5 for affluent schools) schools and more quintile 5 schools, whilst Eastern Cape Province, KwaZulu Natal Province and Limpopo Province had more quintile 1 and fewer quintile 5 schools. (6) The mean score for learners' achievement in mathematics was higher in urban schools than in rural schools. Spaull (2010) also make some recommendations to policy makers in South Africa;

> *"1) ensure all learners have access to at least one year of quality preschool-education, 2) provide adequate access to reading textbooks, 3) increase the frequency of homework in poorer schools, 4) improve school management and discipline, 5) improve the ability of teachers to convey their subject knowledge, and 6) learn from other African countries who produce better results with fewer resources. These interventions are likely to improve the performance of primary-school students, particularly so for those from poorer backgrounds" (Spaull, 2010: 26)*

These findings and recommendations from the study on SACMEQ III reveal valuable information on the status of education in South Africa that is well documented in the TIMSS studies where South Africa has participated. The next section revisits the initial aims of systemic assessments. The objective in mind is to look at the ontology of the South African context, lessons learnt from participation in systemic regional and international assessments, as well as what need to inform the processes of Annual National Assessments.

2.2.4 National Systemic Assessments

There are three policy issues that national assessments must address and these are**:** *issues relating to quality*; these refer to quality teaching and learning that is aligned to the implementation of curriculum, what mathematical knowledge learners exhibit as a result of engaging with the type of national assessment; *issues relating to equity*; the national assessment can help to determine how the education system is responding to issues related to gender, socioeconomic diversities, ethnic groups and school governance (public or private); *issues related to provision*; a national assessment can provide evidence on provision of education such as the challenges of curriculum reform, learner retention rate and its effect to teaching and learning, restructuring of the education system, and factors associated with achievement (DFID [S.a.]).

Reflecting on these policy issues that a systemic assessment must address, the studies reviewed points at equity and provision as being the only issues addressed by TIMSS and SACMEQ which South Africa is participating. Issues relating to quality are partially addressed, like in the TIMSS video studies, however, a lot of challenges are encountered here which raises questions on the validity of such results (Leung, 2014). These are reported by Ferrini-Mundy and Schmidt (2005) that the video studies were conducted in three countries out of forty seven countries that participated in the 2003 TIMSS. This is an indication that the video study did not reflect enough on participating countries. The issue on quality remain not addressed by the TIMSS and subsequently by SACMEQ studies. The study by Koretz (2009) highlights the fact that systemic assessment for student achievement is important and in the context of the United States it continues to fail to provide clarity on performance. I observe that the same challenge that America is experiencing is also evident in South Africa which justifies the need for an alternative way of reporting and interpreting results of national systemic assessments.

Now I focus on the South African Annual National Assessments. The principal aim of ANA is to monitor learner attainment at regular intervals, using nationally or provincially defined measuring instruments (DoET, 2002a). According to DoET (2002a) this form of evaluation compares and aggregates information about learner achievements so that it can be used to assist in curriculum development and the evaluation of teaching and learning. As mentioned in the current study, national systemic assessments must address three issues, quality, namely, provision and equity (DFID, [S.a.]). As such, the depth in what ANA is addressing as mentioned in DoET (2002a) does not cover these three issues.

Reports such as DBE (2012a, 2013b & 2014) presented the following results on ANA testing between 2012 and 2014: (1) National average in Grade 9 mathematics of (13%) in 2012, (14%) in 2013 and (11%) in 2014; (2) Western Cape province with the highest achievement in Grade 9 mathematics with (16.7%) in 2012, (17%) in 2013 and (13.9%) in 2014. (3) The second province was Gauteng, (14.7%) in 2012, (15.9%) in 2013 and (12%) in 2014. (4) The worst performer, positioned last, was the Limpopo Province with 8.5 percent in 2012, (9%) in 2013 and (4%) in 2014. (5) Female learners achieved slightly higher results than males nationally and provincially in Grade 9 mathematics in the 2012, 2013 and 2014 respectively. (6) Analysis of schools in terms of the poverty index, called quintiles, indicated that poor schools performed worse than affluent schools (DBE, 2012a & 2013b, 2014). Most of these findings are consistent with those reported in South Africa's participation in all TIMSS and SACMEQ studies.

The results reported above reveal that there is enough information on issues of provision and equity that is at the disposal of the DBE to redress these issues. These results only report on mere aggregated scores between provinces on, gender, poor and affluent schools (DBE, 2014a). These results have been reported in consecutive years and are still consistent. The need to use another lens is pressing and the DBE needs to change its' focus to view other obstacles that the national systemic testing must address to respond to the context and the needs of South Africa. To be concise, it is evident that rich schools perform better and poor schools perform badly. Rich parents provide better education for their children (they send them to rich schools) whilst poor parents cannot afford to provide better education for their children (they send them to poor schools) and rich provinces are ahead in addressing issues related to provision and equity (Dunne et al., 2002; Koretz, 2009). The challenge is now on the issue of quality which has not been adequately addressed by systemic assessments both at international and national level.

A study by Graven and Venkat (2014) was conducted with 54 teachers in 21 township and suburban primary schools in Johannesburg and Grahamstown. Their focus was on the teachers' experiences with ANA. The findings revealed the following; (1) Learners in Grade 3 had a serious problem reading the test questions and interpreting these tests by their teachers compromised their validity. (2) Learners were subjected to write content in ANA without being taught that content. (3) Teaching towards ANA at the expense of quality mathematics learning. (4) The marking guidelines of ANA did not allow multiple solution strategy which was seen to disadvantage learners. Monitoring by district and Provincial Education has revealed a need to empower teachers and subject advisors with knowledge and skills needed to develop quality learning and assessment materials such as tests, assignments and projects (DBE, 2014a). An observation here is that teachers, as well as subject advisors, lack skills that are critical in the challenges facing achievement in ANA testing.

A qualitative textual analysis carried by AMESA (2012) on the ANA 2012 Grade 9 mathematics focused on content coverage and cognitive level requirements. The findings of the analysis were; (1) only word problems, a small portion of the test in question 3 challenged second language speakers and this revealed that mainly the test did not pose serious language problems. (2) In some questions, learners struggled to answer the questions because the formulae were not given. (3) Some content such as transformational geometry, data handling (statistics) and probability; 15.7% of the test which is taught after September were tested which could have disadvantaged some learners. (4) The stakes for the ANA tests were low which made learners not take it seriously. The analysis suggested that obstacles that could have resulted in low achievement were, psycho-genetic, didactical and epistemological obstacles. These results gave direction to the current study hence it took the epistemological perspective.

The report on the 2012 ANA testing, DBE (2012b) identified language and mathematics knowledge and skills as key challenges to learners who participated in the 2012 ANA. Learners' scripts were randomly collected and remarked and it was found that a majority of learners lacked skills and knowledge of the grade in which they were placed. This was an indication that as learners progressed there was lack of systematic progression in their mathematics knowledge and skills learned in consecutive grades. The main challenge here was to locate the origins of the problem. To identify the niche of these challenges, the following conceptual questions raised: Is the problem in the teaching and learning? Or, Is the problem in the ANA tests themselves? This is not mentioned in this report and needs to be researched. It may be important to revisit the aims of national systemic assessment testing and find alternatives that are relevant for the discursive educational context of South Africa.

National systemic assessments evaluate an education system, schools, students and sometimes teachers in a quest to provide evidence on learners' achievement at a particular stage of education in identified curriculum discourses (DFID, [S.a.]). In achieving these aims of national systemic assessment, Kellaghan et al. (2009) argue that the testing takes two forms: (1) census-based in which all schools and learners at targeted population participate, and (2) sample-based which uses only sampled schools. In the DFID [Sa], it is explained that a sample-based has three advantages; the cost is less, turn-around time is faster and higher quality of data due to possible higher supervision. In the context of South Africa, census-based provides information about all school, all districts, all provinces, and the education system in general, (DBE, 2012a). The cost is higher, there is low quality of data due to low quality of supervision, and the turn-around time must be long (DFID, [Sa]).

If poorly performing schools are sanctioned and results published, the assessment become 'high stakes' which may have the following negative effects; neglect of curriculum main areas in favour of the national systemic assessment, teaching that is characterised by rote memorisation and drill for the national systemic assessment at the expense of higher order reasoning, rich mathematics and problem solving skills; teachers focus on low performing learners to make the school results look good, (DFID, [Sa]; Kellaghan et al. 2009). Such challenges are evident in the ANA testing in South Africa. For example, aggregated scores reported in provinces and districts in Grade 9 mathematics show a downward slide, and this could be a reflection of the practices that the DBE performs in the ANA (DBE, 2014a). Such challenges may be related to the quality of the data, turn-around time or quality of supervision which the current study must address.

2.2.5 The Role of Annual National Assessment in South Africa

The DBE ascertains that there are three basic purposes of ANA as follows: firstly, ANA compares aggregated scores yearly to gauge how the education system is performing in phases, in classrooms, districts, provinces and nationally; second, to provide diagnosis in terms of areas that need improvement; and, third, to lay a sound foundation in terms of good practices in teaching and learning (DBE, 2011). To achieve the first purpose, the expectation is that ANA must convey valuable information on how the system is performing (Pournara, Mpofu & Sanders, 2015). Subsequently, to accomplish the second purpose, it is envisaged that ANA should provide data on what learners can or cannot do (Bansilal, 2017). Correspondingly, ANA results need to be used to improve learner performance in the identified areas of weakness (Sibanda, 2017). I explore the details of these artefacts below.

The three years of implementation of ANA has produced valuable data on the performance of learners, in classrooms, within districts and nationally (DBE, 2014). Challenges that have emerged in implementation have resulted in ANA being paused due to the misuse of data by the DBE (SADTU, [Sa]). Correspondingly, poor ANA Grade 9 mathematics results in the year 2012, 2013 and 2014 was used for performance instead for improvement (Kanjee & Moloi, 2016). Schools and provinces that underperformed in ANA were labelled underperforming schools and provinces instead of identifying learner weaknesses in content areas and address them (Spaull, 2016). As such the purpose of ANA has been inflated to performance instead of improvement.

There has been analytical, comparative, longitudinal studies and diagnostic reports on ANA that have revealed how the education system is performing in South Africa through the provision of data on what learners can or cannot do (AMESA, 2012, Kanjee & Moloi, 2016; Long & Wendt, 2017; Modzuka, 2017). An analysis by AMESA (2012) identified three important obstacles that could have caused low performance in the 2012 Grade 9 mathematics ANA, which are; epistemological, psychogenetic and didactical obstacles. First, epistemological obstacles refer to the nature of content and how it is pitched at the correct Grade level (AMESA, 2012). In their study Pournara et al. (2015) used cognitive levels, knowledge, routine procedures, complex procedure and problem solving to analyse the content of the 2012, 2013 and 2014 ANA Grade 9 mathematics. Their findings revealed that the three tests were relatively different and could not produce consistent results. Additionally, some of the content overlapped with Grade 10 which were difficult for learners in Grade 9. Correspondingly, Bansilal (2017) pointed out that the ANA was more difficult than school assessment prepared by teachers. The main cause of this dichotomy could be attributed to teachers' inability to develop assessments (Bansilal, 2012).

The DBE provided exemplar papers during ANA implementation and teachers had the tendency of teaching using the exemplar papers which were relatively different from the ANA papers (Graven & Venkat, 2014). As such, teachers compromised key concepts that were tested by ANA and a requisite for the curriculum, which created misalignment between content that was taught, assessed by ANA and intended by the curricula (Kanjee & Moloi, 2016). Second, the psychogenetic obstacles refer to fragmented cognitive structures in relation to difficulty of test items (AMESA, 2012). The misalignment of what is assessed in the classroom and the ANA that was observed in lower grades had high possibility that it affected the development of learners' cognitive structures (Spaull, 2016). Thirdly, didactical obstacles refer to the quality of teaching that learners obtain as a result of engaging with the present curriculum (AMESA, 2012). The main challenge that were observed during ANA implementation were; discrepancy in marking by teachers, disjointed preparation for ANA, and teachers' lack of knowledge of using data from ANA to improve teaching, learning and assessment practices (Graven & Venkat, 2014; Sibanda, 2017; Spaull, 2016).

2.2.6 Alignment of Systemic Assessments

Content of instruction and cognition are vital in determining what students achieve in learning (Porter, 2002). Policymakers need information on content of instruction and the level at which teaching materials, such as textbooks and assessments, advance the content of instruction (Ndlovu & Mji, 2012). Furthermore, Porter (2002) points out that content of instruction, teaching and learning materials and standards are essential in monitoring curriculum change. Hence, there is a need to align what is taught and what is learned (Nazeem, 2010). The absence of alignment between what is taught in reference to the curriculum and what is tested, may result in teachers focusing their teaching upon what is tested (Graven & Venkat, 2014). Alignment has been described as the extent to which standards and assessments agree. This guides the education system on what learners must learn and achieve (Webb, 2007). The use of Webb's alignment procedures was widely documented in the United States of America. Researchers such as Martone and Sireci (2009) used Webb's alignment to view how content, instruction and assessment were linked. Other researchers such as Porter, Smithson, Blank and Zeidner (2007) have modified the Webb's alignment and justified alignment with a quantitative index.

In an attempt to contribute to the issue of quality, the current study compares the ANA tests with the 2011 TIMSS Grade 8 response items using the Porter's alignment index to verify the content message that the ANA relate to the learning and teaching of mathematics points as a determinant of student achievement (Porter, 2002). Studies on systemic assessment in America have shown that low performance in America in the 1996 and 2003 was caused by insufficient knowledge of content (Ferrini-Mundy & Schmidt, 2005; Suter, 2000). A study by Schmidt and McKnight (1998) revealed that American learners performed poorly in the 1995 TIMSS, especially in Geometry due to insufficient teaching of content. Furthermore, they pointed out that the American curricula lacked coherence and provided students with less rigour. Another important consideration by Ferrini-Mundy and Schmidt (2005) was that it was significant to compare the richness in message that the NAEP, TIMSS and PISA to make sense of some noticeable progress in mathematics education. This is the rationale that made this researcher decide to align the Grade 8 TIMSS response items and the Grade 9 ANA to compare the content message that these studies are posing in the mathematics discourse.

According to Porter (2002), alignment is a worthwhile tool for measuring content message between the content of instruction and content of instructional materials such as content standards, textbooks and achievement tests. Similarly, in their description of the alignment index, Porter et al. (2011) argued that the alignment index defines the contents of intersections of topics and cognitive demands and also assesses the extent at which two documents have the same content message. In addition, Fulmer (2011) describes the alignment index as the measure of the degree to which assessments adequately measure standards to help schools achieve accountability. Finally, Edwards (2010) described alignment as the degree of agreement of standards and assessments which further inform one another and shape students' learning, with emphasis on the quality of the relationship between the two.

Polikoff, Porter and Smithson (2011) support coherence in tested outcomes and the curriculum which can only be visible by calculating the alignment index. They argued in the context of systemic assessments and show how aligning assessments to state standards in America has been essential:

> *'Twenty years on since systemic reform and the state systemic initiatives, instructional coherence remains an important part of standards-based reform in its current incarnation, the No Child Left Behind Act (NCLB). The text of NCLB echoes the framework for systemic reform laid out in the early 1990s, claiming that improving student achievement and ensuring access to a high quality education for all will be accomplished first and foremost through "ensuring that high-quality academic assessments, accountability systems, teacher preparation and training, curriculum, and instructional materials are aligned with challenging State academic standards so that students, teachers, parents, and administrators can measure progress against common expectations for student academic achievement" (No Child Left Behind Act of 2001, 2002, pp. 1439-1440). NCLB mentions alignment dozens of times, specifically focusing on the alignment of assessments with content standards. States are required to show to the Department of Education that their assessments are aligned with the content specified in their standards, and they are also required to assist local entities in identifying curricula that support those standards. Clearly, the coherence of the system remains of utmost importance in the vision of standards-based reform under NCLB"* (Polikoff et al., 2011: 2).

In their study Polikoff et al. (2011) tried to address the following question; Are state standards and assessments aligned with one another? Their findings led them to conclude that the answer was *NO*. However, they acknowledged that this was dependent on how alignment was defined. Such discrepancies in calculating alignment made them conclude as follows: (1) when the definition of alignment was relaxed in the levels of cognitive demand, alignment increased to 0.5 or even higher. (2) Some content is over-tested and other under-tested. (3) State standards and assessment in the specimen of states were not aligned as per intentions. (4) Once states had been made aware of the low alignment, the alignment between state standard and assessments increased. (5) There were many topics that were specified in the state standards that were not tested at all and so in some cases the state standards did not agree with the cognitive levels resulting in misalignment.

In their study, Ndlovu and Mji (2012) calculated the alignment between the TIMSS and the South Africa's Revised National Curriculum Statement (RNCS). Their findings revealed that there was poor alignment between the RNCS and the TIMSS which shows a lack of attention in benchmarking the RNCS and international assessments. They suggested that policymakers must engage urgently to align the TIMSS and the RNCS to make South Africa's curriculum reform relevant. This however, is an ambitious recommendation that the current researcher does not fully agree with in view of the fact that RNCS and TIMSS do not serve the same purpose.

When measuring the alignment index, Porter (2002) aligned of the state assessments and the state standards by calculating the alignment index. Cells of cognitive demands and topics were formed which were described as tools for measuring content and alignment. An example of this is a content matrix shown below.

Table 2.1: Content matrix for cognitive demand and topics (Porter, 2002:4).

Topic	Category of cognitive demand				
	Memorise	Perform procedures	Communicate understanding	Solve nonroutine problems	Conjecture/ generalise/ prove
Multiple-step equations					
Inequalities					
Linear equations					
Lines/slopes and intersect					
Operations on polynomials					
Quadratic equations					

To calculate the alignment index, matrices are formed which match assessments and curriculum in cells and the alignment index is calculated per cell. The items are assigned in the cells and, once finished, alignment is calculated where: Alignment index = 1 - $\frac{\sum |X-Y|}{2}$ where x and y stand for proportion in cell i for documents x and y respectively. The alignment index ranges from 0 to 1, 0-0.50 range depict no alignment to moderate alignment, 0.51 to 1.0 range from moderate to perfect alignment (Ndlovu & Mji, 2012; Porter, 2002; Porter, McMaken, Hwang & Yang, 2011).

2.2.7 Mathematics Cognitive Levels in South Africa

The Subject Assessment Guideline for Mathematics (SAGM) that is outlined in DoE (2007) specifies that core mathematics cognitive levels for mathematics at senior phase are: *knowledge*; *routine procedures*; *complex procedures* and *problem solving*. According to DoET (2007), the cognitive levels are described as follows: *knowledge* involves knowing and using formulae, or algorithms; *routine procedure* involves automative calculations that involve identifying a known procedure. *Complex procedures* are those that are an unfamiliar, indirect route to the solution, and are abstract and involve dealing with complex procedures. Lastly, there is *problem solving*, which involves the solving of non-routine problems, extrapolating from unfamiliar contexts, and the analysis of a problem by breaking it down into manageable parts.

These cognitive levels were adapted as they are from the TIMSS studies in 1995 and subsequent studies (DoET, 2002a). Problems were reported in the use of these cognitive levels in the TIMSS and most recently they were changed to knowing, applying and reasoning (HSRC, 2013; IEA, 2013; Reedy, 2006). However, the SAGM in South Africa's senior phase still remain the same irrespective of challenges reported in the TIMSS and the use of the new cognitive levels in the 2011 mathematics TIMSS (IEA, 2013).

In their critique of the SAGM, Berger et al. (2010) identified two key problematic issues of using the South African cognitive levels. (1) The first is that it conflates the cognitive levels with the type of mathematical activity which leads to challenges that are associated with assessing examinations at varied levels of complexity. (2) The SAGM inhibits the development of essential elements of mathematical reasoning such as conjecturing and justification. They further advise that this may result in problems associated with weak alignment between the curriculum and examination items.

Most studies that associated themselves with cognition in the examination items in South Africa have raised concerns related to examination questions not testing enough higher order skills, and this confirms observations made by Berger et al. (2010). Luneta (2015a) observed that in rural mathematics classrooms in South Africa, assessment tasks set by teachers were far below those standards that involved learners in higher order and critical thinking. Another study by Luneta and Dhlamini (2012) which analysed SMP in Grade 12 examinations pointed out a deficit of tasks that examined adaptive reasoning. Another study by Luneta (2015b) that analysed Grade 12 geometry examination questions revealed that when examinations questions tested level 3 and 4 of van Hiele's hierarchy of geometric thought, most learners were operating at level 2. A study by Ally (2011) in a district called UMgungundlovu in KwaZulu Natal South Africa revealed that, the quality of the promotion of SMP by teachers was very weak and obscure, especially reasoning which was absent in Grade 6 classes. This is a clear indication that the SAGM need to be revised by the DBE because the researcher believes in the epistemology that was once expressed by Schoenfeld and confirmed by Groves (2012) that '*what you test is what you get (WYTIWYG)*' (Groves, 2012: 124).

Problems associated with the SAGM need urgent attention by policy-makers. I will address two issues that are relevant for the current study, first, that the SAGM conflates cognitive levels, and second that the SAGM neglects reasoning, conjecturing and proof. First, the use of the components in the SAGM which are: (1) knowledge, (2) routine procedure, (3) complex procedure and (4) solving problems (Berger et al. 2010) do not allow the varying of cognitive complexity in test items. In the curriculum document (DoET, 2007), the items appear in this order, here called them level 1-4 in this study. The trend in cognitive levels, like the Bloom's Taxonomy (both old and revised), Porter's Taxonomy, Soho's Taxonomy, Stein's Taxonomy and NAEP Taxonomy (just to mention a few) is that level 1 is the lowest and level 4 is the highest cognitive level. If this context of Taxonomies is applied in the SAGM, problem solving assumes level 4, the highest cognitive levels. Some of the elements of problem solving are as follows:

> *"(1) Non-routine unforeseen, (2) interpreting and extrapolating from solutions obtained by solving problems based in unfamiliar contexts, (3) using higher order level cognitive skills and reasoning to solve non-routine problems, (4) breaking down problem into constituent parts, (5) non-routine real context." Berger et al. (2010: 39).*

The five components of problem solving in the SAGM are actually not cognitive levels, except for number 4. The component number 4 may be equated to cognitive level 4 (analysis) in the Bloom's Taxonomy which is not the highest cognitive demand. In total contrary this is not a higher order taxonomy level in Bloom's Taxonomy (Hess, 2006). The Porter's Taxonomy identifies three components of reasoning, conjecture, generalise and prove as the higher order cognitive level. The taxonomy states the following components;

> *"Determine the truth of a math pattern or proposition, write formal or informal proofs, recognise, generate or create patterns, find a mathematical rule to generate a pattern or number sequence, make and investigate math conjectures."* (Berger et al., 2010: 39).

The problem with the Porter's Taxonomy is that is portrays reasoning as the highest cognitive demands. Contrary to this, the NAEP taxonomy includes an element of problem solving, 'creative thought' as higher order cognitive level and the Stein taxonomy includes 'self-regulation' as the higher order cognitive level.

In this current study, it is essential to use the NAEP Taxonomy adapt it to suit the context of the study. The current study focuses on ANA questions and learner responses, and so then it is relevant to focus on the complexity of the question items to understand if it is low complexity, moderate complexity or high complexity (Hess, 2006). The NAEP taxonomy outlines the complexity of test items as follows: (1) low complexity; *recall, recognise concepts, perform mechanical procedures, explain work, but no reasoning* (2) moderate complexity; *flexible choice of procedures, multiple representations and explain procedures with no justification* and (3) high complexity;

reason, plan, judge, analyse, create, reason and justify mathematical statements (Berger et al., 2010).

2.2.8 A Model for Monitoring the Mathematics Education

Drent et al. (2013) outlined a model that helps guide the functionality of an educational system. The model is outlined in Figure 2.2;

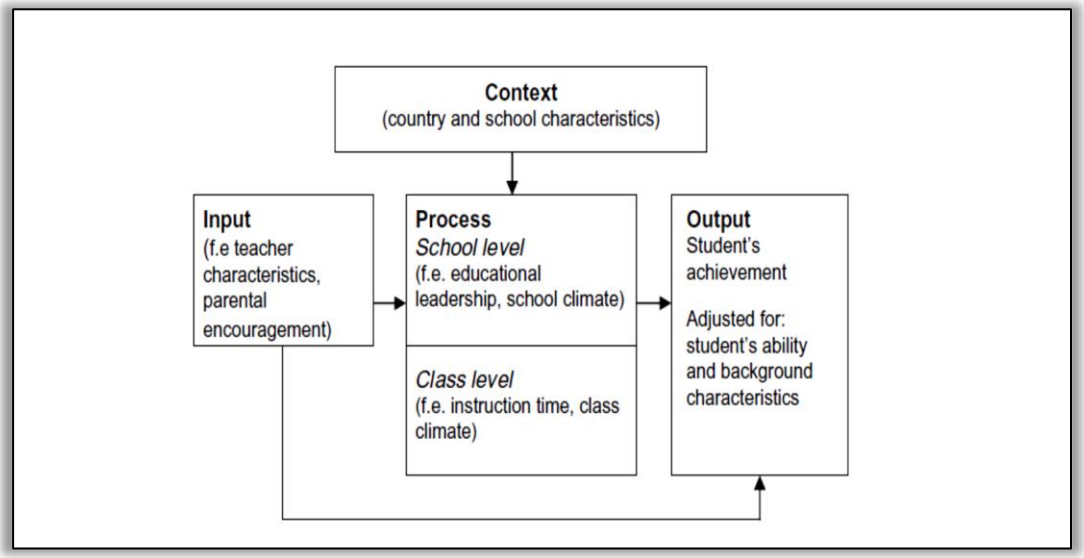


Figure 2.2: Basic system model on the functioning of education (Drent et al. 2013: 201)

Drent et al. (2013) named this model an input-process-output that shows the functionality of a system from teachers and parents to school and classroom factors and how these contribute to students' achievement. An important observation on the model in is that it does not allow for the system to redress. The relationship between the factors is linear and show a one-way process. This model assisted in the development of a model that may be used by policymakers to monitor the quality of mathematics education in South Africa and elsewhere.

This model may be adopted by policy makers during the implementation of the ANAs. It was mentioned earlier that the Minister of Basic Education indicated that the Grade 9 mathematics ANA testing posed a serious challenge to the DBE and researchers must assist (DBE, 2013b). It was further observed that this area of research has a deficit of literature, therefore a model will assist other researchers with areas that need to be researched.

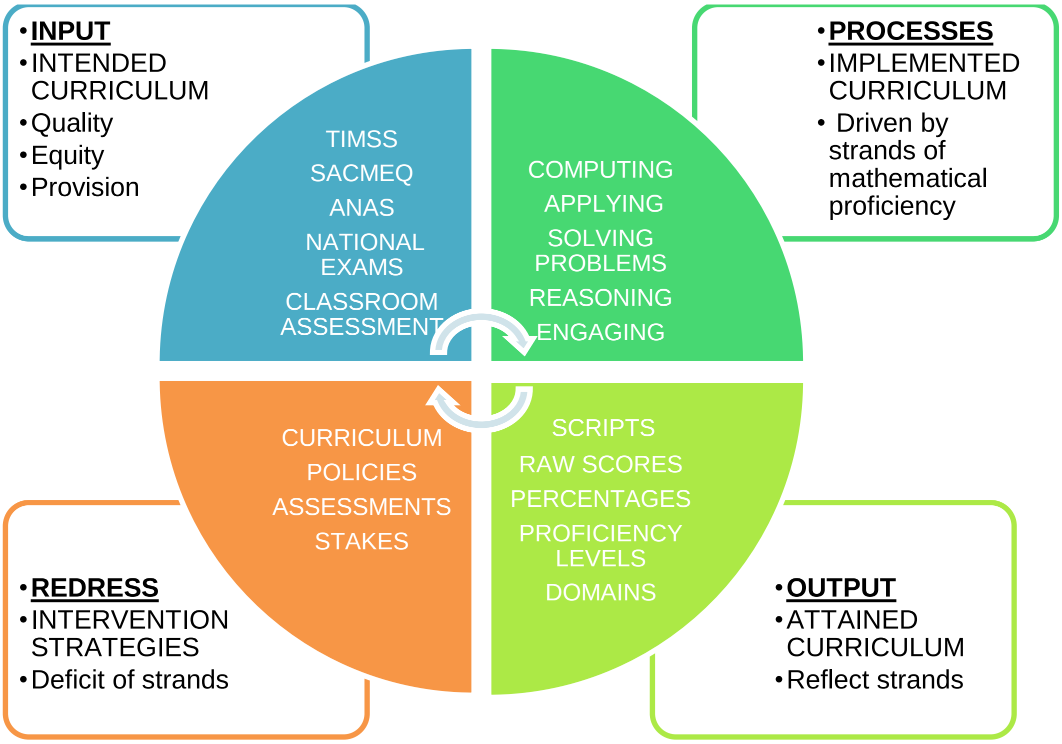


Figure 2.3: Model for monitoring the standard of mathematics education in South Africa as adapted from Drent et al. (2013; 201) with modifications

The model outlined in Figure 2.3 has four parts, input, processes, outputs and redress. The issues in the Cartesian plane are cyclic which shows that there is no specific order in which they may occur and it is a continuous process. These issues are elaborated below.

❖ INPUT

The input stage is characterised by discursive teaching practices such as, teacher characteristics and parental encouragement (Drent et al., 2013). In the context of systemic assessment, both national and international, considerable recent research has reported consistent results in this discourse. Affluent schools have the ability to produce numerate learners unlike poor schools which grapple to produce numerate learners (Spaull, 2010; Wang et al., 2012). Furthermore, rich parents have the ability to send their kids to affluent schools and assist them with school work. Other results were reported in the United States of America that effective teaching and substantial coverage of content with reference to prior knowledge by teachers had a positive effect on learner achievement (Creghan & Creghan, 2013; Koretz, 2009; Tchoshanov, 2011). A study by Moraa and Nyaga (2015) revealed that parental involvement, especially fathers, who frequently monitored their children's work helped their children perform better. They warned that fathers must be directly involved in monitoring their children's' work instead of just providing money. Additionally, Poyraz, Gulten and Bozkurt (2013), as well as Kiwanuka, Van Damme, Noortgate, Anumendem & Namusisi (2015) highlighted that students' success in mathematics is enhanced by a high level educational background of their parents and parental support.

In their study that was seeking to find factors that promote effective schools, Dobbie and Fryer (2013) as well as Angrist, Pathak and Walters (2013) found that class size, teacher certification, per-pupil expenditure, teacher certification and teacher training did not correlate with school effectiveness. Contrary, factors that had a considerable impact on school effectiveness were: increased teaching time, frequent tutoring and high expectations. These results contradict findings from a comparative study on Grade 9 mathematics achievement in Botswana, Kenya and South Africa by Carnoy, Ngware and Oketch (2015) that reported the following: (1) school resources contribute positively to student's learning gains; (2) and an increase in teacher skills and the quality of teaching and learning have a positive impact on the attainment of leaning outcomes by students. Kenyan Grade 6 mathematics students, from both high performing schools and low performing schools had better achievement as compared to South African and Botswana students. This is an indication of differences in the quality of teaching and learning in the three countries. A study by Phillips (2010) reported that having highly qualified teachers does not imply enhanced student achievement, but the findings confirm those by Dobbie and Fryer (2013) that quality teaching and learning is the main determinant of high student achievement.

Reflecting on the model that the current study contributes, it is important when embarking on systemic assessment to be clear on the policy issues that a systemic assessment must address which are mentioned earlier. These are, *equity*, *quality* and *provision* as mentioned by the DFID [Sa]. In the context of systemic assessment, much has been reported on learner achievement, school resources, benchmarking and gender, but this process limits reporting to issues of equity and provision. Less has been done on quality, and this justifies that systemic assessments have not done enough justice to all these three issues. In the model, three issues must direct systemic assessments in the TIMSS and SACMEQ as well as in the ANA. The observation is that these issues must reflect in classroom assessments and National examination which should be the intended curriculum. The study by Ferrer-Esteban (2016) as well as that by Dieltiens and Meny-Gilbert (2012) reported that educational provision is polluted by social segregation in the form of divisions among parents, such as rich parents accessing quality schools. This notion is supported by Spaull (2010) who suggested that in South Africa, equal quality education remains a myth.

- **PROCESSES**

In the model by Drent et al. (2013), processes were limited to the school level which covered the school leadership and school climate. Some researchers agree, especially Halverson, Kelley and Shaw (2014), that when evaluating performance, the focus must not only be on learners and teachers, but also on leaders who must as well improve. A lot has been done on these two issues, school leadership and school climate and most studies reported that schools with effective transformational leadership as well as positive school climate contributed positively to learner achievement (Adu, Akinloye & Adu, 2015; Hofman, Hofman & Gray, 2015; Naicker, 2015; Sosibo, 2015).

The model contributed by the current study suggests that processes, which form the implemented curriculum; must be driven by the SMP. Various studies have shown the importance of using the SMP in mathematics teaching, learning and assessment (Ally & Christiansen, 2013. Groves, 2012; Hauserman & Stick, 2013; Maharaj, Brijlall & Narain, 2015; Schoenfeld, 2007; Suh, 2007). There is no doubt that these strands have been central to mathematics teaching and learning for the past two decades due to their coherence in covering core issues in mathematics education (Kepner &

Huinker, 2012; Kilpatrick et al., 2001). In the model that is contributed by the current study, the intended curriculum, processes are driven by the SMP. This researcher used action verbs that are used in some subsequent studies on the strands with one aim, that is to bring the drive that these processes ought to bring to the discourse of systemic assessment: *computing,* usually called procedural fluency; *applying* usually called conceptual understanding; *solving problems* usually called strategic competence; *reasoning* usually called adaptive reasoning; and *engaging* usually called productive disposition (Kilpatrick et al., 2001; Schoenfeld; 2007; Star; 2005).

Figure 2.4 outlines the pentagonal Singapore Mathematics Curriculum (SMC) which is driven by the strands of mathematical proficiency. At the core of the framework is problem solving which Kilpatrick et al. (2001) call strategic competence. Groves (2012) noted Kilpatrick’s comments on the similarity of the Singapore Mathematics Curriculum in that they both cover the five strands and that mathematics education is not limited to skills and knowledge as earlier pronounced by Skemp (1976). Other strands are at the periphery of the framework, to justify their presence in driving the framework. According to the SMC, problem solving is characterised by learners acquiring mathematical concepts and skills in various contexts (Naroth & Luneta, 2015). As a result, Singapore performance in TIMSS has been consistently high especially in mathematics amongst participating countries (IAE, 2013; Reddy, 2006).

A study by Naroth and Luneta (2015) monitored the implementation of the SMC in South African foundation phase, which is Grade R to three, and found the following results: (1) SMC teacher’s guides were easily accessible to the teacher’s mathematical knowledge. (2) Learners grappled with the English language used in the tasks. (3) The depth and breadth and time spent on topics was useful during learning. (4) The Concept-Pictorial-Abstract approach assisted learners who were second language speakers in understanding mathematics concepts. (5) The SMC was problematic to learners because teachers had limited pedagogical knowledge of problem solving. These findings reveal the nature of the South African curriculum which lacks focus on problem-solving, hence the model contributed by the current study suggests that the implemented curriculum be driven by the SMP. The policymakers may decide on the strand that drive the curriculum at any given instance and such then filters down to the whole education system such as teacher training, professional development a curriculum design and classroom practice. A pertinent example is shown by Koh, Tan and Ng (2012) who justify the notion of thinking schools in Singapore that it emphasised quality over quantity in educational reform.

❖ OUTPUTS

The model by Drent et al. (2013) portrays output as learner achievement that is characterised by students' abilities and other discursive educational factors. Different studies have attributed learner achievement to various factors. A study by Adu and Olaoye (2015) revealed that proficiency in the language of instruction is key to learner's success in algebra. Furthermore, they observed that allowing learners to make discoveries enhances their performance. Other studies pointed improvement in learner achievement to the availability of resources, both human and material (Dobbie & Fryer, 2013; Jantjies & Joy; 2015; Kotze & Strauss, 2006). One case study by Emmett and McGee (2013) pointed student achievement to extrinsic motivation in high stakes large-scale assessments. In the essence of teaching and learning all these results confirms that there are multiple and discursive factors that enhance student achievement. A study by Froneman, Plotz and Vorster (2015) indicated that outcome-based education (OBE) cohort of students lacked procedural algebraic skills of algebra. However, learners had an improved conceptual knowledge of algebra. Such results indicate the need to have the mathematics curriculum being driven by the strands. The deficit of the procedural skills of algebra displayed by the OBE cohort is a serious challenge and Kilpatrick et al. (2001) advocated for intertwined, interconnected and interdependent SMP which are also evident in the model that the current study proposes for the monitoring of the education system.

❖ REDRESS

The positives and negatives that were reported by Graven and Venkatakrishnan (2013) that happened during the administration of the ANA raise a concern about what policy-makers do in response to ANA results. Some of the negatives that were reported pointed to the content coverage, validity of the tests and timing of the ANA testing. These issues were also noted in other countries that administered National systemic assessments. In the study by Klieger (2016) there were different perceptions from teachers and principals on the role of large-scale assessments. Subject teachers viewed them as formative, while head teachers viewed them as summative due to the stakes attached by policy-makers. The initial aim of these assessments was to track students' achievement and monitor how the education system was performing (DFID, [Sa]). However, policy-makers used the large-scale assessments to hold schools accountable for poor performance hence teachers and principals felt threatened (Klieger, 2016). An important observation made by this study is to educate both teachers and principals on the role of large-scale assessments, hence the need to have curriculum in South Africa driven, either by the SMP is evident.

2.3 Theoretical Framework: Strands of Mathematical Proficiency

In this study, the SMP are used as theory and lens to view knowledge, skills and values that the ANA tests examine, as well as those exhibited by learners in their response to the ANA tests. The notion of mathematical proficiency has been widely accepted recently in the mathematics domain and there has been evidence by researchers that it is useful and portrays mathematics as coherent (Groves, 2012; Khairani & Nordin, 2011; Luneta & Dhlamini, 2012; Schneider & Stern, 2010; Schoenfeld, 2007; Suh, 2007; Star, 2005).

In their study, Kepner and Huinker (2012) pointed out four important claims that justify the use of the SMP to strengthen mathematical processes that show learners' behaviours and often lead to a mathematically proficient learner.

> '*Claim 1. Concepts and Procedures: Students can explain and apply mathematical concepts and interpret and carry out mathematical procedures with precision and fluency.*
> *Claim 2. Problem Solving: Students can solve a range of complex well-posed problems in pure and applied mathematics making productive use of knowledge and problem solving strategies.*

Claim 3. Communicating Reasoning: Students can clearly and precisely construct viable arguments to support their own reasoning and to critique reasoning of others.

Claim 4: Modelling and Data Analysis: Students can analyze complex, real world scenarios and can construct and use mathematical models to interpret and solve problems.' (Kepner & Huinker, 2012:29).

These claims are what drives this theoretical framework, that the SMP cover a wide range of knowledge, skills and values that are essential for mathematics learning. Furthermore, such claims elucidate the notion of coherence in mathematics learning advocated by Schmidt and Houang (2012) and Venkat and Adler (2012). In their study, Schmidt and Houang (2012) conducted analyses to determine the rigour and coherence that is outlined by the Common Core State Standards in Mathematics (CCSSM). The results of their analysis revealed that, there is substantial similarity between the CCSSM and the standards of highest achieving nations in the TIMSS and that it was suggested that the CCSSM were coherent and focused. One of the highest performing nations in the mathematics TIMSS, Singapore, was noted by Kilpatrick (2011) to have the same message that are in SMP, see Figure 2.4 below.

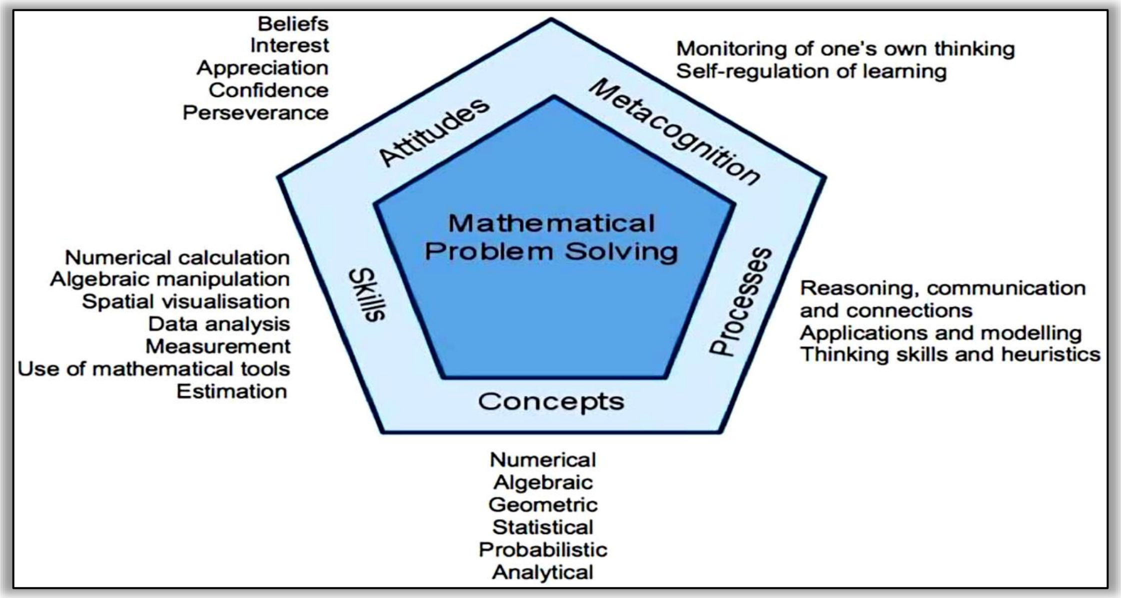


Figure 2.4: The Singapore Mathematics Framework (source: Naroth & Luneta, 2015: 3)

Venkat and Adler (2012) analysed mathematical discourses of concepts to show coherence which involved representing mathematical relations in a way that make them communicate more relevant information. They gave this example; 'Find the turning point of:

" $f(x) = x^2 - 8x + 9$
$= x^2 - 8x + 16 - 16 + 9$
$= (x - 4)^2 - 7$"
Venkat and Adler (2012: 1).

The steps on the function examine skills of converting a general form to a perfect square so that some features of the function, such as the turning point can be explicit. The processes involved, adding and subtracting of 16 on one side of the function, introducing brackets, factorisation to a perfect square, justify coherence in this mathematical discourses which has intertwined proficiencies. As such, a coherent curriculum must provide the following. (1) The contents must be sequenced in such a way that concepts develop from previously developed ideas. (2) A selection of well-structured tasks and representation that enhance coherence.

To view SMP it is important to reflect on constructs mentioned by Schoenfeld (2007) on how mathematically proficient learners project themselves and these are;

> *'What does it mean for a student to be proficient in mathematics? What should students be learning? How can we measure proficiency in mathematics? How can we tell if we are succeeding?'* (Schoenfeld, 2007: 59).

When advocating for SMP, Kilpatrick et al. (2001) pointed out that they are intertwined, interdependent and interconnected. These are: procedural fluency, conceptual understanding, strategic competence, adaptive reasoning, and productive disposition. Most studies that use the notion of SMP, uses two (procedural fluency and conceptual understanding) and some use four without using the last strand, productive disposition (Graven & Stott, 2012; Luneta & Dhlamini, 2012; Schneider & Stern, 2010; Suh, 2007). This is contrary to what Kilpatrick et al. (2001) suggest which is that SMP are intertwined and interconnected, hence the current study uses all five SMP as theory to coherently view knowledge, skills and values examined by the ANA tests and exhibited by learners in their response to the tests. In their study, Dhlamini and Luneta (2016) pointed out that procedural fluency and conceptual understanding form mathematical knowledge, while strategic competence and adaptive reasoning form mathematical skills

For learners to be mathematically proficient they must understand concepts, put meaning to learned procedures, use efficient methods to solve mathematical problems, justify and defend their reasoning (Cragg & Gilmore, 2014; Groves, 2012; Guberman & Leikin, 2013; Stylianides, Stylianides & Shillling-Traina, 2013). However, this study must consider the dangers of focusing on proficiency at the expense of efficiency which results from discursive challenges that learners face in various mathematics classrooms and maybe revealed in baseline assessments (Neal, 2010).

Ally (2011) carried out a study to question the extent at which opportunities for developing mathematical proficiency are made available. The study of four mathematics Grade 6 classes in South Africa, in the KwaZulu-Natal province, Ally looked for empirical evidence in 242 video recordings from 30 lessons. From the lessons, 90 percent were opportunities to develop procedural fluency, 17 percent for conceptual understanding, close to 2 percent for strategic competence, 8 percent for adaptive reasoning and 20 percent for productive disposition. The high frequency in procedural fluency was due to the teaching of procedures and in productive disposition was due to the inclusion of everyday examples (Ally, 2011:90).

Luneta and Dhlamini (2012) conducted a study on a question by question analysis of ninety Grade 12 scripts for learners' responses to four topics on final examination questions. Their focus was to view the SMP that learners exhibited in response to the examination questions. The results showed that learners were proficient in only one topic, analytical geometry, irrespective of the fact that these examination questions did not demand the higher order reasoning that is a pre-requisite for natural sciences courses at tertiary level such as engineering.

Another study by Sanni (2009) identified forty six verbs in the South Africa's Grade 7 Revised National Curriculum Statements (RNCS) assessment standards that were categorised into levels of mathematical practices and SMP. The focus of his study was to discuss the extent to which the South African Mathematics Curriculum Statement promotes the development of SMP and mathematics practices. The results of the textual analysis revealed that most of the verbs used in the assessment standards were ideally suitable for each of the five SMP and the three mathematics practices. Most of the assessment standards required learners to exhibit conceptual knowledge and adaptive reasoning. Some of the SMP such as adaptive reasoning, strategic competence and productive disposition were not promoted by the assessment standards. However, this study did not focus on what learners exhibited as a result of engaging with the assessment standards. This could have been done by analysing learner responses to some tests that were set using these assessment standards. Moreover, Sanni (2009) pointed out that some tenets are provided by the curriculum and not implemented by teachers as they may be difficult to implement.

The significance of using SMP as a theoretical framework in the current study is justified by their coherence and their effect in the teaching and learning of mathematics. The following sections focus on making SMP compatible for document analysis, a problem reported by Graven and Stott (2012).

2.3.1 Procedural Fluency

Procedural fluency is referred to as the learners' ability to carry out exact, supple and correct procedures (Kilpatrick et al., 2001). In addition, Star (2005) describe procedural fluency as computing, which involves performing mathematical procedures such as addition, subtraction, multiplication and division of numbers flexibly, accurately, efficiently and appropriately. Groves (2012) advises that computational fluency is more than just that speed and accuracy which have previously been regarded as essential in performing procedures. Kilpatrick et al. (2001) stressed that procedural fluency also involves knowing *'when and how to use them appropriately, and skill in performing them flexibly, accurately, and efficiently'* (Kilpatrick et al., 2001: 121). Furthermore, Groves (2012) suggests that, while accuracy and efficiency are aligned to computations of algorithms it is essential that learners exhibit flexibly and fluently their ability to perform mental computations. Flowers, Kline and Rubenstein (2003) explained that their experience in teaching for computational fluency implies that classroom activities must be characterised by reasoning, problem solving, communicating mathematically and computations that are rich and meaningful. Such learning is characterised by students constructing meaning through the use of what they know to solve what they do not know, called routine and non-routine problems (Naroth & Luneta, 2015). Russell (2000) suggested that computationally fluent procedures are accurate, efficient, flexible and consolidated. The only difference that Russell brings to the definition of procedural fluency is the notion of consolidation. Flowers et al. (2003) pointed out that there are challenges in promoting computational fluency where teachers are also not themselves computationally fluent which results from how they were taught or an over-dependence on one, meaning pencil and paper computations that rely on a calculator.

Flowers et al. (2003) illustrate with examples how teachers master computational fluency using consolidated computations using subtraction, as the authors believed that subtraction offers affluent opportunities for computational fluency. The question read as follows: $'How\ much\ is\ 1006 - 98?$ (Flowers et al., 2003: 332). Teachers shared their reasoning first in small groups and after convincing their peers then they shared their computations with the whole class. In one example the teacher subtracted as follows;

"$1006 - 98 = ?$
$98 + ? = 1006$
$98 + 2 = 100$
$100 + 900 = 1000$
$1000 + 6 = 1006$
$2 + 900 + 6 = 908$" p332

According to Flowers et al. (2003) the teacher who carried out this computation justified it as follows;

Teacher A: 'It works because in the beginning you took 6 away from 1006 and added onto 98. When you subtract 100 from 1000, you get 900. Now you have to compensate from what you did in the beginning. You add 6 back on because you subtracted it from 1006 to get 1000. Then *you add 2 on because you added it to 98 to get 100. The reason you add 2 and not subtract it is because in subtraction, instead of subtracting in the in the end for compensating, you need add it on.'* (Flowers, at al., 2003; 331).

Such computations justify competence in the algorithm followed and the confidence displayed by teacher A shows mastery as opposed to pencil and paper computations that are calculator based.

An illustrative example given by Groves (2012) to explain procedural fluency in a Grade two lesson that was video recorded working on word problems that allowed multiple solution strategy while collaborating on their solutions. One of the problem was: *'Sarah's team scored 49 runs and Jason's team scored 63 runs. How many more runs did Jason's team score?'* (Groves, 2012: 130). From transcribed data, two girls start by counting from 49, 50, 51… until they suggested that they do not have more fingers to count and used a calculator to record the remaining numbers. Another boy said he worked out 49 plus 11 and got 60 then added another 3 to get 14. This problem is not like a written computation that guarantees a single answer. Such mental computations allowed the learners to think make sense of the problem.

Ebdon, Coakley and Legnard (2003) give illustrative examples of strategies that may enhance computational fluency and efficiency. One example shows Second Grade learners adding '84 + 16' which results in four strategies; (1) math map; (2) the storing strategy; (3) the giving strategy; and (4) the tens-party strategy as shown in Figure 2.5. In the math map strategy, the second grader, student A, explained how in the brain there are road maps which look like arrows which connect tens and ones and this enhanced the student's ability to join correct numbers. For the storing strategy, student B, decided some constituents of the number, place on one side of the brain, deal with the remaining numbers and then later retrieve those in the storage and compute (see figure 2.5). In the tens- party strategy student C explained that they work out tens and units separately then add at the end. The giving strategy, student E explained that they give numbers to make all the numbers tens then it became easier to add numbers that only had tens. This is one example that promotes multiple solution strategy to solve a single problem.

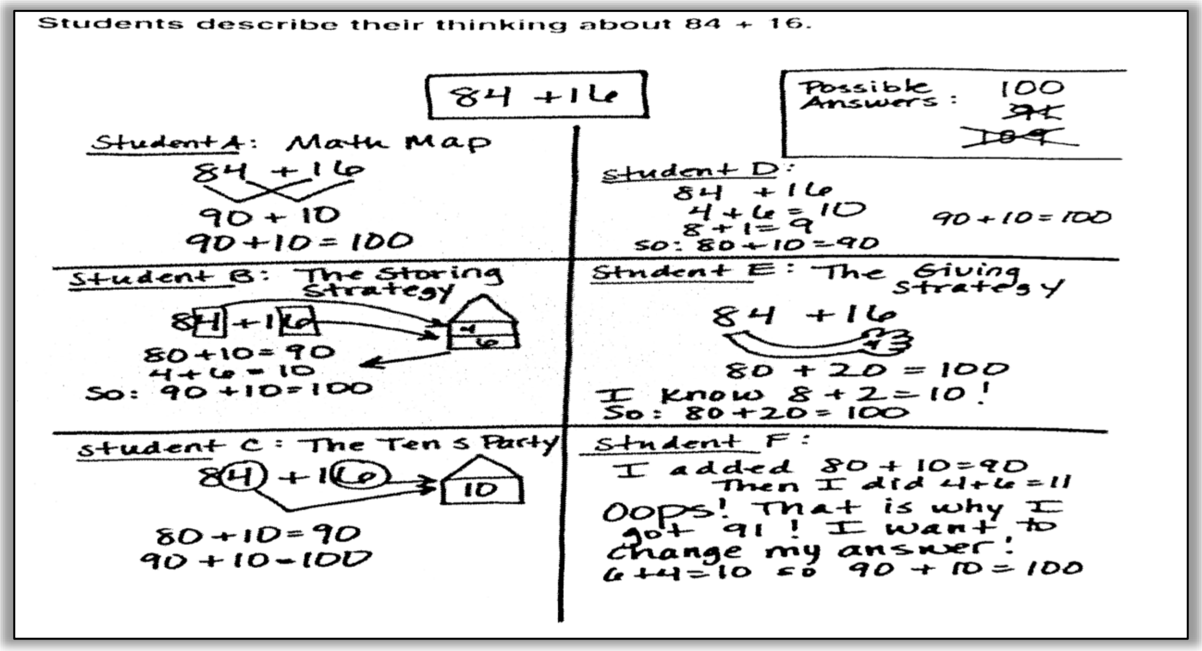


Figure 2.5: Reasoning strategies by second grade students (Ebdon et al., 2003: 488)

A study by Graven and Stott (2012) used the SMP of Kilpatrick et al. (2001) as a framework to analyse written responses by learners towards test items. These test items examined the speed and accuracy of mathematical operations as well as mathematical identities and dispositions on procedural fluency. In their attempt to view these artefacts in the learners' responses, the authors admitted that it was not possible to use the SMP of Kilpatrick et al. (2001) as they are to analyse documents. Therefore, they developed what they called a spectrum of efficiency and fluency in a bid to adapt procedural fluency to help assess learner's proficiency in their response to the test items. An example of how they used their spectrum is shown below;

'By way of an example, for question part 'a) add 10 to 92' we found that the learners answered the question using these possible methods:

• Less efficient ways would be to: draw the sum on paper (e.g. concrete representation of the number by drawing 10 and 92 dots and recounting), or count on 92 from 10 using fingers (or counters).

• A slightly more efficient way may be to use a standard vertical algorithm such as: 92 +10.

• More efficient 7 ways would be to count on 10 from 92 in ones or by using fingers.

• The most efficient method would be to mentally add 10 to 90 and then add 2 (or 10 to 92 with knowledge of the pattern of adding tens) thereby involving conceptual understanding of commutativity, patterns and place value.
Similarly, it can be seen, that the less efficient methods would become even less efficient ways of answering Task 10 part b and d ' (Graven & Stott, 2012; 152)

The findings from this study revealed that the less efficient methods were not appropriate when dealing with procedures of bigger numbers and more efficient methods were efficient when dealing with bigger numbers. Graven and Stott (2012) pointed out that much work needs to be done to adapt the SMP of Kilpatrick et al. (2001) to assess learner progression. What I identify as problematic in this study is that the codes used by Graven and Stott (2012) did not clearly provide a coherent spectrum of all the SMP. Their study was only limited to one SMP, procedural fluency. Graven and Stott (2012) ascertained that the spectrum was developed by using definitions of the SMP. For an example, Kilpatrick et al. (2001) described procedural fluency as follows:

"Procedural fluency refers to 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: 121).

An observation here is that Graven and Stott (2012) used the key words in their development of their spectrum of procedural fluency and they used the words, 'flexible' and efficient' and could not use the word 'accurately'. In the current study, the researcher used all key words that define the SMP to unpack and adapt all the five strands to be compatible for textual analysis. There is a need to use all the strands because Kilpatrick et al. (2001) suggested that the SMP are intertwined, interlinked and interrelated.

The work done by Graven and Stott (2012) suggests that this strand, procedural fluency, is composed of the four concepts that are in the definition by Kilpatrick et al. (2001) which are, flexibility, accuracy, efficiency, and appropriation. The first part of a procedure, *flexibility* refers to knowing various procedures and relative efficiencies of the procedures (Star, 2004). The second part*, accuracy* often refers to the correct use of signs to carry out basic operations that involve addition, subtraction, multiplication and division to reach the required answer consistently (Bass, 2003). The third part, *efficiency,* is often visible in learners who conduct procedures certainly, and use intermediate outcomes to execute the problem (NCTM, 2000). The last part, *appropriation,* refers to learners who are conscious of the right time to apply a procedure (Sherin & Fuson, 2005).

Figure 2.6, is an example from Grade 9 mathematics ANA exemplar questions that are aimed at assisting in preparation for the ANA test. This procedure demands that learners have knowledge of definition of these concepts and does not test any computational fluency. For learners to respond to these questions, they need knowledge of the number concept and must then tick the appropriate box. The question does not allow learners to exhibit procedural fluency in these concepts.

	NATURAL NUMBER	WHOLE NUMBER	INTEGER	RATIONAL	IRRATIONAL	REAL
e.g. $3\frac{1}{2}$				√		√
$\frac{7}{15}$						
$\sqrt{2\frac{1}{8}}$						
$\sqrt[3]{0{,}081}$						
2π						
$-\sqrt{16}$						
0,528						
2,6						
$\frac{6}{2}$						

Figure 2.6: Question 1.8 Grade 9 Mathematics Exemplar (DBE, 2012d)

In adapting procedural fluency to be compatible for document analysis, I reflect on work done by Bass (2003) on computational and algorithm fluency. Bass (2003) describes a computation as a specialised form of problem solving that involves fluency in the four basic operations (+, -, ×, ÷). According to Bass (2003), fluency indicates the nature of an algorithm to be efficient, accurate and generalisable. These algorithms are described here as a general solution for all related problems which has a concise specified series of steps that lead to a coherent solution of a class of problems. Bass (2003) proposed qualities of an algorithm which involves the following processes;

- '**Accuracy or reliability:** *the algorithm should always produce the correct answer.*
- **Generality:** *the algorithm applies to all instances of the problem, or class.*
- **Efficiency (or complexity):** *This refers to whether the cost (the time, effort, difficulty, or resources) of executing the algorithm is reasonably low compared to the input size of the problem.*
- **Ease of accurate use:** *use (versus error proneness). The algorithm can be used reasonably and does not lead to high frequency of error in use.*
- **Transparency (versus opacity):** *what steps of the algorithm mean mathematically and why they advance us towards the problem solution, is clearly visible.*' (Bass, 2003: 32).

Bass (2003) explains that for a procedure to qualify as an algorithm, it must have the characteristics of accuracy and generality. The other algorithm checks such as ease of accurate use, efficiency and generality are just as important in ensuring that student perform procedures fluently. I consider a subtraction example by Bass (2003); 36 – 19 =17. In trying to solve this problem, learners in elementary mathematics performed some procedural errors. Some learners did the following: 36 – 19 = 23, whilst common correct solution is performed in this manner;

$$\begin{array}{r} {}^{2}3\,{}^{1}6 \\ -\ \underline{1\ 9} \\ \underline{1\ 7} \end{array}$$

Bass (2003) explained this algorithm in the context of the American classroom, the borrowing method and common practice by teachers explaining this algorithm is; (1) you look at the number on the units in the upper number and if it is smaller than the one in the second number, (2) you borrow one from three and remain with two. (3) Then it is sixteen minus nine to get seven. (4) Then it's two minus one to get one. Thus, an answer of seventeen is obtained.

Checking algorithm usefulness: (1) accuracy (2) generality (3) efficiency (4) ease of accurate use (5) transparency, it can be seen that learners have the tendency of subtracting as; 36 – 19 = 23. It shows that the algorithm fails to conform to number (2) then it does not fulfil the entire algorithm check. Bass (2003) suggests that if an algorithm fails 1 & 2 ‘algorithm usefulness’ then an alternative algorithm must be found that will be consistent in all the checks. For example; teachers must use 36 – 19 =? (3x + 6) – (x + 9) = (2x – 3), x is the number of tens in both numbers being subtracted. Therefore 36 – 19 will be calculated as; (30 + 6) – (10 + 9) = (20 – 3) = 17. Testing this algorithm with other numbers; 53 – 38 = ? yields ; 53 – 38 = (5x + 3) – (3x + 8) = (2x – 5) = (20 – 5) =15. This is proof that the second procedure is consistent with accuracy and generality therefore it qualifies as an algorithm. However, this algorithm uses symbols. This is an abstraction, and learners who have not developed proper cognitive functions may not perform this algorithm (Breen & O’Shea, 2010). This means that learners need to be proficient in this algorithm to fulfil efficiency, ease of accurate use and transparency.

These processes coherently address qualities of an algorithm and they bring meaning to this SMP, procedural fluency. Bass (2003) explains the use of these qualities further below;

> *‘To qualify as an algorithm, a procedure must be characterized by accuracy and generality. Ease of accurate use, efficiency, and transparency also are desirable qualities, although they often are in competition with one another. An algorithm that will be used to program a machine must be efficient, to achieve computational speed, but does not have to be transparent. If humans will learn and use the algorithm, however, transparency and ease of accurate use are important’* (Bass, 2003: 324).

These processes outlined by Bass (2003) are relevant to revealing the nature of routine and non-routine problems. Reformed mathematics learning demands competency in the functionality and coherence of mathematics which implies that learners need to create quality connections and representations in the conceptual and social aspect of mathematics (Breen & O’Shea, 2010; Chapman, 2013; Greenes, 2014; Mwakapenda, 2008).

2.3.2 Conceptual Understanding

Conceptual understanding is often referred to as the learners' ability to connect mathematical ideas, processes and commonalities and it is evident when learners begin to connect new ideas to old ideas, and to reconstruct new ideas once they are forgotten, (Kilpatrick et al., 2001). As such, conceptual understanding is *'knowing how and how'* which suggests that this is an extension of procedural fluency (Rittle-Johnson & Schneider, 2015). Furthermore, a learner with conceptual knowledge can easily integrate mathematics ideas, and their knowledge is organised in a coherent whole which enables them to recall and reconstruct knowledge once forgotten. Additionally, Star (2005) explained conceptual understanding as 'understanding' the underlying principles of mathematical symbols, diagrams and procedures which is done by comprehending them with appropriation. Thus, learners are empowered to think generatively with content, select relevant procedures and strategies for steps of problem solving (Davis, 2005; Kazemi & Stipek, 2001; Richland, Stigler & Holyoak, 2012). The DBE (2011) in one of its specific aims for mathematics education is that learners must develop '*deep conceptual understandings in order to make sense of mathematics*' (DBE, 2011: 8). However there seems to be a challenge when studies such as that of Khashan (2014) reported that teachers have only average distorted conceptual and procedural knowledge of rational numbers. Such may have a serious consequence for the knowledge that learners exhibit to assessments such as ANA.

I indicated earlier in the preamble on SMP that conceptual understanding and procedural fluency form mathematical knowledge (Dhlamini & Luneta, 2016) I noted some extensive history in these forms of knowledge. Skemp (1976) introduced mathematical knowledge using two concepts, instrumental and relational understanding. Firstly, instrumental understanding referred to the functional grasp of mathematical rules which process the solution without giving reasons and contrary relational understanding was equated to knowledge of both rules and its underlying reasons for their use. I note that most researchers in mathematics education often and incorrectly equate instrumental understanding to procedural knowledge and relational understanding to conceptual understanding. Most specifically, Skemp's view is superficial and reduced to rules, whilst 'fluency' as used by Kilpatrick et al.

(2001) protrudes beyond rules. In fact, fluency refers to the ownership of efficient, accurate and generalisable methods (NCTM, 2000) called algorithms that are used to compute number relations. Skemp's work on understanding was advanced by the introduction of two types of knowledge, procedural and conceptual knowledge. Knowledge of automation of operators and the setting of their use to achieve certain goals is often called procedural knowledge (Davis, 2005; Hiebert, 1986). Consequently, automated knowledge is restricted to insufficient cognitive resources and may not be converted to higher cognitive thinking processes (Schneider & Stern, 2010). Conceptual knowledge is abstract and general knowledge coupled with interrelations of mathematical processes, procedures and concepts (Hiebert & Lefevre, 1986; Tevfik & Ahmet, 2003). Subsequently, this form of mathematical knowledge is supple and enhances the transfer of learning and understanding of algorithms and procedures (Rittle-Johnson & Alibali, 1999; Star & Stylianides, 2013). Accordingly, the suppleness of conceptual knowledge postulates a generic grasp of both the abstract and concrete knowledges (McCormick, 1997).

Vigorous deliberations took centre stage (Hallett, Nunes, Bryant & Thorpe, 2012; McCormick, 1997; Rittle-Johnson & Alibali, 1999; Schneider & Stern, 2010) on the sequencing and blending of procedural and conceptual knowledge, some assertions were 'procedures first then concepts' and other assertions were vice versa. Subsequently, Rittle-Johnson and Schneider, (2015) lately suggested four views on the connection between conceptual knowledge and procedural knowledge and holding a particular view they hypothesise the connection of conceptual and procedural knowledge as follows: (1) *Concepts first view*, learners often obtain concepts first (Tevfik & Ahmet, 2003) which they use to comprehend procedures through frequent practice. (2) *Procedures first view*, means that learners grasp procedures exploratively, (Siegler & Stern, 1998) which is followed by generalisations that lead to conceptual knowledge. (3) *Inactivation view* posits that conceptual and procedural knowledge progress parallel to each other (Haapasalo & Kadijevich, 2000). (4) *The iterative view* elucidates that the frequent increase in procedural knowledge concurrently escalates conceptual knowledge (Baroody, Feil & Johnson, 2007) and vice versa. Most recently, studies (Rittle-Johnson & Schneider, 2015) have revealed worthwhile quantitative results that validate the iterative view as being effective in the acquisition of coherent mathematics knowledge. Hence, I pronounce my subscription to the iterative view of knowledge.

Furthermore, the iterative view is explained by Star (2005) in the concept of re-conceptualising of procedural and conceptual knowledge, as quality knowledge. A relevant view is the critique argued by Baroody et al. (2007), that procedural knowledge is either superficial or deep. In contrast quality mathematical knowledge is often associated with connections (Hiebert & Lefevre, 1986) which proposes that quality knowledge is always conceptual (Star, 2005). Ironically, my adherence to the iterative view clashes with this point of view and proposes the existence of quality in both procedural knowledge and conceptual knowledge. In addition, Star (2004) and Schneider, Rittle-Johnson and Star (2011) posed conflicting views on learners' prior conceptual or procedural knowledge. Some assertions were that learners with limited procedural knowledge enhances conceptual development and vice versa. As such, much debate is still essential to confirm this view and other numerous assertions that remain untested on the debate on conceptual and procedural knowledge.

Since this study is informed by SMP, Kilpatrick et al. (2001) advise that SMP are interconnected, inseparable and interwoven. Subsequently, my view is that using the iterative view posits a useful simultaneous grasp of both procedural and conceptual mathematical knowledge (Cragg & Gilmore, 2014; Khashan, 2014) and when in the perspective of SMP, it enhances coherent mathematics learning. The incorporation of procedural and conceptual knowledge into SMP by Kilpatrick, et al. (2001) hinges on five coherently combined categories, three of which are, strategic competence, adaptive reasoning and productive disposition that are used in this study. These are discussed in the next sections. Next, I explore studies on procedural and conceptual knowledge. Below I explore studies in this discourse.

Tevfik and Ahmet (2003) conducted a study on the comparison of the equality between the levels of conceptual and procedural knowledge of 182 students in the Grade 10 attending one school. Data was collected from two tests, 1st test on procedural knowledge given to 42 students and the 2nd test on conceptual knowledge given to 40 students. A total of five students were interviewed about their acquisition of procedural and conceptual knowledge. A t-test was used to analyse the data from the tests. The results revealed that there is a significant difference between procedural and conceptual learning. The results also revealed that, even though both procedural and conceptual knowledge were both important in mathematics learning, procedural learning was more important. According to this study, the emphasis was on performing correct computations at the expense of understanding (Schoenfeld, 2007).

A study by Niemi (1996) focused on students' ability to represent their conceptual knowledge in a variety of task contexts and formats while comparing their performance across these tasks. A total of 540 students in the fifth Grade across 22 schools from the Washington State participated in the study. Various tasks were used to collect data. Data was analysed structurally to view mathematical understanding in terms of conceptual fields and schemas which had a series of relations to these elements: symbols, concepts and operations. Results from this study were as follows: (1) representational knowledge, justification and explanation tasks provide worthwhile information on students' knowledge of mathematical representations that assessed conceptual knowledge. (2) Students need to be exposed to a variety of representations and the use of a single representation cannot lead to learner's mastery of other representations. (3) Assessment items used to test understanding should include incorrect and misunderstood representations. Such results justify the importance of conceptual understanding over procedural fluency contrary to the study by Tevfik and Ahmet (2003).

Kaulinge (2012) carried a study that focused on the analysis of two tasks on the relationship of procedural and conceptual knowledge of Grade 3 learners from a numeracy workbook. Results from tasks analysis revealed that: (1) Learners required both conceptual and procedural knowledge to get correct responses in both tasks. (2) Conceptual and procedural knowledge co-exist and cannot exist in isolation. (3) Conceptual and procedural knowledge are equally essential for learner's success in mathematics learning. Emphasis here is a balance in conceptual and procedural knowledge is contrary to results by Niemi (1996) as well as those of Tevfik and Ahmet (2003). However, Hallett et al. (2012) pointed out that there are individual learner differences of Grade 9 mathematics learners on their conceptual and procedural knowledge of fractions. This may be attributed to the rate at which learners grasp the different conceptual and procedural knowledge of fractions. Gardee and Brodie (2015) suggested that if teachers probe and embrace errors made by learners, this may enhance the conceptual knowledge of the learners.

To adapt conceptual understanding for document analysis, the current researcher reviewed how Kilpatrick et al. (2001) defined conceptual understanding:

> *'Conceptual understanding refers to an integrated and functional grasp of mathematical ideas. Students with conceptual understanding know more than isolated facts and methods. They understand why a mathematical idea is important and the kinds of contexts in which is it useful. They have organized their knowledge into a coherent whole, which enables them to learn new ideas by connecting those ideas to what they already know. Conceptual understanding also supports retention. Because facts and methods learned with understanding are connected, they are easier to remember and use, and they can be reconstructed when forgotten'* (Kilpatrick et al., 2001:118).

Looking closely to this definition, there are some key constructs that explore the meaning of conceptual understanding; *integration* of or *connecting*, *known isolated facts* and *methods*, *mathematical ideas* and *coherence*. Earlier the notion of coherence is explained as the sequencing of content in such a way that concepts develop from previously developed ideas, and a selection of well-structured tasks and representation (Goldman, 2006; Schmidt & Houang, 2012; Venkat & Adler, 2012; Watanabe, 2007). In trying to justify the importance of connections, Mwakapenda (2008) navigates into the curriculum statements and identify how mathematics is perceived in the South African curriculum as shown below:

> *'Mathematics enables creative and logical reasoning about problems in the physical and social world and in the context of mathematics itself. It is a distinctly human activity practised by all cultures. Knowledge in the mathematical sciences is constructed through the establishment of descriptive, numerical and symbolic relationships. Mathematics is based on observing patterns; with rigorous logical thinking, this leads to theories of abstract relations. Mathematical problem solving enables us to understand the world and make use of that understanding in our daily lives. Mathematics is developed and contested over time through both language and symbols by social interaction and is thus open to change (DoE, 2003:9).*

Mwakapenda (2008) pointed out that making connections involves the following: (1) Begin with everyday practices that make sense to the learner while identifying the development of sound mathematical notions. (2) Focus on what is taught then find a relevant everyday practice of student that links well with the development of the concept. Such connection of procedures enhances conceptual understanding that is coupled with sense making (Lee, 2012; Mhlolo, Venkat & Schafer, 2012).

In considering the definition by the DoET, I reflect on work done by Russell (2008) with teacher colleagues. They connected concepts and themes in such a way that their students made meaning with shapes such as, square, rectangle, triangle, circle, quadrilaterals and polygons. These shapes were put in wall charts together and mapped their properties in trying to connect them. Students made sense of this activity and their learning of concepts and facts was no longer isolated. Such learning activities show coherence and logic in the connections made on the content that need to be learned. Such activities made enabled this researcher in the current study to identify key features of connections that are relevant, and these are; (1) conceptual connections, (2) everyday connections, (3) algorithmic connections, and (4) symbolic connections.

Communication is vital in assisting students to make meaningful connections in the physical, pictorial, graphic verbal and mental conceptions of mathematical ideas (Capps & Pickreign, 1993). These authors proposed that clear and correct mathematical words would help students connect informal, intuitive notions as well as abstract symbolism of mathematics. Furthermore, they also pointed out that in everyday connections, mathematics operations are not explicit in problem solving. They gave this example:

> *'Jesse and his dad are making cookies. The recipe calls for two eggs for a single batch of cookies. If they plan to make a triple batch, how many eggs will they need?''* (p11).

In response to this question, most student gave an answer, 3×2=6. This puzzled Capps and Pickreign (1993) on how the learners came with the algorithm, and through semi-structured interviews it was revealed that the secret is in the connective words, such as triple which means ‘3’ and the context of the problem leads to multiplication. This shows how language in a task connects symbols, in this instance, (2 & 3), concepts (multiplication) as well as the everyday (cookies, recipe, eggs, and triple batch).

Kalchman (2011) warned that, due to the high stakes attached to assessments, there is a growing need to prepare test items that use routine practices of learners without compromising the depth and breadth of the mathematics curriculum. The argument was that connecting mathematics to everyday life enhanced the student’s ability to communicate, recognise as well as apply mathematical ideas with competence in functional mathematical situations. Kalchman (2011) outlined the impact of fifth grade students’ authentic experience they experienced outside the school setting that required them to use mathematics. The students showed how the situations demanded mathematics and their approach to solve the problems. Reflecting on the case of Hannah who was developing fluency in addition, subtraction, multiplication and division of whole numbers using the following example: Hannah wanted to know how much her babysitter earned per hour if she earned $75 for working six hours. Hannah divided $75 by six and got twelve remainder three. She was confused about the remainder and decided to say that the babysitter earned $12, 03 per hour, meaning that the remainder of three was equal to three cents. Hannah did not realise the results of division and her teacher had to intervene and show her how to apply division to get the required answer.

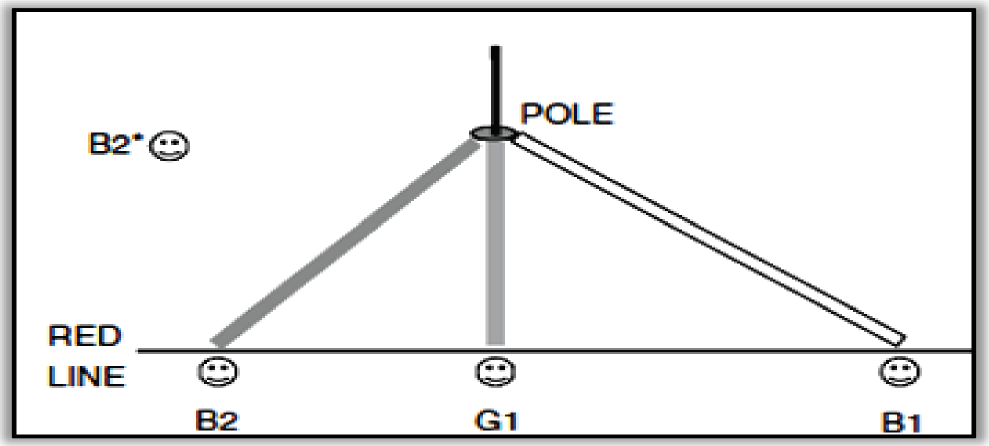


Figure 2.7: Student B_2’s solution in the quoits game (Groves, 2012: 126)

Considering the circle lesson (Figure, 2.7) that Groves (2012) video recorded in a Japanese Grade 3 class of eight children, Mr J. wanted to teach the concept of a circle and decided to engage students in a meaningful exercise that involved them in an activity that resulted in the learning of integrated concepts of a circle. The teacher then placed students in positions so that they could throw rings, in a pole. The teacher began by placing the students B_2, G_1 and B_1 in three positions.

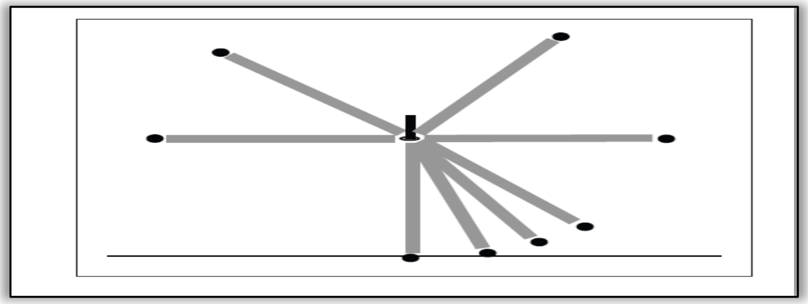


Figure 2.8: Paper representations of children's positions (Groves, 2012: 127)

Student B_2 (Figure 2.8) drew this diagram and noted that it was unfair because some were closer to the pole and were advantaged. Then the teacher asked them about what needed to be done to make it fair and the students positioned themselves equidistant from the pole. In the transcripts provided by Groves (2012) the teacher then helped them arrange themselves in various ways until they realised that a circle had been formed and there were actually infinite ways in which they could arrange themselves. This activity assisted learners to learn various concepts of a circle simultaneously and these were; the radii, the circumference, and the definition of the concept of a circle. Such learning was connected to this social experience; when learners argued about being unfair in the positioning and repositioned themselves equally they were actually forming the radii, when moving to multiple positions they were coming up with the definition of a circle as a locus and at the same time coming up with the circumference. This was a Grade 3 class and this example illuminates the idea that conceptual understanding can take place in all grades.

2.3.3 Strategic Competence

Strategic competence is the learners' ability to frame, symbolise and solve mathematical problems, and this is better known as problem solving and problem formulation (Kilpatrick, et al., 2001). Learners who see themselves as having the ability to formulate mathematical problems, devise problem solving strategies using mathematical concepts and procedures appropriately are known to be applying (Star, 2005). Below I explore some detail on the three components of strategic competence.

First, problem formulation is associated to task organisation or problem posing (Kontorovich, Koichu, Leikin & Berman, 2012; Land, 2017) which is classified into three types, which are, structured tasks, semi-structured tasks and free problem posing. It has been established that ANA questions are tasks (Stein, Grover & Henningsen, 1996) designed for evaluative purposes as stipulated in curricular and policy documents (Amit & Fried, 2002; DBE, 2011). Although the three types of task posing as used by Kontorovich et al. (2012) were in the context of teaching and learning of mathematics, I observe that most of its contents are also relevant to posing tasks for evaluative purposes. The first class, structured tasks, as outlined by Silver, Mamona-Downs, Leung and Kenney (1996) are mathematical tasks that test novel problems resulting from known problems and familiar solution strategies. Contrary, semi-structured tasks are those that pose a mathematical question based on a contextualised story normally called word problems (Kontorovich et al., 2012). Ironically, a free problem posing task is used in mathematical competitions to appeal to the competitors and differs from the other two as it is not confined to specific mathematical content (Stoyanova, 1998). I observe that for ANA purposes, only structured and semi-structured tasks may be used. Most importantly, formulation of tasks in the context of structured or semi-structured, examiners must ponder if a task stimulate the most of following: (1) Formulating data that links with known or unfamiliar problems; (2) locating a useful hypothesis; and (3) sketching a plan to answer the problem (Guberman & Leikin, 2013) as vital mathematical problem solving abilities. Moreover, problem formulation is reliant on the aesthetics of the mathematical problem (Singer & Voica, 2013) which may be in the form of a discovery or a process. Therefore, formulating a problem in the context of a discovery, tests learners' capability to be invent, innovate and persevere (Guberman & Leikin, 2013). By contrast, a process task requires learners to conceptualise pertinent hypotheses using deep knowledge to solve the problem and additionally monitor, assess and modify their thought processes during problem solving (Kim, Park, Moore & Varma, 2013) which postulates metacognition.

Second, problem representation needs to consider, (Stein et al., 1996) the presence of multiple solution strategies and a range of problem solving methods. Consequently, (Schoenfeld, 1985) observed that learners often follow a sequence of stages during mathematical problem representation as follows: (1) sketch figures; (2) discover related problems; (3) re-articulating problems, working back; and (4) assessing and validating procedures. Shapes are often drawn to signify geometric problems (Hsu & Silver, 2014). Consequently, diagrams are necessary when solving geometric problems (Kramarski, Mevarech & Arami, 2002) to explore the dynamics of the problem. In addition, the testing and verifying procedures, which are often exhibited by metacognitive learners, are all vital when ratifying the validity of solution strategies (Schoenfeld, 1985).

Third, solving mathematical problems relies on the aesthetics of particular mathematical questions, which may either be familiar or non-familiar (Guberman & Leikin, 2013; Roth, Ercikan, Simon & Fola; 2015; Sigley & Wilkinson, 2015; Sullivan, Borcek, Walker & Rennie, 2016). Consequently, the former posits reproductive thinking because the learner only needs to recall and use familiar procedures of solving those mathematical problems (Kilpatrick et al., 2001). However, the latter demands productive thinking as learners are required to invent procedures of problem formulation, problem representation and problem solving.

Lately, research (Granberg, 2016; Schoenfeld, 1985) has focused attention on the following four problem solving attributes: 1) understand the problem; 2) devising a plan; 3) acting; and 4) reflection that were adapted from Polya (1945). Such problem solving attributes were regarded as aiming on the situation of the problem by Schoenfeld, (1985) who explored them further to serve for the cognitive aspect through the use of five steps, which are; (1) reading a given problem, (2) exploration of prior knowledge, (3) drawing a suitable plan for solving the problem, (4) implementing the plan by solving the problem; and (5) verifying the correctness of the answer. Consequently, when solving mathematical problems, learners often spend time reading the problem, revisiting their previous knowledge, adjusting their previous knowledge to suit the present knowledge and re-doing the problem in instances where learners make errors or where they use wrong methods for their solution strategies (Granberg, 2016). Subsequently, a metacognitive learner is able to modify erroneous solution strategies to get the correct solution (Hsu & Silver, 2014). A non-metacognitive learner will reach a dead end due to failure to modify erroneous mathematical problem solving strategies (Kramarski et al., 2002). This researcher believes that proficiencies and in-proficiencies of solving mathematical problems are also reliant productive thinking skills that learners exhibit. The next section concentrates on studies that addressed this strand.

Guberman and Leikin (2013) worked on a study with twenty seven prospective teachers who were in their third and fourth year of Bed program in elementary school mathematics in a University in Israel. Their study focused on views of prospective teachers' perceptions on the difficulty of mathematical problems. Prospective teachers worked with mathematical tasks that integrated various topics and demanded knowledge of various mathematical concepts. The tasks were structured such a way that they produced multiple solutions with a variety of representations. Mathematical problems that were in the post-test were more complex as compared to those in the pre-test. They indicated that problem solving is regarded as a powerful means of developing robust and connected mathematical knowledge. They also pointed out that challenging mathematical problems must be included in the curriculum in a spiral way across the hierarchy of elementary and high school education and that the tasks must provide the following: (a) motivate students; (b) exclude known procedures; (c) initiate students into the problem solving activity; and (d) provide multiple approaches to the solution. The results of the study revealed that as prospective teachers participated in the problem solving course, their problem solving ability improved for both high and low achievers. It was observed that there was a significant shift from the use of trial and error methods that prospective teachers used in the pre-test towards the use of systemic strategies in the post-test. There was also an improved fluency and flexibility among the prospective teachers in their ability to exhibit multiple solution tasks to a particular problem.

A study by Thanheiser (2014) was conducted with fourth year prospective teacher's conceptions when working with well-designed tasks, documenting their successes and carrying out an analysis of their conceptual difficulty. Two tasks were developed, the first being one to determine the prospective teachers' conceptions at the beginning of the study and the other task to advance conceptual knowledge of working with multi-digit numbers. The findings of this study revealed that engaging prospective teachers with the two tasks improved their conceptions of multi-digit numbers, especially their conceptual knowledge and sense making. Another observation was that when prospective teachers access students' mathematical thinking through, helping them with the tasks helped address the prospective teachers' conceptions.

Prochazkova (2013) conducted a study with teacher trainees who worked on teaching experiments with 13 -14 year old students in high schools. The teacher trainees prepared lesson plans that took into consideration the content and language integrated into learning mathematics. They did micro-teaching before working with students. The focus of the study was to evaluate the effectiveness of lesson plans as well as monitor the development of content and language of mathematics with intensified exposure of students. Prochazkova (2013) argued that students' low proficiency in the language of tasks and learning the mathematical language enhances the following skills: (1) looking at the mathematical content of problems from multiple perspectives; (2) acquisition of the mathematical language of the task; (3) the creation of bridge between the student's mother tongue and the language of the task assist in expanding the expression of mathematical concepts and mathematical processes; and (4) changing the language of the task increases procedural fluency, motivates students and enhances higher order thinking skills. Results of the study revealed that mathematics students opted for more complicated concepts that were in conjunction with the development of the language of the task. Mathematics teacher trainees improved lesson plans to be more innovative after reflecting on their effectiveness.

Greenes (2014) conducted a study with Grade 5 to 8 students on problem solving with nine tasks that demanded a variety of problem solving strategies. The focus of the study was to initiate students into tasks that promoted hard thinking. The study was also aimed at stimulating students' curiosity, perseverance and flexibility to collaboratively solve problems for which their answers were not immediately known. The study came up with the following findings: (1) students developed confidence in the mathematics language of the tasks; (2) developed some innovative and creative skills in approaching their solutions; and (3) students developed an awareness of using mathematical representations as aids in their problem solving.

Kilpatrick et al. (2001) explained strategic competence as shown below:

> '*There are mutually supportive relations between strategic competence and both conceptual understanding and procedural fluency, as the various approaches to the cycle shop problem illustrate. The development of strategies for solving nonroutine problems depends on understanding the quantities involved in the problems and their relationships as well as on fluency in solving routine problems. Similarly, developing competence in solving nonroutine problems provides a context and motivation for learning to solve routine problems and for understanding concepts such as given, unknown, condition, and solution.'* (Kilpatrick et al., (2001: 127).

Kilpatrick et al. (2001) explained that there are routine and non-routine problems. Routine problems are problems that can be solved by learners using their previous experience on methods by applying known procedures to solve the problem. Kilpatrick et al. (2001) argue that routine problems require reproductive thinking where learners only reproduce and apply known problem solving strategies. Non-routine problems require productive thinking where learners form new schema in inventive ways to become familiar with the problem and hence solve the problem. Schoenfeld (2007) argued that a learner who performs routine problems has generative skills and one who performs non-routine problems has evaluative skills. Generative skill is explained as the ability to generate known procedures to solve a problem and evaluative skills require learners to deduce correct procedures to solve the problem. Schoenfeld (2007) explained that learners with the ability of solving mathematics problems limited to ‘how’ the problem is solved have procedural fluency and learners who solve problems showing ‘how and why’ have conceptual knowledge.

> *‘Strategic competence refers to the ability to formulate mathematical problems, represent them, and solve them. This strand is similar to what has been called problem solving and problem formulation in the literature of mathematics education and cognitive science, and mathematical problem solving, in particular, has been studied extensively’* (Kilpatrick et al., 2001: 124).

These features of problem solving are essential and assist the current study to advance this SMP and they are: *problem formulation*, *representing the problem* and *solving the problem*. However, Kilpatrick et al. (2001) failed to show a clear problem solving structure of the routine and non-routine problems. In developing some structure in problem formulation, problem solving as well as problem formulation, the researcher in the current study makes use of Guberman and Leikin (2013) demands of problem solving.

> ‘problem solving requires mathematical knowledge that facilitates (1) semantic understanding of the problem, (2) ability to connect the given problem with appropriate pieces of information learnt in the past, (3) use of multiple representations in solving the problems, (4) ability to recognize similarities in the structure of different problems, (5) metacognitive analysis of a problem and its solution; and (6) meta-mathematical awareness of the esthetics of the problem and its solutions’ (Guberman & Leikin, 2013: 34)

From these demands of problem solving, as well as the other literature reviewed in the current study on this SMP, five constructs emerge that assist the researcher adapt this strand and they are: (1) readability or semantics; (2) multiple representations; (3) reproductive thinking; (4) meta-cognition; and (5) productive thinking. The next sections review these constructs in detail.

❖ READABILITY OR SEMANTICS

Proficiency in language of learning and teaching is essential for thinking and communication in mathematics education (Gough, 2007; Morgan; 1996; Riordain & O’Donoghue, 2008; Roth, Ercikan, Simon & Fola, 2015). Hoffert (2009) suggested that problems in standardised tests should be structured in a language that is accessible to students. Furthermore, she pointed out that, even though mathematics is considered a universal language, second language speakers of English struggle to do mathematics in English. In South Africa, in a majority of mathematics classrooms, learners are taught mathematics in English, where for some of them English is their second, third or fourth language, (Jantjies & Joy, 2015). Importantly, Jones, Hopper, Franz, Knott and Evans (2008) as well as Wium and Louw (2012) conceptualised mathematics as a language that is used by people to become numerate in communicating, solving, reasoning and justifying mathematical relations, algorithms, computations and conjectures. Furthermore, Greenes (2014) and Prochazkova (2013) stressed that, for students to have the desire to attempt unfamiliar problems, they need language proficiency that will enable them to learn subject-specific terminology to be able to articulate and express their mathematical thinking.

Research (Bailey, Blackstock-Bernstein & Heritage, 2015; Roth, Ercikan, Simon & Fola, 2015; Sigley & Wilkinson, 2015) has shown that students often share their mathematical ideas and conjectures with their peers as well as teachers to make sense and further justify their thinking. Mathematical communication in problem solving often takes the form of symbols, drawings, graphs and manipulatives which results in a complex linguistic structure of mathematics (O'Halloran, 2015; Sigley & Wilkinson, 2015). Qualities of mathematical language make it unique to the English language and hence mathematics proficiency and English proficiency are not linked (Bailey et al., 2015). Most importantly, Prochazkova (2013) pointed out that learning mathematics in a different language provides access for students to a variety of perspectives of the content and different epistemological dimensions of mathematical knowledge. This author gave an example of the formula for calculating the area of a triangle in learning mathematics in English and learning mathematics in Czech: $S = \frac{a.v}{2}$;and $A = \frac{1}{2}b.h$. In these two instances Prochazkova (2013) pointed out that the labels and dimensions of the objects change, allowing students to explore the new meaning in English with added rigour.

Sigley and Wilkinson (2015) suggested three important features of the mathematics register which were: (1) *mathematics lexicon* this refers to unique usage of ordinary language and technical words found in mathematics such as, cosine, sine, trigonometry (technical) and length, prove, value (ordinary). The technical words are not accessible in ordinary language and the ordinary words contain contradictory meaning in mathematics and in ordinary language. An example is the word *prove,* in ordinary language, this means give evidence whilst in mathematics it means use either logic, analogies or empirical evidence. (2) *Mathematics syntax* refers to a specialised use of words in mathematics as connectors or to relate an idea such as, 'product of' in mathematics means multiply and in ordinary language it has a different meaning. (3) *Mathematics discourse features* refers to linguistic features, sequencing, procedures and justifications during mathematics teaching and learning. Sigley and Wilkinson (2015) pointed out that teachers often switch their modes as follows; for explanation teachers often use explicit inferences directed at students' understanding of mathematical concepts, and for justification they use logical inferences that justify the appropriateness of the explanation.

To explore the mathematics register, Bailey et al. (2015) reflected on how students evolve their conceptions when engaging with mathematics tasks at various levels of their schooling years. Students seem to use simple sentences in their lower grades which apparently become more complex in higher grades. However, the challenge still remains concerning proper attention that teachers afford learners in their mathematical conversation as well as written responses. One way of bridging such a challenge is to connect mathematics to learners' everyday experiences (Mwakapenda, 2008).

❖ MULTIPLE REPRESENTATIONS

The significance of the use of multiple representations in the teaching and learning of mathematics as well as in research is widely acknowledged (Hyde, George, Mynard, Hull, Watson & Watson, 2006; Perry & Atkins, 2002; Tsamir, Tirosh, Tabach & Levenson, 2010; Wong, Yin, Yang, & Cheng, 2011). Students navigate mathematics concepts and relations by creating, comparing and using a variety of representations (NCTM, 2000:280). Such representations are objects, pictures, symbols, charts and graphs which reveal to students that there are various ways of representing and solving a problem.

Researchers have shown that successful problem solvers have the ability to use a variety of representations to solve a problem. However, this depends on various factors such as, type of test, type of context, and type of learning aids provided (Hwang, Su, Huang & Dong, 2009; Jonassen, 2001; Wong et al. 2011). The notion of representations in mathematics is outlined in Figure 2.9 that is proposed by Stephen, Pape and Tchoshanov (2001) in their description of the role of representation in enhancing mathematical understanding. In this synopsis, learners navigate through representations such as numerals, graphs, equations, tables and diagrams referred to as '*written five'.* These are called external representations, where learners use their mental imagery of counting and numeracy.

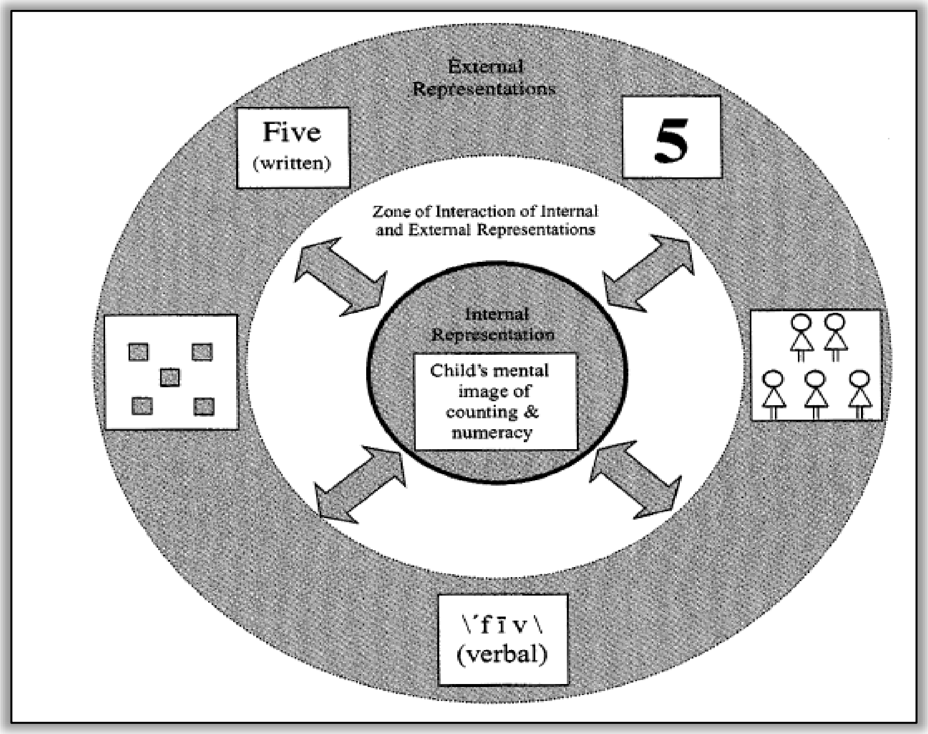


Figure 2.9: The relationship between internal and external representations in developing the child's understanding of the concept of numeracy (Stephen et al., 2001: 119)

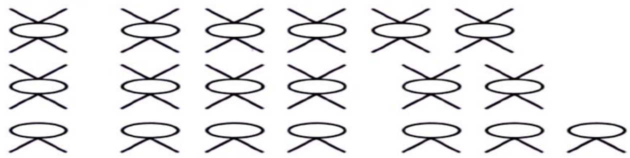
A pertinent classroom example that shows a teacher giving students opportunities to present a mathematical idea is shown in Figure 2.10, in the '*taking the stock*' by Tripathi (2008: 440). "*A farmer had 19 animals on his farm-some* *chickens and some cows. He also knew that there were a total of 62 legs on the animals on the farm. How many of each kind of animal did he have?*" (Tripathi, 2008; 440). Figure 2.10 shows three different representations that helped students get the answer.

Strategy 1: Use a table

No. of Chickens	No. of Legs	No. of Cows	No. of Legs	Total Legs
0	0	19	76	76
1	2	18	72	74
2	4	17	68	72
3	6	16	64	70
4	8	15	60	68
5	10	14	56	66
6	12	13	52	64
7	**14**	**12**	**48**	**62**
8				

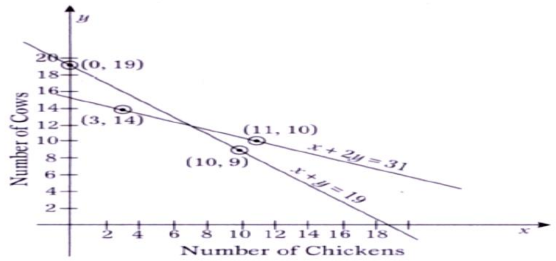
Increase the number of chickens by 1. Write down the corresponding number of cows. Count the total number of legs until they total 62, which will occur once 7 chickens and 12 cows are listed.

Strategy 2: Draw a pictorial model



The ovals represent the animals. Give each animal 2 legs. The total number of legs is $19 \times 2 = 38$. Add 2 legs to each animal until the total number of legs becomes 62, representing 7 chickens and 12 cows.

Strategy 3: Use a graph



The solution is the intersection of the two lines that represent the linear equations $x + y = 19$ and $x + 2y = 31$. The coordinates of the point of intersection are (7, 12). Thus, there are 7 chickens and 12 cows.

Figure 2.10: Visual strategies for solving the Taking stock problem (Tripathi, 2008: 440)

❖ REPRODUCTIVE THINKING

In explaining the meaning of reproductive thinking, the current study, reflects on learners' cognition, memory and computational fluency. Research on mathematics cognition has indicated that learners have the tendency of expecting computations that are familiar where they have already acquired their solution strategy (Foster, 2011; Kramarski, Mevarech & Arami, 2002; Stanic, Silver & Smith, 1990). What is referred to as reproductive thinking in this study are features of a task that require a recall of mathematics procedures and it is clear that such are embedded in both computations as well as complex procedures. The procedure may be complex but there are elements in it of pieces of knowledge that constitute the computations (Flowers et al., 2003; Russell, 2008).

Teachers have the tendency of funnelling students which gives the impression that that tasks must have known answers (Brodie, 2007; Luneta, 2015a). As such, this has an influence on learners' cognition which, according to Tchoshanov (2011), results in learners having what is called '*type 1' knowledge*, ability to recall and apply basic mathematics facts, rules, algorithms and procedures. This is what is referred to as reproductive thinking in the current study. Such learners have procedural knowledge and Tchoshanov (2011) gives the following example; "*if a teacher is able to recall a rule for fraction division or to solve simple fraction division problem such as* $1\frac{3}{4} \div \frac{1}{2} =$*, we say that she has procedural knowledge of fraction division*" (Tchoshanov, 2011: 142). This implies that reproductive thinking is limited to procedural knowledge in routine procedures.

❖ METACOGNITION

In various studies, metacognition is often described as knowledge control of personal cognition as well as regulation, monitoring and modification of behaviours as learners navigate through mathematical tasks (Callahan & Garofalo, 1987; Fagnant & Crahay, 2011; Kim et al., 2013). In their study, Kim et al. (2013) pointed out that there are three levels of metacognition: *individual level*, which is reflection, monitoring and regulation of personal cognition; *social level*; when individuals get stuck during a complex problem solving, they often seek external assistance such as collaborating with other people and sources of information; *environmental level* means that in some cases, learners seek empirical evidence from the environment to justify their conceptions. A balance in these three levels is essential for learners when solving non-routine, complex and authentic problems. Answers to these problems are not readily available and learners often need to regulate their thinking when they get stuck, make errors, reflect on their mistakes, hence, they engage in productive thinking which is discussed in the next section (Kramarski et al., 2002; Neuenhaus, Artelt, Lingel, Schneider, 2011; Yilmaz-Tuzun, & Topcu, 2010; Veloo, Krishnasamy & Abdullah, 2015). Research has proved that there is an achievement of optimal learning where mathematical tasks always foster high-level student thinking coupled with reasoning and proof (Koichu & Leron, 2015; Levenson, 2013; Pedemonte & Balacheff, 2016; Sullivan et al., 2016)

❖ **PRODUCTIVE THINKING**

Non-routine problems are those that are not readily accessible to the learner in terms of solution strategy, hence the learner needs to invent a particular way to manipulate and solve the problem (Kilpatrick et al., 2001). Guberman and Leikin (2013) pointed out that, when learners experience non-routine problems, they require the ability to connect the problem with relevant pieces of information. Hence, productive thinking refers to inventive ways of solving a problem that does not have a pre-conceived answer. Similarly, Tchoshanov (2011) argued that non-routine procedures require an accelerated quantity and quality of connecting mathematical procedures and relations which is conceptual in nature. In the fraction division problem, learners in this instance are required to connect various procedures to solve the problem as shown here; "*Solve the following fraction division problem* $1\frac{3}{4} \div \frac{1}{2} =$*, in more than one way (e.g., draw a diagram or illustrate it with manipulatives)*" (Tchoshanov, 2011: 142).

2.3.4 Adaptive Reasoning

Adaptive reasoning is the learners' ability to exhibit logical thought, reflection, explanation, conjecturing and justification (Kilpatrick et al., 2001; Komatsu, Jones, Ikeda & Narazaki, 2017). This gives an indication that the SMP are paramount in making sense of procedures, connecting them and showing the mastery of problem solving processes. However, most research has revealed the gap between classroom learning activities that promote and the type of learning that allows deeper understanding of complex reasoning processes (Bergqvist & Lithner, 2012; Blanton & Stylianou, 2014; Brodie, 2010; Palha, Dekker, Gravemeijer & van Hout-Wolters, 2013; Sumpter & Hedefalk, 2015; Zazkis, 2015). Furthermore, it has been reported in other studies that teachers' limited knowledge of proof as well as constrained opportunities that they make available to their students inhibit constructive mathematical reasoning at the expense of imitative reasoning (Jonsson, Norqvist, Liljekvist & Lithner, 2014; Ramful, 2014; Samkoff & Weber, 2015; Savic, 2015; Soto-Johnson & Troup, 2014; Yopp, 2015; Zandieh, Roh & Knapp, 2014). Other studies such as Bleiler-Baxter (2017) reports that student teachers lack important elements of proof, a key element of reasoning, which are: presenting, discussing, conjecturing and critiquing. This is despite inferences mentioned by Kilpatrick et al. (2001) on the importance of adaptive reasoning as mentioned below:

> *"In mathematics, adaptive reasoning is the glue that holds everything together, the lodestar that guides learning. One uses it to navigate through the many facts, procedures, concepts, and solution methods and to see that they all fit together in some way, that they make sense. In mathematics, deductive reasoning is used to settle disputes and disagreements. Answers are right because they follow from some agreed upon assumptions through series of logical steps. Students who disagree about a mathematical answer need not rely on checking with the teacher, collecting opinions from their classmates, or gathering data from outside the classroom."* (Kilpatrick et al., 2001: 129).

According to this explanation by Kilpatrick et al. (2001), in all the SMP, adaptive reasoning holds them together. This implies that the strength in mathematical knowledge base, skills and values depends on the ability to think logically. One of the specific aims for Curriculum and Assessment Policy Statements, DBE (2011) is *'learn to listen, communicate, think and apply the mathematical knowledge gained'* (p9). This is possible when learners have fully developed this strand of mathematical proficiency. Communication of mathematical reasoning is essential to the confidence that learners show to their classmates, to stand in front of them and justify their mathematical thoughts (Cantlon, 1998; Whitacre, Azuz, Lamb, Bishop, Schappelle, & Philipp, 2017). However, such efforts are fruitless when considering the findings of the TIMSS Video Study which was conducted in eight countries and in 87 Australian lessons which reported that there were no signs of formal or informal reasoning (Stacey & Vincent, 2009). These results are consistent with those reported by Ally (2011) who found that there was an absence of reasoning in some Grade 6 mathematics classrooms in South Africa.

Star (2005) argued that adaptive reasoning is 'reasoning' which describes the learners' ability to use logic to explain and justify problem solving strategies and furthermore expand reasoning from known procedure to those not yet known. Schoenfeld (2007) pointed out that adaptive reasoning is seen when learners justify the use of concepts, procedures, and algorithms and justify solving a problem in a particular way. Proof (Hanna, 2000) is not limited to justifying an argument but rather includes promoting conceptual understanding, acting as an automated machinery in problem solving and justifying the authenticity of the solution as well as the solution strategy. In mathematics education, reasoning means a clear procedure of axiomatic reasoning, coherence deductions and formal inferences (Reid, 2002).

Knuth (2002) found that a proof in its utility by Grade twelve students enables them to do the following: (1) identify reasoning and proof as important aspects of mathematics; (2) construct and explore mathematical conjectures; (3) exhibit and scrutinise mathematical arguments and proofs; and (4) select and utilise several kinds of reasoning and methods of proof. In the current study, the researcher considers adaptive reasoning as a glue to the other SMP if learners exhibit these traits in their mathematical reasoning. According to the researcher's understanding, reasoning and justification of the use of mathematical concepts and procedures as well as the problem solving strategies must be a valid, reliable piece of knowledge and be applicable to learning transfer. Hanna (2000) suggested some significant functions of proof as shown below;

"• *Verification* (concerned with the truth of a statement)
• *Explanation (providing insight into why it is true)*
• *Systematisation (the organisation of various results into a deductive system of axioms, major concepts and theorems)*
• *Discovery (the discovery or invention of new results)*
• *Communication (the transmission of mathematical knowledge)*
• *Construction of an empirical theory*
• *Exploration of the meaning of a definition or the consequences of an assumption*
• *Incorporation of a well-known fact into a new framework and thus viewing it from a fresh perspective." (Hanna, 2000: 8)*

A true statement needs to be verified against mathematical knowledge. One must become aware of it and structure it in a logical way (Hanna, 2000). Furthermore, it must show innovations, convey mathematical knowledge, experiential theory, and a search for meaning as well as conjecturing. However, Quinn, Evitts and Heinz (2009) observed that a majority of students found that proofs are very challenging and teachers had problems in helping their students write proofs. What seems a challenge is that that teachers themselves struggled to structure tasks that tested higher order and critical thinking skills (Luneta, 2015a). In their study on reasoning, Lachmy and Koichu (2014) as well as Zazkis et al. (2015) pointed out that students' mathematical reasoning often reflected on the structure of the proof and that students fail to make a distinction between empirical and deductive reasoning, as well as the power associated with deductive inferences.

Tillema and Hackenberg (2011) used multiplicative notation with fractions to explain reasoning. Their illustrations are shown in Figure 2.11.

The Ribbon Problem:

"I have $\frac{3}{4}$ of a yard of ribbon. My friend needs $\frac{2}{5}$ of that amount. Draw a picture of how much of a yard she needs. Then, write mathematical notation to represent your reasoning" (Tillema & Hackenberg, 2011: 29).

The student first wrote $\frac{2}{5} \times \frac{3}{4} = \frac{6}{20}$. The authors argue that, although this notation is correct, it ignores that uses notation as traces student's reasoning. The response in Figure 2.11 helps the student to be aware of the traces of reasoning on the ribbon problem by reflecting on the abstraction of the problem and make the solution visible. The use of the representation in Figure 2.11 acts as an aid for the student and gives meaning to the solution (Tillema & Hackenberg, 2011).

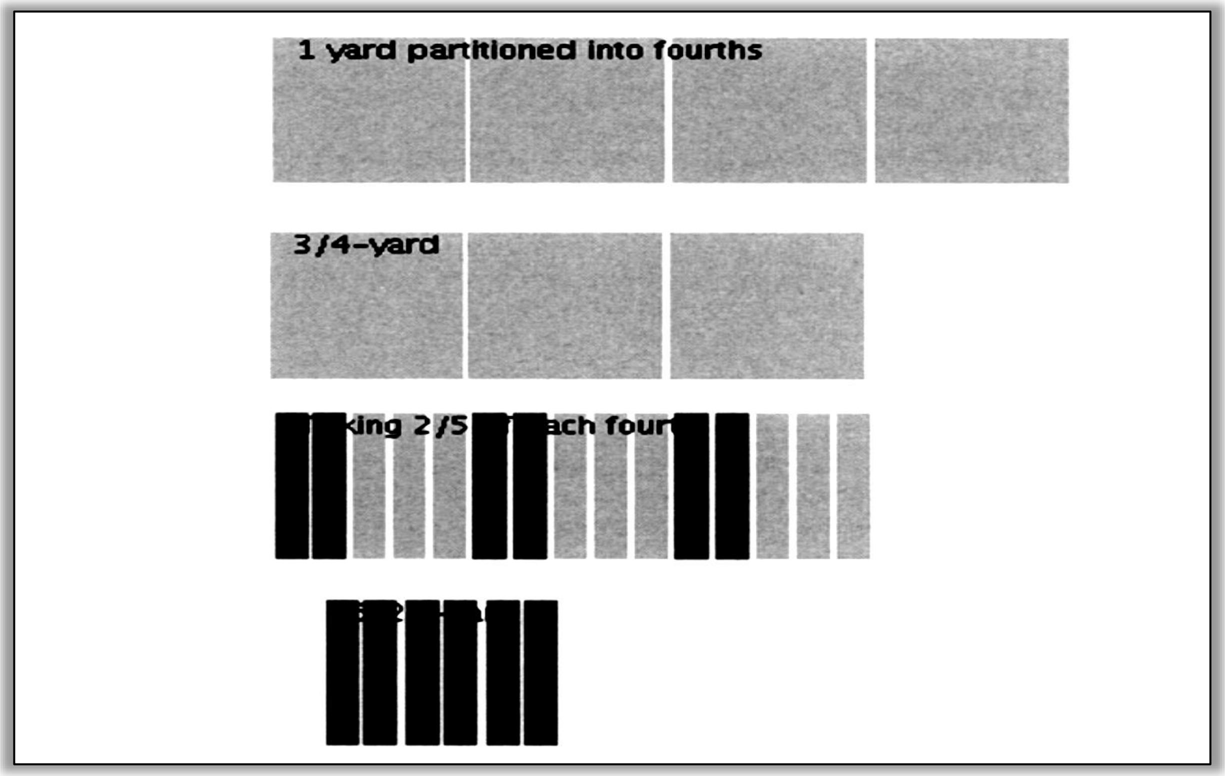


Figure 2.11: Illustration of distributive reasoning (Tillema & Hackenberg, 2011: 29)

Richardson, Carter and Berenson (2010) conducted a study using teaching experiments with Grade 5 students. The paramount goal of the study was to encourage mathematics teachers to use connected tasks in their mathematics teaching as means of promoting reasoning. There were three tasks, the tower task, pizza task and taxicab task. The results showed that the connected tasks provided students with the opportunity to make discursive predictions and conjectures about patterns that they observed from the tasks. The use of information from tables and pattern blocks enhanced students' justification of their own conjectures in an attempt to validate their conjectures. Connected tasks were key in providing students multiple representations, making known their ideas and linking their previous knowledge.

A case study by Bieda (2010) of seven middle-school experienced teachers on the implementation of tasks that enhanced the production of mathematical proofs came up with the following findings: (1) Students attempted these tasks by formulating conjectures and half of which were proved. (2) Teachers provided insufficient feedback to justify attempts from students to prove the tasks. (3) Most of the students' proofs used empirical means to justify their inferences. (4) Students attempts to prove were inhibited by the teacher's lack of ample time to review proof attempts by students. (5) Irrespective of a curriculum that encouraged the formulation of proofs, the instruction that was carried out by experienced teachers was too shallow in providing reasoning opportunities. These findings revealed that teachers' knowledge of proofs was insufficient.

Boesen, Lithner and Palm (2010) conducted a series of studies with the focus on analysing discursive learning environments to view students' opportunities of learning different types of mathematical reasoning. Two types of tasks were used based on the reasoning opportunities they provided: *imitative reasoning* which was suitable for routine tasks and allowed students to exhibit memorised reasoning and algorithmic reasoning. *Memorised reasoning* required students to recall through memory an answer and strategy to reason an argument. Algorithmic reasoning demanded recall of a sequence of rules used to solve a task through mathematical reasoning. The second type of tasks demanded what they called *Creative Mathematically Founded Reasoning*. Such tasks were identified as having three properties: (1*) novelty,* which was a new (to the problem reasoned) sequence of reasoning exhibited by students or a recreated sequence to a forgotten reasoning; (2) *plausibility,* these were statements that supported why arguments were true; and (3) *mathematical foundation* arguments that exhibit key elements that are embedded in properties of the reasoning. In the context of their study Boesen et al. (2010) described an argument as a substantial explanation that convincingly shows that a reasoning is appropriate. The results of this study and the study by Bergqvist and Lithner (2012) revealed that a majority of students exhibited imitative reasoning and lacked creativity in constructing correct reasoning. Furthermore, students could easily exhibit creative mathematically founded reasoning. However, this was their alternative when they had forgotten algorithmic reasoning or memorised reasoning. Lastly, students had a tendency to use mathematical foundations to justify their imitative reasoning. Such results reveal that learners preferred imitative reasoning rather than creative reasoning.

Stylianides et al. (2013) focused on teachers' knowledge and their challenges with reasoning and proving in their prospective classrooms. Teachers engaged their students with tasks that demanded the formulation of generalisations that were in form of conjectures and the development of arguments created for the truth or falsity of the generalisations. They worked closely with three prospective teachers who developed tasks and lesson plans on reasoning and proving. The findings of the study were as follows. (1) Teachers faced difficulty in figuring student's prior knowledge on reasoning and proving when engaging them with new topics. (2) Implementation of high level tasks was a serious challenge even to experienced teachers. (3) As students got used to working with tasks on reasoning and proving, they shifted from being passive recipients of knowledge to active constructors of knowledge. (4) Teachers had the tendency of funnelling students to a particular direction, instead of allowing them to develop their own conjectures.

Kilpatrick et al. (2001) described adaptive reasoning as shown below:

> '*Adaptive reasoning refers to the capacity to think logically about the relationships among concepts and situations. Such reasoning is correct and valid, stems from careful consideration of alternatives, and includes knowledge of how to justify the conclusions. In mathematics, adaptive reasoning is the glue that holds everything together, the lodestar that guides learning. One uses it to navigate through the many facts, procedures, concepts, and solution methods and to see that they all fit together in some way, that they make sense. In mathematics, deductive reasoning is used to settle disputes and disagreements*.' (Kilpatrick et al., 2001: 129).

Kilpatrick et al. (2001) shows that mathematics learning is grounded in deductive reasoning and failed to show that other forms of proofs are essential in mathematics reasoning. In expanding this SMP, adaptive reasoning, and this study examined how Amir-Mofidi, Amiripour and Bijan-zadeh (2012) explained mathematical reasoning. The authors say mathematical reasoning is divided into three parts; (1) *Inductive proof*, (2) *Deductive proof,* and (3) *Analogical proof*. However, Kilpatrick et al. (2001) explained that mathematical reasoning is not confined to formal proof but also informal proof. In the current study the researcher identified these three forms of reasoning formal proof. The fourth type of reasoning was extracted from an analytic framework by Stylianides (2008), especially the mathematical component of the framework as shown in Figure 2.12.

	Reasoning-and-proving			
	Making Mathematical Generalizations		Providing Support to Mathematical Claims	
Mathematical Component	Identifying a Pattern	Making a Conjecture	Providing a Proof	Providing a Non-proof Argument
	• Plausible Pattern • Definite Pattern	• Conjecture	• Generic Example • Demonstration	• Empirical Argument • Rationale
Psychological Component	What is the solver's perception of the mathematical nature of a pattern / conjecture / proof / non-proof argument?			
Pedagogical Component	How does the mathematical nature of a pattern / conjecture / proof / non-proof argument compare with the solver's perception of this nature? How can the mathematical nature of a pattern / conjecture / proof / non-proof argument become transparent to the solver?			

Figure 2.12: Analytic framework (Stylianides, 2008:10)

The mathematical component of the framework yields the four components of mathematical reasoning. Under the component making mathematical generalisations, *conjecturing* is there and also *analogical* reasoning which is identifying patterns, and in providing support to mathematical claims, an empirical argument is *inductive* and lastly, generic examples are *deductive.*

In this instance, clarity is provided on informal proof and put it in the perspective of the South African education context. DoET (2007) explained informal reasoning as *conjecturing*, and this is the fourth reasoning. To bring clarity to what adaptive reasoning entails, the researcher focuses attention towards how Stylianides et al. (2013) described a proof, and this clarifies the qualities of a mathematical proof.

> ‘*According to this conceptualization, a proof in the context of a classroom community at a given time is a mathematical argument with the following three characteristics: (1) It uses statements accepted by the classroom community (set of accepted statements) that are true and available without further justification; (2) It employs forms of reasoning (modes of argumentation) that are valid and known to, or within the conceptual reach of, the classroom community; and (3) It is communicated with forms of expression (modes of argument representation) that are*

appropriate and known to, or within the conceptual reach of, the classroom community. (Stylianides et al., 2013: 1466).

❖ INDUCTIVE PROOF

Inductive proof in the context of mathematics education refers to proving from observable characteristics that can be generalised from facts. This is essential for students when they reason with contextual tasks (Amir-Mofidi et al., 2012; Baroody, 2005; Simon, 1996). According to Yopp (2010) inductive proof takes the form of a process that learners undergo to form conclusions from examples. He argues with illustrative examples that not all inductive arguments generate formal proof. In addition, Klauer and Phye (2008) argues that the paramount aim of inductive reasoning is to detect generalisations, rules and regularities. Morris (2002) focuses her definition of inductive proof on a premise that provides some feasible actions, not necessarily evidence for the conclusion. She argues that an inductive inference begins with a particular phenomenon leading towards a generalisation. She outlines the following example:

'*e.g., X and Y are qs: X and Y have property p; therefore all qs have property p)*' (Morris, 2002: 80.

Knuth (2002) acknowledges that the key aspect of a proof is to find the truth of a result. He gives an example of a problem that has been proved by induction as shown in below:

'*Prove: the sum of the first n positive integers is* $\frac{n(n+1)}{2}$.

For n=1 true since $1 = \frac{1(1+1)}{2} = 1$

Assume it is true for some arbitrary k, that is, $S_{(k+1)} = S_k + (k+1)$

$$= \frac{k(k+1)}{2} + k + 1$$

$$= \frac{k(k+1)}{2} + k + 1$$

$$= \frac{k(k+1)}{2} + k + 1$$

$$= \frac{(k+1)(k+2)}{2}$$

Therefore, the statement is true for k+1 if it is true for k. By induction, the statement is true for all n' (Knuth, 2008:487).

The proof above is considered a valid proof because it proves the opening statement. This proof tests if the statement holds for all positive integers, from *'1'* to *'k'* and finally *'k+1'* then it is from these tests that a generalisation is reached that the initial statement is true.

Figure 2.13 is a proof by induction. However, it is too abstract and may not convey substantial meaning to learners (Knuth, 2002; McLeod & Briggs; 1980). There is a suggestion by Knuth (2002) that the use of representations in a proof enhances sense-making why the sum must be equal to $\frac{n(n+1)}{2}$ as shown in Figure 2.13 above. Learners are able to see that the $'n^2'$ results from the area then dividing by two gives two congruent triangles. This proof is described by Knuth (2002) as a '*proof that explains as well as proves'*, which can be used to cater for the diversity in learning abilities in mathematics classrooms.

Klauer and Phye (2008) conducted a study to detect various thinking processes of 3600 kindergarten and pre-primary children on seventy four training experiments dealing with procedural inductive problems. The results of the study indicate that the use of procedural inductive proof enhances both learners' molten intelligences and improved educational opportunities concerning the content of mathematics for learners with diverse learning abilities. These results do not concentrate on how inductive reasoning affect learner's fluid intelligence.

In another study on inductive proof, Tomic and Klauer (1996) conducted training experiments with 34 Dutch students and 23 German students in primary schools. A pre-test which was an intelligence test, was first administered to the students before training them on inductive proof strategies. The study indicated the following results; (1) Training on inductive proof on both cases seemed to be effective in improving the intelligence of the students. (2) In the Dutch case, the improvement was not transferred to mathematics achievement and it was argued that this may be as a result of the cultural learning differences in both countries, inexperience of the trainers on inductive proof as well as the limited time spent on the training. (3) In the German case, the improvement was transferred to mathematics achievement and this was attributed to the experienced trainers on an inductive proof.

Prove: The sum of the first n positive integers is $n(n + 1)/2$.

We can represent the sum of the first n positive integers as triangular numbers.

1   1 + 2   1 + 2 + 3   1 + 2 + 3 + 4

The dots form isosceles right triangles, with the nth triangle containing

$$S(n) = 1 + 2 + 3 + 4 + \cdots + n \text{ dots.}$$

Overlaying a second isosceles right triangle of the same size so that the diagonals coincide produces a square containing n^2 dots plus n extra dots because of the overlapping diagonals. To illustrate, the figure below represents the fourth isosceles right triangle and another of the same size overlaid so that the diagonals coincide. In this case, a square containing 4^2 dots plus 4 extra dots caused by the overlapping diagonals is produced.

Therefore, in the general case (using the nth triangle), the number of dots produced by the two overlapping triangles is $2S(n) = n^2 + n$, so $S(n) = (n^2 + n)/2$.

Fig. 2
A proof that explains as well as proves,
adapted from Hanna (1990)

Figure 2.13: A proof that explains as well as proves (Knuth, 2002: 487)

A study by Kagan, Pearson and Welch (1966) on 155 first-graders on the conceptual impulsivity and inductive proof was conducted in Newton, Massachusetts. Findings of their study revealed that: (1) impulsive learners make many errors in their inductive proof because they reflect on the accuracy of their conclusions; (2) impulsive learners do not seem too wary about errors and mistakes; (3) reflective students are keen to identify their errors because they evaluate their inferences before voicing them in the public domain; and (4) teachers contribute to the failure of impulsive students because they always view them as having insufficient knowledge and fail to assist them to become reflective learners. Such findings show a need for professional development of teachers' ability to handle these diverse cases.

It is imperative to assist students to bridge the gap between inductive and deductive reasoning. Quinn et al. (2009) observed that inductive proof alone is not enough and that some instances students may reach incorrect conclusions. They showed that some students revealed that deductive arguments were more convincing than just giving examples. This statement prompted the researcher to look at deductive proofs below.

❖ DEDUCTIVE PROOF

Deductive proof disregards empirical evidence and is reasoning that takes into cognisance the derivation of facts drawn from a logical chain of reasoning, ideas or theories based on formal truth that have been accepted for their accuracy (Amir-Mofidi et al., 2012; Baroody, 2005; Simon, 1996). Additionally, Morris (2002) argues that deductive proof is not a mere derivation of facts, but is a logical argument that is true and valid followed by a true conclusion as shown below.

e.g., *'X is p or q; X is not p; therefore, X is q)'* (Morris, 2002: 80).

Furthermore, deductive proof is rigorous abstract logical proof of a mathematical argument to reach a valid and true conclusion, considering the underlying structure of mathematical reasoning (Ayalon & Even, 2008; Baroody, 2005; Markovits & Doyon, 2011). Proving deductively remains essential in mathematics classrooms, especially in reformed curriculum where sense-making originates from real-world problems (Lee, 2016; Stalvey & Vidakovic, 2015). Amir-Mofidi et al. (2012) illustrated deductive proof with an example that required them to show that the addition of two odd numbers always results in an even number as shown below.

'*Solve: suppose that (2m+1) and (2n+1) are odd numbers that m and n are natural numbers. Then its addition is following as;*$(2n+1)+(2m+1)=2m+2n+2=2(m+n+1)$' Amir (Mofidi et al., 2012: 2918).

The argument shown above leads to a conclusion that when adding two odd numbers the result is always an even number. This is shown by the co-efficient of two. Any number multiplied by two the answer is an even number. The conclusion,$2(m+n+1)$ is generic to all cases and is therefore a true and valid argument. This proof is a process of generating a generic statement that will generate an even number for the sum of any two odd numbers. This is a deductive proof and does not necessarily need to use an infinite sum of odd numbers to generalise.

Reflecting on the problem that was proved inductively by Knuth (2002), the same problem can be proved deductively as shown below.

$'Prove: The\ sum\ of\ the\ of\ the\ first\ n\ positive\ integers\ is\ \frac{n(n+1)}{2}$

$Let\ S_n = 1+2+3+\cdots+n$

$Then\ S_n = n+(n-1)+(n-2)+\cdots+1$

$Taking\ the\ sum\ of\ these\ two\ rows.$

$2S_n = (1+n)+[2+(n-1)]+[3+(n-2)]+\cdots+(n+1)$

$2S_n = (n+1)+(n+1)+(n+1)+\cdots+(n+1)$

$2S_n = n(n+1)$

$Therefore\ S_n = \frac{n(n+1),}{2}$ (Knuth, 2002: 488).

In this problem Knuth (2002) shows deductively that the sum of two positive integers will always be $\frac{n(n+1)}{2}$ and this approach uses only to sums to generate the conclusion instead of testing infinite sums then generalise.

Ayalon and Even (2008) conducted a study to examine the perceptions of twenty one professionals who were involved with deductive proof. The group of the professionals were made up as follows: four researchers in mathematics education with 2 PhD in mathematics education and 2 PhD in mathematics, six curriculum developers, teacher educators with 6 PhD in mathematics education and 1 MEd in mathematics education, 4 senior and junior school teachers with 2 BSc & MA mathematics, 1 MSc mathematics education & 1 Bed mathematics education, 3 mathematicians with PhD in mathematics, two Logicians of philosophy with PhD in

Philosophy, and two researchers in Science Education with 1 PhD in biology, 1 PhD science education. This indicates that seventeen of these professionals were directly involved in mathematics and the remaining four were not directly involved in mathematics' but rather in some form of logic and deductive proof. The study revealed two approaches towards the nature of deductive reasoning: the *systematic approach,* and the *logic approach*. The systematic approach was more concerned with the utility of deductive proof outside mathematics in step-by-step problem solving and the logical approach concerned with deductive reasoning as rules of formal logic. The rigour in the usefulness of logic was associated with the background of the professionals. Thus, the mathematicians were concentrating more on logic than the philosophy logisticians who were interested in onto the utility of more flexible deductive reasoning outside mathematics. This study indicates that rigour in deductive reasoning is associated with professional's content knowledge of mathematics.

In his study, Morris (2002) examined 30 pre-service elementary and middle school teachers' ability to extricate essential deductive and inductive features of a mathematical inference. Results of the study revealed that a majority of the pre-service teachers failed to differentiate between a deductive and inductive inference from statements of proof. There were a few pre-service teachers who were able to distinguish between deductive proof and inductive proof from mathematical proofs. This was as a result of their experiences with the two forms of reasoning. These findings may have negative effects to pre-service teaching of mathematical reasoning. Markovits and Doyon (2011) pointed out that deductive proof is difficult for adults who are well educated, hence this may have negative effects upon classroom practice.

❖ ANALOGICAL PROOF

Analogical reasoning is described by researchers (Amir-Mofidi et al., 2012; Lee & Sriraman, 2011) as the students' ability to reason about contrasting corresponding commonalities between mathematical relations to gather integral components such as equality and proportionality. In trying to review the definition of an analogy, Lee and Sriraman (2011) analysed the task used and revealed that it had assisted learners to identify three types of commonalities, *surface, transitional* and *relational similarities.* Surface similarity allowed learners to recall concepts that are common in the base object and the target object. Transitional commonality involves relating properties of the base object with the target object, and relational commonality is the construction of new concepts of the target object in form of a conjecture.

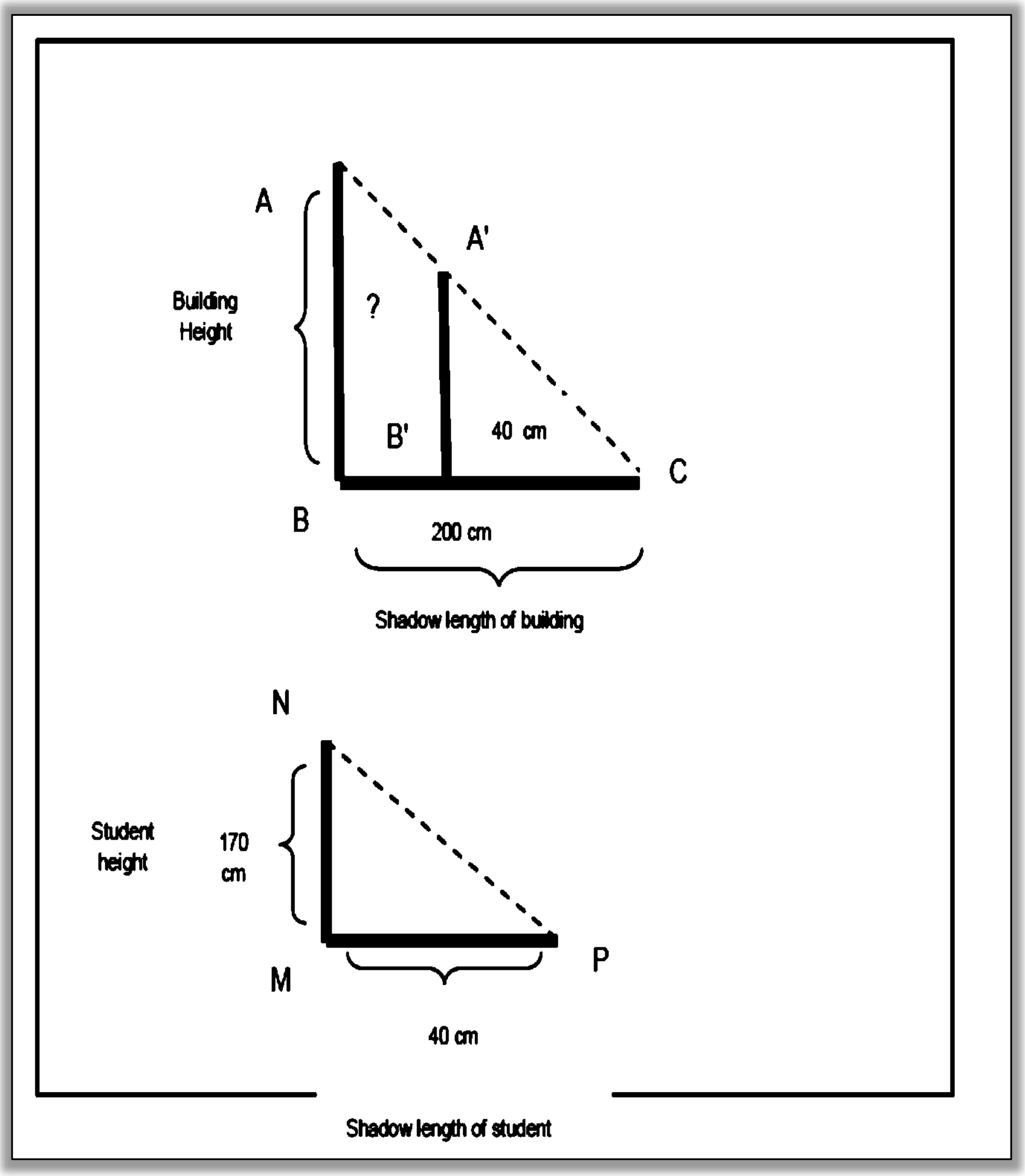


Figure 2.14: Analogical reasoning-mathematical (Amir-Mofidi et al., 2012: 2919)

In Figure 2.14, above, conceptualises an analogy as shown in Amir-Mofidi et al. (2012), which begins with students' calculations of the height of the school building. This exercise leads students to the use of the Thale's law on similarity, $\frac{AB}{A'B'} = \frac{BC}{B'C'} = \frac{AC}{A'C'}$. Students used the concept of congruent triangles to identify analogies such as;

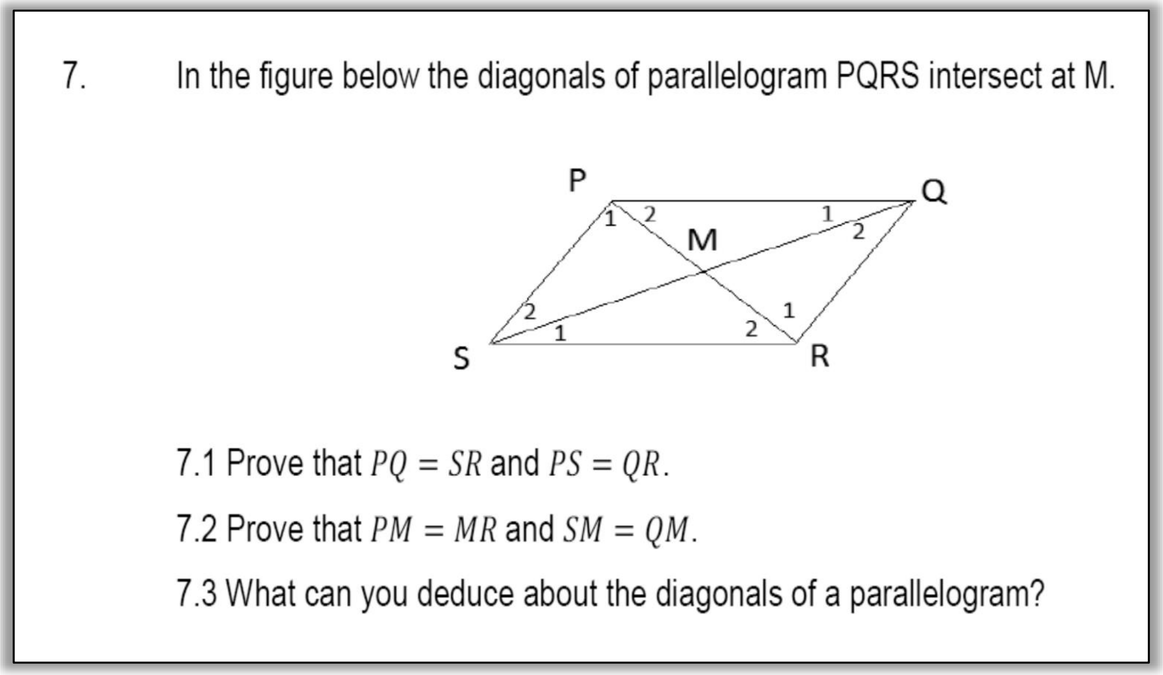
> '∠C is common in two angles; Δ ABC and Δ A'B'C' and ∠B and ∠B' are equal and right. Then ∠A, ∠A' are equal too. Also two angles; Δ ABC and Δ NMP are congruent triangles therefore angles of two respected triangles are equal too. On the other hands (MN = A'B'), (MP = B'C) and (NP = A'C)' (Amir-Mofidi, et al., 2012; 2920).

In their study Amir-Mofidi et al. (2012) investigated the effectiveness of analogical reasoning skills as means of improving learning in a mathematical context. Two mathematics examinations were used with thirty eight high school students, one exam as a pre-test and the other as a post-test with word problems on analogies. Findings from this research showed the following: (1) analogical reasoning improved creativity in students; (2) there was an improvement in student's abstract thinking that related to real life of students; (3) analogical reasoning developed student's reasoning ability and enhanced motivation in problem solving; (4) the use of analogical reasoning examples improved students' learning transfer; and (5) long-term retention of mathematical concepts was enhanced.

A study by Lee and Sriraman (2011) investigated eighth graders ability to generate conjectures using analogies of geometric figures. The study revealed four useful findings as follows. (1) Classical analogy was essential to make possible the transition from surface similarity to transitional as well as relational similarity. (2) Learners were able to generate a handful of transitional similarities. (3) The task used allowed learners to apply processes to base objects and target objects. (4) Pedagogical content knowledge of the instructor on analogies was essential in assisting learners to build worthwhile analogical conjectures that were acceptable. These results indicate that analogical reasoning of teachers is better as compared to their inductive and deductive proof.

Markovits and Doyon (2011) conducted a study on how to use analogies as the bridge between known content and abstract reasoning. The first study was carried with 256 university students with an average age of 23. The second study was carried with 102 adults with an average age of 19 college students. The results from both studies indicated the following: (1) Concrete problems does not necessarily enhance abstract reasoning. (2) Students failed to reason abstractly because they have not mastered this form of reasoning. (3) Students who reasoned better with analogies did not necessarily transfer that to abstract reasoning. These results highlight the fact that analogical proof does not enhance abstract reasoning.

Figure 2.15 is an example that is based on analogies and in the marking guideline such answers were suggested as possible solutions. Learners needed to prove that triangles are congruent by identifying analogous sides and angles as in question 7.1; SQ (common side), $<Q_1$ and $<S_1$, and $<P_2$ and $<R_2$.


7. In the figure below the diagonals of parallelogram PQRS intersect at M.

7.1 Prove that $PQ = SR$ and $PS = QR$.

7.2 Prove that $PM = MR$ and $SM = QM$.

7.3 What can you deduce about the diagonals of a parallelogram?

Figure 2.15: ANA mathematics Grade 9 question 7 exemplar 2012 (DBE, 2012d)

❖ **CONJECTURING**

The NCTM (1989), Curriculum and Evaluation Standards for School mathematics argues as follows; '*Conjecturing and demonstrating the logical validity of conjectures* *are the essence of the creative act of doing mathematics'* (NCTM, 1989: 81). Cantlon (1998) describes a conjecture as conclusions that are derived from inconclusive evidence. Alcock and Inglis (2008) add that conjecturing is the process of writing conceived hypotheses that are yet to be justified. Cantlon (1998) suggested that conjectures are useful for making connections, promoting of sense-making, providing students with opportunities to construct mathematical knowledge and promotion of conceptual knowledge. Furthermore, conjecturing makes learner's respect and value students thinking as they are open to scrutiny and other learners may refute or confirm their conjectures (Aaron & Herbst, 2015; Alcock & Inglis, 2008). Reid (2002) suggested that conjecturing involves testing a rule, exploring its usefulness, and then modifying it or rejecting it.

In trying to explain a conjecture, Reid (2002) reflect on work done by Lakatos on fallibility and stated the following: (1) Mathematical activity is characterised by proofs and refutation. (2) Mathematics education is not as it used to be portrayed as a purely deductive science of accepted truth and theorems. (3) Rather, mathematics education is a human activity of providing examples to prove conjectures as well as counterexamples to refute conjectures. Reid (2002) also reflects on work done by Polya that in a mathematical activity, there is rigorous reasoning which generates conjectures that are formalised using logical inferences.

Cantlon (1998) suggested three useful uses of conjectures that she employed in her classroom; conjecturing to demonstrate mathematical power, conjecturing to make connections and conjecturing to develop the learning community.

- *'empower students by promoting ownership and inquiry,*
- *provide a means for students to construct mathematical knowledge, and*
- *foster opportunities for students to make connection.'* (Cantlon, 1998:108).

In her quest to show the power that conjectures give to students Cantlon (1998) shows some conjectures from her fourth graders as they responded to the following questions on fractions: *'(1) What are fractions? (2) How can you represent fractions? (3) When do you use fractions in your own life?'* (Cantlon, 1998: 110). In response to these questions, classroom discussions came with conjectures such as this one that was prompted by one student called Cody: $5 \div 10 = \frac{1}{2}; 10 \div 20 = \frac{1}{2} and\ 20 \div 40 = \frac{1}{2}$. When asked to explain this, Cody came up with more examples: $2 \div 4 = \frac{1}{2}$; $3 \div 6 = \frac{1}{2}; 4 \div 8 = \frac{1}{2} and\ 5 \div 10 = \frac{1}{2}$ and finally he wrote the conjecture: 'any number divided by twice itself equals $\frac{1}{2}$ except 0' (Cantlon, 1998: 111). This conjecture prompted other learners to use other examples to extend this conjecture, such as Bob, Nick and Kelly. Finally, learners had the following conjectures:

'(1) any number divided by twice itself equals $\frac{1}{2}$ except 0.

(2) Any number divided by three times itself equals $\frac{1}{3}$ except 0.

(3) Any number divided by four times itself equals $\frac{1}{4}$ except 0.

(4) Any number divided by five times itself equals $\frac{1}{5}$ except 0.

(5) Any number divided by six times itself equals $\frac{1}{6}$ except 0.

(6) Any number divided by seven times itself equals $\frac{1}{7}$ except 0.' (Cantlon, 1998: 111).

According to Cantlon (1998), these examples given by her students justified the fact that, conjectures demonstrates mathematical power, and this allows students to extend, accept or rebuke conjectures developed by their peers. This confirms that conjectures are powerful to stimulating reasoning in mathematics education.

Reid (2002) conducted a study with Grade five students where they conjectured general rules, tested the rules, explored the rules, rejected some of the rules and modified some of the rules. The study came up with the following findings: (1) through rigorous observation of patterns, students were able to come up with conjectures. (2) When testing these conjectures, students either rejected or confirmed the conjectures based on sufficiency of mathematical justifications. (3) Conjectures were generalised into logical inferences and further explored. These results confirm that conjecturing can act as a bridge for formal reasoning.

A study by Alcock and Inglis (2008) reviewed literature on the use of examples in proving conjectures made on abstract mathematics, which they called syntactic and semantic reasoning strategies. They used examples which were of the following kinds: (1) examples that were called *start-up* which provoke basic definitions and illuminate essential insights; (2) examples called *reference* examples, because in their nature they are basic and extensively pertinent in connecting findings and theories; (3) *model examples*, which are universal examples that propose and summarise findings and theories; and (4) counterexamples, which refute a conjecture and refine merits between theories.

These examples took two forms, *semantic* and *syntactic* reasoning. A semantic statement is when a single symbol is used to represent some class of objects which may require an individual to use a set of examples to be applicable in in generic situations (this is inductive). In syntactic reasoning, a proof is generated by using logic to provide systematised definitions of concepts which need to be proved or refuted using logic and definitions, $'If\ p_1\ and\ p_2\ are\ primes, then\ p_1p_2\ is\ not\ abundant.'$ (p115). Such a conjecture needs to be proved using logic and the knowledge of the definition of the term abundant (this is deductive).

2.3.5 Productive Disposition

Productive disposition is when learners are seen to be valuing mathematics as useful, rational and valuable, (Kilpatrick et al., 2001). Furthermore, if learners fully develop all the first four SMP, to these learners mathematics makes sense and they value mathematics because they are proficient in the procedures (Mellone, Verschaffel & Van Dooren, 2017). They produce quality connections between concepts and procedures. They have the ability to solve problems beyond the routine ones, and furthermore, they can produce a series of justifications that are valid and reliable (Karakoc & Alacaci, 2015; Groves, 2012). In their description of the SMP, Kilpatrick et al. (2001) describe productive disposition as follows:

> *'Productive disposition refers to the tendency to see sense in mathematics, to perceive it as both useful and worthwhile, to believe that steady effort in learning mathematics pays off, and to see oneself as an effective learner and doer of mathematics. If students are to develop conceptual understanding, procedural fluency, strategic competence, and adaptive reasoning abilities, they must believe that mathematics is understandable, not arbitrary; that, with diligent effort, it can be learned and used; and that they are capable of figuring it out. Developing a*

productive disposition requires frequent opportunities to make sense of mathematics, to recognize the benefits of perseverance, and to experience the rewards of sense making in mathematics.' (Kilpatrick et al., 2001: 131).

Various studies have focused on attitudes, anxiety and mathematics achievement (Cheung, 1988; Gierl & Bisanz, 1995; Hannula, 2002; Thomas, 2000), and these shed some light in some key issues that are raised by Kilpatrick et al. (2001) on this strand. Moreover, one of the specific aims of CAPS outlined by DBE (2011) is that learners must develop 'a spirit of curiosity and love for mathematics.' (DBE, 2011: 8). How then do learners develop the necessary dispositions that make them value and make sense of mathematics? Below are some studies that show how learners have developed dispositions.

Gierl and Bisanz (1995) carried out a study to determine the existence of mathematics anxiety and attitudes among Grade 3 and 6 learners. The findings of their study revealed that learners become more anxious as they progress through their schooling. Furthermore, the researchers found that older learners show more positive attitudes towards mathematics. Another study by Cheung (1988) which was carried out in Hong Kong revealed that, if students perceive mathematics as useful in their daily lives, this enhanced learners' consideration of mathematics as a creative subject. However, the study warned that positive attitudes cannot always determine mathematics achievement.

A study conducted by Thomas (2000) determined the influences of educational productivity factors on mathematics achievement and attitudes among Grade 8 ethnic groups. There were eight productivity factors investigated: *ability*, *motivation and attitude*, *quantity*, *quality*, *classroom*, *home*, *peers* and *use of out-of school*. Some comprehensive results emerged from this study. (1) Asians and whites performed better in mathematics as compared to African Americans. (2) Optimisation of the productivity factors enhanced mathematics achievement and had little effect on attitude towards mathematics. (3) Quantity and quality, which focused on engaging learners with many tasks for problem solving had a significant effect on concept building. Such findings reveal that the time spent on mathematics tasks enhances their conceptions and hence stimulates their dispositions.

Hannula (2002) conducted an analysis of a case study of a lower secondary mathematics learner, Rita, to develop a framework on emotions, associations, expectations and values related to mathematics. The results of the study revealed that the framework was useful in describing attitudes towards mathematics and how it changes over time. This confirmed the findings by Gierl and Bisanz (1995) that, as learners progress in their schooling, their attitude towards mathematics improves. Other key findings were that a negative attitude can be a defence mechanism for positive self-concept and can be used as a coping strategy. This was seen when Rita kept on referring to a mathematics problem as being not useful to her life when she failed to solve it.

Mueller, Yankelewitz and Maher (2011) carried out a study that examined commonalities in learning environment, classroom interactions and cognitive tools that enhanced positive dispositions. Their framework observed that key mathematical practices that developed positive disposition were, problem-solving, reasoning and proof, real-world connections, multiple representations and mathematics connections. Most of these are elements are the SMP which Kilpatrick et al. (2001) urged should be developed in order to support productive disposition. In their study, Mueller et al. (2011) justified the need to develop intrinsic motivation which they said it originates from learners' interests and autonomy to a given task that allows them to be self-reliant and make sense of their solution strategy. Such tasks allowed learners to demonstrate sense-making through higher order reasoning and to defend mathematical solutions.

The results from the study by Mueller et al. (2011) revealed the following: (1) Learners developed confidence as they justified their solutions especially in the presence of their peer without the authority of the teacher. (2) Students developed some sense of ownership of such justifications (3) Motivation that was coupled with positive dispositions towards mathematics resulted in mathematical reasoning which enhanced understanding. (4) Learners relied on reasoning as opposed to memory of facts to convince their conceptions and that of other learners about what made sense.

Some important constructs emerge from this study, sense-making, confidence as well as enhanced understanding as key features of dispositions.

Martin and Kasmer (2010) advocated for reasoning as a way of making sense. This justifies the view of Kilpatrick et al. (2001) that to develop productive disposition the other SMP must be fully developed. Martin and Kasmer (2010) described reasoning as a process of reaching a conclusion through evidence or preconceived assumptions. Furthermore, reasoning allows learners to formulate conjectures, justify their claims, make predictions and reach a meaningful generalisation. They described sense-making as the development of understanding of a mathematical concept or context through connecting that with existing knowledge. This is an element of conceptual understanding, and is one of the SMP. Reflecting on the strands of mathematical proficiency, Martin and Kasmer (2010) insist that for learners to make sense in mathematics, they must master conceptual understanding and adaptive reasoning.

Lee and Chen (2015) carried a study with one hundred Grade 5 mathematics learners in Taipei. The purpose of the study was to investigate how manipulatives enhanced learning performance as well as attitude of fifth Grade learners towards mathematics. They referred to the attitude towards learning mathematics as: (1) learner’s perceptions of mathematics resulting from experimental teaching; (2) learner’s enjoyment of mathematics; (3) learner’s motivation to do mathematics; and (4) learners’ anxieties that resulted from studying mathematics. These were measured against basic flexible thinking skills and advanced thinking skills. The findings of their study revealed that learning with virtual manipulatives had the same learning effect on mathematics as physical manipulatives. Virtual manipulatives can increase learning enjoyment more effectively than physical manipulatives.

Jones et al. (2008) discussed the utility of mathematics in the context of viewing mathematics as a second language. Although they acknowledged that mathematics is more than just a language, mathematics involves meaning-making of symbols that do not have the same flexible meaning as ordinary language. Jones et al. (2008) equated proficiency in the language of mathematics to the acquisition of the Polya’s framework of problem solving: (1) understanding the problem; (2) devising a plan; (3) carrying out a plan; and (4) looking back. For the first step Jones et al. (2008) saw it as a stage where a novice learner the rules of the language (in this case mathematics language), the structure, definitions, reading of sentences in relation to a context. Jones et al. (2008) also suggest some proficiency levels that come as a result of learning mathematics language. These are: *zero level of mathematics*; at these level learners struggle to articulate basic mathematical conventions such as naming objects, mathematical argumentation and mathematics anxiety. The second is *very limited mathematics*; learners attempt conversations with fear, seldom initiate mathematical arguments and answer short phrased questions. *Limited mathematics*; learners easily respond to short phrased questions, cannot elaborate answers in thought provoking questions and seldom initiate mathematical conversations. *Limited fluency in mathematics*; here learners articulate mathematical phrases with ease, often initiate conversations and switch freely between languages used in conversations. Furthermore, fluency in mathematical language develops positive dispositions when learners freely articulate mathematics meaningfully.

This section on SMP concludes by reflecting on a narrative by Wheeler (2001) on a concept of '*Mathmatisation as a Pedagogical tool'*. This is a relevant determinant of the notion of Kilpatrick et al. (2001) on productive disposition that it develops as a result of a mastery of the other four SMP. Wheeler (2001) refers to mathematisation as the processes that exist mentally and its product is mathematics and this comes with an understanding of these processes. The author located a mathematics activity within mental operations and took cognisance of some useful discursive practices within the process of mathematisation. Mathematisation is described as '*putting a structure on a structure'.* The structure consists of holistic mathematics processes to mention a few; '*language, notation, graphical representation and imagery*.' This is a process that involves awareness of the structure, getting along with the structure, formulating hypotheses concerning the structure, proving hypotheses of the structure as well as generating new ideas about the structure. Finally Wheeler (2001) summarises this concept as follows;

"1. The educator must be able to reorganise substantial portions of the content of the mathematics curriculum so that they can be mathematized. Anyone who has tried to make a mathematical film knows how the mathematical content has to be detached from its normal context and considered afresh to see how it might be apprehended 'from scratch'. 2. The educator must be able to select suitable 'proxy' experiences which indicate to the students how certain situations have been or could be mathematised. The example of the cube given earlier in this paper is an instance. History provides another resource- the teacher might reconstruct, for example, how Leibniz mathematised the problem of finding the sum of an infinite geometric series. 3. The teacher must be able to take advantage of the spontaneous events of the classroom that will occur when students are given the freedom to employ their own mathematising abilities. Indeed, this is a function that only the classroom teacher can perform. One of the best descriptions of the process by an aware teacher is given in Mathématiques sur mesure by Madeleine Goutard" (Wheeler, 2001: 52).

Although this narrative by Wheeler (2001) is in the context of teaching, which is not the focus in the current study, it shed a great deal of light to the dispositions that may be observed in textual analysis of test questions as well as learners' mathematics responses to those questions. Using this perception on pedagogy is useful for a critical analysis of learners' responses that they exhibit in response to mathematics tests.

Schoenfeld (2007) gave an illustration of what he called beliefs and dispositions by using an arithmetic problem from the 1983 National Assessment of Educational Progress (NAEP) as shown below:

'An army bus holds 36 soldiers. If 1128 soldiers are being bussed to their training site, how many buses are needed?' (Schoenfeld, 2007: 69).

According to Schoenfeld (2007), the solution to this problem is clear, dividing 1128 by 36 the answer is 31 with a remainder of 12. This implies that 32 buses are needed to carry all soldiers. In 45000 students that responded to this problem, the results were as follows:

'29% gave the answer 31 remainder 12, 18% gave the answer 31, 23% gave the correct answer, 32, 30% did the computation incorrectly' (Schoenfeld, 2007: 69).

Reflecting on these results, it means that (70%) of all the students calculated correctly, however (23%) of the students rounded up correctly. The question that arises here is, how can (29%) of the total students (which is 10350 students) give an answer that the number of buses needed a remainder? What does a remainder mean when the context of the problem is buses? The experience of the researcher in school mathematics education both as a teacher and head of department, in the past gives an idea about the causes of this problem. Learners often believe that mathematics does not make sense to them and mathematics involves only working with symbols, then learners will always produce responses that do not make sense such as the remainder of buses. Hence learners’ beliefs about mathematics is essential and such believes are developed by learners in n their classroom learning experiences (Schoenfeld, 2007).

The review of literature on productive disposition illuminates some constructs that the current study explores for data analysis and these are: (1) *sense-making* (using context, representations and reasoning to make sense of mathematics); (2) *utility of mathematics* (using mathematics to solve real-life problems; (3) *valuing mathematics* (usefulness of mathematics to solve complex problems); and (4) *enhanced understanding* (problems that need learners to persevere characterised by creativity and innovation).

2.4 Conclusion

In this chapter, the literature on systemic assessment, nationally, regionally and internationally was reviewed. Also, policy issues on systemic assessments have been set out as well as the cognitive levels in South African mathematics. Literature on calculating the Porter’s alignment index was reviewed. The SMP are adopted to be compatible with document analysis. The next chapter outlines the research methods, data collection, research process and data collection which have been done to respond to the purpose of the study.

3. CHAPTER THREE
RESEARCH METHODOLOGY AND PROCEDURE

3.1 Introduction

The previous chapter reviewed the literature on systemic assessments and adapted a theoretical framework, the SMP. This chapter presents the research methods and design used to collect and analyse data. It begins with a comprehensive discussion of the research methods, followed by the sampling procedure and a narrative of the data collection and analysis process. The qualitative data collection and mixed methods data analysis within the exploratory sequential paradigm are explained. The research ethics, validation and verification of data are set out. Finally, the challenges, strengths and limitations of this study are presented.

3.2 Research Methodology

This study used a mixed methods research approach within the transformative paradigm (Creswell, 2014) to document epistemological obstacles that ANA poses to systemic testing in South Africa (AMESA, 2012; Graven & Venkatakrishnan 2013). The mixed methods approach integrates quantitative and qualitative approaches by simultaneously using both qualitative and quantitative data (Gay et al., 2014). The mixed methods approach was appropriate for this study to coherently address the epistemological obstacles posed by the Grade 9 ANA mathematics systemic tests as well as document learners' responses to these tests (Harrits, 2011; Luyt, 2012). As described by Creswell (2014) mixed methods has a variety of designs, in this study I used the exploratory sequential design which begins with the process of gathering qualitative data to explore a phenomenon, followed by the collection of quantitative data that examines the patterns found in the qualitative data (Gay et al., 2014; McMillan & Schumacher, 2014). The research process in mixed methods happens in two phases (Creswell, 2014).

In the context of this study, the Grade 9 mathematics ANA question papers, the Grade 8 TIMSS 2011 question items as well as the learners' responses to the 2014 ANA test provided documents for an initial qualitative analysis. According to McMillan and Schumacher (2014), documents are records of events of the past which are either written or printed. In this study the documents took the form of the Grade 9 ANA mathematics question papers, Grade 8 TIMSS 2011 mathematics test items and the learners' scripts.

Firstly the Grade 9 ANA mathematics question papers were analysed using a theoretical framework developed from the SMP (Kilpatrick et al., 2001). These results were used to generate qualitative results. Secondly, the learner's responses were analysed using codes and themes that came from the developed theoretical framework and were classified in four categories: correct, partially correct, no response and incorrect. Last, the qualitative results from the matrices of content and cognitive demands were analysed further to yield quantitative results.

Document analysis approach was selected to develop coherent insight in the SMP that the Grade 9 ANA mathematics question papers examined as well as those that learners exhibited in response to ANA testing (Gay et al., 2014). The data from the aligning of the ANA with the TIMSS (Howie 2003), together with the data from the analysis of the question paper as well as the data from the learners' responses allowed presented different perspectives of viewing epistemological obstacles that could have been experienced in the ANA testing. These were also viewed as alternative ways of reporting results of national systematic assessment results.

3.3 Exploratory Sequential Design

An exploratory sequential design was used (Creswell, 2014), where the first phase of collected qualitative data is followed by a second phase which collects quantitative data (Creswell, 2014). Gay et al., (2014) also explain that this design weights qualitative data more than quantitative data (Table 3.1). The exploratory sequential design is appropriate for this study because of the following. First, qualitative data was generated from a thick description of SMP examined by ANA. Subsequently, the analysis of generated quantitative data, frequencies, means and mean deviations. Secondly, qualitative data came from learners' responses to ANA questions. Subsequently, the analysis generated quantitative data, frequencies, means and standard deviations. Lastly, qualitative data was from a description of content and cognitive levels of ANA tests and TIMSS response items. Subsequently, the analysis generated Porter's alignment index, means and mean deviations. The design is summarised in Table 3.1

Table 3.1: A summary of the research design

RESEARCH DESIGN		
Qualitative		*Quantitative*
Part one Phase one From ANA questions code SMP tested, themes emerge.	Builds to	**Part one Phase two** Generate descriptive statistics, means and mean deviations to explore SMP examined by ANA and mathematics cognitive levels in three consecutive years.
Part two Phase one Assess learners' responses to find proficiency levels.	Builds to	**Part two Phase two** Generate descriptive statistics, means, mean deviations and standard deviations to explain proficiencies exhibited by learners in their responses to ANA.
Part three Phase one Using content and cognitive levels in ANA and TIMSS develop matrices.	Builds to	**Part three Phase two** Calculate the Porter's alignment index of ANA and TIMMS, generate descriptive statistics, means and mean deviations to explain content and cognitive levels in corresponding cells.

3.4 Sampling Procedure

The following sampling procedures were applied to address each research questions of the study: The first research question for the current study is: *How are cognitive levels of mathematics tested by ANA reflective of SMP?* Convenience sampling is a procedure that focuses on those who are available at a certain time (Springer, 2010). The 2012, 2013 and 2014 ANA Grade 9 mathematics test papers were available data (McMillan & Schumacher, 2014) hence the question papers were conveniently sampled due to their availability. More specifically, the suggested answers in the marking guidelines were used as they provided various SMP tested by ANA. Subsequently, the suggested answers in the content areas were used as follows: 1) numbers, operations and relations (n = 48); 2) patterns, functions and algebra (n = 73); 3) space and shape (n = 52); and 4) measurement n = 18 and data handling and probability (n = 22) The second research question; *what levels of mathematical proficiency do learners' exhibit in response to the ANA tests?* When setting up the time lines for this study, the analysis of learners' scripts was planned for the year 2014. During this period the available learners' scripts were the 2014 scripts hence I used convenience sampling. According to Creswell (2014), convenience is when a researcher samples respondents that are accessible at that time. Seven schools, in the Limpopo Province, Capricorn District, were accessible due to their location and so these schools were conveniently sampled. Among the seven schools that were sampled, n = 1250 scripts that had responses to most questions and were conveniently sampled for the purpose of analysis (Springer, 2010). Finally, three questions, i.e., question three, question six and question ten in the 2014 learners' scripts, were purposively sampled, (Gay et al., 2014), due to their representativeness of the content in the 2014 ANA. The third research question for the current study is: *How does the content and cognitive levels tested by ANA compare with TIMSS?* Test papers that were readily available were the 2012, 2013 and 2014 (DBE, 2012a) and conveniently sampled. The TIMSS 2011 Grade 8 response items were conveniently sampled due to access having been granted (Creswell, 2014, McMillan & Schumacher, 2014). Table 3.2 provides a summary of the profile of the schools.

Table 3.2: Profiling of participating schools

Profiling of schools accessed for the 2014 Grade 9 mathematics ANA scripts			
Schools	*School settings*	*Script numbers*	*Number of scripts sampled*
School A	The school is in a suburb in Polokwane city, Limpopo Province. It is a government school formerly private school before1994. Staff members are predominantly black, with a few white and Indian.	SA1-215	215
School B	The school is in the Mankweng circuit 40km east of Polokwane City, Limpopo Province.	SB1-212	212
School C	The school is in the Bahlaloga circuit 40km west of Polokwane City, Limpopo Province.	SC1-50	50
School D	The school is in the Mankweng circuit 30km east of Polokwane city, Limpopo Province. The majority of the learners are Sepedi speaking with a few Tsonga. The staff is all black.	SD1-189	189
School E	The school is in the Mankweng circuit 34km east of Polokwane city, Limpopo Province.	SE1-349	349
School F	The school is in a suburb of Polokwane city, Limpopo Province. It is a government school built after 1994 during the democratic government.	SF1-163	163
School G	The school is in the Koloti circuit 45 km east of Polokwane City, Limpopo Province.	SG1-72	72
Total			**1250**

3.5 Assumptions of the Study

Epistemologically and philosophically the researcher accepts that research is subjective and needs to be experienced, deduced and described, hence the current study adoption of the mixed method for this study. The current study is informed by the transformative paradigm in the context of the epistemological transformative assumptions. A transformative paradigm is a framework that allows researchers to prioritise addressing inequalities in society and the promotion of social justice (Mertens, 2010b). The epistemological assumptions are that, the relationship between the researcher and stakeholders is centred on the need to create accurate knowledge (Mertens, 2010a). Within the transformative paradigm, the epistemological assumptions are that this study must contribute substantially to change in national systemic assessment (Mertens, 2007).

Various studies on systemic assessments have reported that in South Africa, rich schools have the ability to produce numerate learners and rich parents afford quality education whilst poor schools struggle to produce numerate learners and this is where poor parents can afford to send their children (Howie, 2003 & 2004; Kotze & Strauss, 2006; Spaull, 2010). Furthermore, rich provinces such as Gauteng and Western Cape perform well in systemic assessments, internationally, regionally and nationally (DBE, 2012a, 2013b & 2014a; Howie, 2004; Spaull, 2010).

This study proposes a different approach in reporting the results of national systemic assessment testing in South Africa. These issues are: (1) the use of SMP; (2) analysing the SMP that learner's exhibit in response to the ANA tests; and (3) aligning the ANA with other systemic assessments such as the TIMSS, to view the content message that ANA pose. The assumptions of this study is that ANA examine a spectrum of knowledge, skills and values. These assumptions originate from philosophical assumptions on what mathematical proficiency is and how it can be tested as suggested by Schoenfeld (2007).

3.6 The research Process

The consideration of the research questions was planned systematically and according to the anticipated data.

3.6.1 Data Collection

Three sets of data were collected to respond substantially to the research problem. First, ANA question papers were available from the website of the DBE Three consecutive papers were accessed, the 2012 (n=59 questions), 2013 (n=62 questions) and the 2014 (n=61 questions) Grade 9 ANA mathematics question papers (total N=182 questions). Second, the TIMSS 2011 Grade 8 mathematics test that had 90 questions were accessed from the internet. Thirdly, the 2014 ANA mathematics learners' scripts (n=1250) were accessed from sampled high schools in the Capricorn District, Limpopo Province. Only question three, six and ten were used in the analysis, as these were representative of the content covered in 2014 ANA (DBE, 2014a). The exploratory design employed in this study meant the data was predominantly qualitative. Hence, in the context of the qualitative paradigm, there is a particular way of ensuring reliability and validity of instruments (Springer, 2010). Qualitative researchers often approach research in a consistent manner with other researchers, to ensure reliability (McMillan & Schumacher, 2014). In contrast, when ensuring validity, qualitative researchers use certain research procedures to ensure the accuracy of their findings (Creswell, 2014). One of those is triangulation which involves the use of a variety of procedures to explore a single phenomenon (Zohrabi, 2013). Consequently, first, the data collected was secondary data and in addition, theory, SMP, had to be used to justify the conceptual analysis (Kimberlin & Winterstein, 2008).

Researchers often select existing instruments that measure the current phenomenon that is being researched (Gay et al., 2014). As such, in the context of this study, secondary data was collected, and so most of the reliability and validity focused on the analysis. First, the table used to capture the codes of SMP in ANA questions were adapted from a study by Dhlamini and Luneta (2016). Second, the table used to capture learners' responses to ANA was adapted from Luneta (2015b). Third, the tables used to document content and cognitive levels during the calculation of the alignment index were adapted from Porter (2002). To ensure validity, qualitative researchers often triangulate data from a variety of sources to confirm findings with similar results (Zohrabi, 2013). As such, in this study, data from the ANA questions, learners' responses and the Porter's alignment were triangulated.

3.6.2 Data Analysis Strategies

The process of data analysis started soon after the 2012, 2013 and 2014 ANA Grade 9 mathematics tests and 2011 Grade 8 mathematics TIMSS response items were accessed. Firstly, the analysis of the 2012, 2013 and 2014 ANA question papers started after accessing the question papers. Secondly, data analysis continued when scripts were accessed from the schools. Lastly, the alignment index was calculated between the 2012 Grade 9 ANA mathematics test and the Grade 8 mathematics TIMSS response items, the 2013 Grade 9 ANA mathematics test and the Grade 8 mathematics TIMSS response items and the 2014 Grade 9 ANA mathematics test and the Grade 8 mathematics TIMSS response items. Documents as records of events were reproduced either by individuals or groups (Cohen, Manion & Morrison, 2011). The trends and patterns that exist in documents are interpreted in a process called document analysis (Creswell, 2014).

❖ ANALYSIS OF ANA QUESTION PAPERS

The analysis of the ANA question was to generate data that responded to the following research question: *How are cognitive levels of mathematics tested by ANA reflective of SMP?* First, the ANA questions for 2012, 2013 and 2014 were organised into five content areas as stipulated in the CAPS document (DBE, 2011) and these were: 1) Numbers, operations and relations; 2) Patterns, functions and algebra; 3) Space and shape; 4) Measurement; and 5) Data handling and probability.

Second, SMP were explored in the ANA questions and subsequent questioning was coded as per the emerging SMP from each question (see Table 3.3 for codes). Consequently, Table 3.3 is a synopsis of the generic codes that emerged. Subsequently, specific codes that emerged from the content areas of ANA and how these emerged are outlined in the next chapter.

Third, these codes were documented in strands which was informed by the theoretical framework, SMP, that they are intertwined, interconnected, inseparable and interwoven (Kilpatrick et al., 2001). Axial coding was followed which allowed this researcher to categorise the emerging codes in strands that outline coherent mathematical activity in each question item (Gibbs, 2012).

Last, the codes were categorised according to their relationship into themes. Subsequently, axial coding allows emerging codes to be matched against some hypotheses, in this case SMP. Hence, the codes in Table 3, the code SP captured interconnected codes for categories of the following themes: *simple procedures*, SP for coding of categories for procedures that neither tested computations nor algorithms (Schoenfeld, 1985) for the theme; *procedural fluency*, PF1-2 denoted coding for categories of fluency in computations and sequence of steps (NCTM, 2000) for the theme, *procedural fluency*, CU1-2 for categories of comprehending concepts, computations and algorithms the theme, *conceptual understanding*, SC1-3 categories for problem formulation, problem representation and problem solving (Granberg, 2016; Land, 2017; Stein, Grove & Henningsen, 1996), for the theme (strategic competence), AR1-3 categories for logical thought, explanation and justification for the theme adaptive reasoning. Most studies (Dhlamini & Luneta, 2016; Graven & Stott, 2012; Maharaj et al., 2015) used the first four SMP except for one study (Graven, 2012). However, that study was limited to mathematics questions that resulted in the development of dispositions. This study closes that gap by using learners' responses to examine productive dispositions that learners develop as they respond to sampled ANA question. The categories of sense making, utility of maths and valuing mathematics emerged from the analysis of ANA question and were coded as PD1-3, for the theme *productive disposition*. Subsequently, these categories were coded separately for the identified content areas due to the conceptualisation by Kilpatrick et al. (2001) conception that the strand productive disposition results from learners' proficiency from the other strands. Conversely, learners who are not proficient in the other strands do not develop productive disposition (Graven, 2012). To respond succinctly to the first research question, the emerging themes were matched with the NAEP to show clearly the mathematics cognitive levels and levels of complexity (Berger, et al., 2010) posed by the ANA testing in the three consecutive years.

Table 3.3: Generic codes for mathematical proficiency

Code	Meaning
SP	The question requires the learner to write the procedure which requires no calculations (simple procedure).
PF1	The question requires some systematic computations to reach the required answer (computation).
PF2	The question requires a sequence of steps of computations, procedures and relations (algorithm).
CU1	The question requires learners to comprehend a variety of mathematical concepts to reach the required answer (conceptual connections).
CU2	The question comprises of two computations, first compute a value that is subsequently used in the computation required by the question (computational connections)
SC1	The question used familiar problems for learners in the grade (routine procedures).
SC2	The question used a diagram or context to represent concepts, procedures and relations (provision of multiple representations).
SC3	The question requires learners to recall already known procedures, concepts and relations to solve the problem (reproductive thinking).
AR1	The question allows learners to give reasons for their answers, which gives them the opportunity to reflect on their solutions and navigate through concepts procedures and relations (mathematical reasoning).
AR2	The question allows learners to make inferences that are subject to acceptance or rejection (conjecturing).
AR3	The question allows learners to invent suitable commonalities of mathematical relations to make a proof (an analogy).
PD1	A mathematical problem is useful to make sense through the use of diagrams and representations (sense making).
PD2	Using mathematics to solve real life problems (utility of mathematics).
PD3	A mathematical problem is useful and important in solving a reasoning or thought provoking problem, which is regarded as a complex problem (valuing mathematics).

- **SCRIPT ANALYSIS**

The three questions that were sampled from the 2014 Grade 9 ANA mathematics questions were analysed using document analysis and the instrument used was adapted from Luneta (2015). There were a total of n=1250 scripts that were analysed by exploring SMP exhibited by learners and categorising them in the following four variables; *correctly answered*, *partially answered*, *incorrectly answered* and *no response*. The assumption of the current study was as follows. (1) A *correct response* in line with the marking guidelines is enough to justify that the learner was proficient in the question and the learner fully exhibited the SMP that the ANA posed. (2) A *partially answered* question only justifies that a learner still requires additional assistance in the SMP that the ANA examined. (3) An *incorrect response* is enough to justify the fact that the learner is not proficient in the SMP that the ANA examined. (4) A *no response* may mean that the learner skipped the question because of lack of proficiency in the SMP that the ANA examined but on the other hand, a no response might also be as a result of not finishing answering the questions which may be the result of time allocated for the test or slow pace of a learner.

The current study uses the standard deviation, mean bar graph and radar to present findings. The use of these descriptive statistics to present and interpret data has been widely accepted in research (Gorard, 2005; Lathrop, 1961; Saary, 2008). Bar graphs are often limited to displaying frequencies (Saary, 2008). However, the standard deviation is useful in instances where there is a need to measure variance of data from a comparable point (Lathrop, 1961). The use of mean deviations has more advantages than the standard deviation (Gorard, 2005) as it easier to understand and suitable for distributions that may have minute errors. As such, the trustworthiness of using radar has been documented and justified in terms of its competence, popularity, recency, corroboration and proximity (Nurse, Agrafiotis, Creese, Goldsmith & Lamberts, [Sa]). Additionally, radar are useful in presenting multivariate data (Feldman, 2013; Saary, 2008).

❖ CALCULATING THE ALIGNMENT INDEX

In calculating the Porter's alignment index, firstly, there was a need to analyse the cognitive levels as well as the content messages conveyed by the Grade 9 mathematics ANA test papers and the 2011 Grade 8 mathematics TIMSS test response items. Secondly, matrices were formed and the hits in the cells were documented using a protocol. This was done to calculate the Porter's alignment index. The question totals in the ANA papers were: cell X_i the 2012 matrix with n=59 questions; cell X_j the 2013 matrix with n=62 questions; cell X_p the 2014 matrix with n=61 questions; and cell Y_i, the TIMSS matrix with n=90 questions. These matrices are shown in Chapter 4, the presentation of findings.

3.7 Ethical Considerations

Research that engages human participants has some methodological as well as ethical challenges (Creswell, 2014). The current study involved schools and learners, and so procedures used to collect and analyse data had to be ethical. Permission was sought and received from the Limpopo Department of Basic Education (Appendix C). The sampled schools were approached before accessing the learner's scripts, four principles of research ethics were discussed with the principals, informed consent, confidentiality, trust and, risks and benefits (McMillan & Schumacher, 2014). Detailed information on these principles is outlined below.

3.7.1 Informed Consent

In addressing ethical issues, Bournot-Trites and Belanger (2005) pointed out that participants need to be informed about the nature of the study so that they decide on their participation (DuBois, 2002). Firstly, in the current study, the question papers were available on the website of the DBE, hence copyrights were upheld. One of the copyrights was to use the question papers for educational purposes and not for business purposes (McMillan & Schumacher, 2014). The current study is educational and not meant for business. Secondly, when accessing learners' scripts from schools, I first met with the principal in each school and explained that the analysis of learners' responses to ANA was meant to view learners' responses to the test and that the school had the right to withdraw from participation.

3.7.2 Confidentiality, Anonymity and Safety in Participation

Participants in a research study expect that information given to the researcher must be treated with confidentiality (Gay et al., 2014). To ensure confidentiality, it is essential to keep names of learners in the scripts concealed throughout the study to ensure their anonymity in participating in the study (McMillan & Schumacher, 2014). In the current study, fictitious names of schools as well as learners were used. Furthermore, I notified the participants about the research ethics that the University of Limpopo uphold.

3.7.3 Trust

Researchers often have the obligation to treat participants with respect and the expectation is that the researcher exercise intrusiveness (Creswell, 2014). In the context of the current study, I should build trust with the principals of participating schools that the scripts were safe during the analysis. Furthermore, I communicated the research process and the reporting of the anticipated results to the principals.

3.7.4 Risks and Benefits

The participants in a study have reasonable expectation that they must be safe and not be harmed as a result of participation (DuBois, 2002). Additionally, participants expect to benefit from the results of a research study (Bournot-Trites & Belanger, 2005). Subsequently, the results of this study are not meant to name and shame participating schools. Consequently, the results will assist the DBE with data on the effectiveness of ANA. Finally, I informed the schools that after the study has been finished, the final research report will be available at the Limpopo Department of Education.

3.8 Quality Criteria of the Study

Various measures were undertaken in the current study to ensure that the results would be trustworthy. In this instance, I used several measures to such as, using available data, adopting research instruments and maintaining the required ethical standards for carrying research with humans. All these elements of the current study are contained in what Creswell (2014) calls approaches to qualitative study and these are, *credibility*, *transferability*, *confirmability* and *dependability*. Below, I elaborate on these approaches and how the current study engaged within its context.

When ensuring *credibility*, the researcher determines rich accounts of the research process by gathering different types of data for the purposes of triangulation (Morrow, 2005). A series of processes were undertaken to triangulate methodologically to ensure that the data were valid. Firstly, through the use of SMP, qualitative data was generated in both the analysis of ANA questions and learners' responses. Secondly, the qualitative data was explored further using descriptive statistics, means, mean deviations and standard deviations. Lastly, when calculating the Porter's alignment index, a question by question analysis of the ANA question papers as well as the Grade 8 TIMSS question paper was undertaken to generate qualitative data. Subsequently, the qualitative data was explored to calculate the Porter's alignment index that was interpreted using descriptive statistics, means and mean deviations (Porter, 2002).

Transferability refers to strategies followed by the researcher follow to ensure that generalisations are made on a subject (Krefting, 1991). The current study, followed mixed methods in the context of the exploratory design, meaning it is dominated by qualitative data. Qualitative research uses a small sample. Generalisation is not the focus and in-depth understanding of the phenomenon is paramount (Creswell, 2014). Therefore, it is safe to report that the current study is more descriptive in nature and the small sample used, especially the learner's scripts did not warrant generalisations. Hence the current study does not ensure transferability.

Dependability refers to the succinct research process to make it consistent over time and with various researchers (Morrow, 2005). This study ensured dependability by the application of the audit trail through rigorous explanation of the coding of data and the emerging categories and themes using SMP. As such, the process ensured that the findings maintained internal validity due to the low inference descriptors shown in the coding process.

Confirmability refers to the inference that research is never objective but rather subjective in terms of the situation under inquiry and avoidance of researcher bias (Hofstee, 2015) as the means of maintaining internal validity. One way of avoiding bias (Gay et al., 2014) is to be impartial as much as possible during the data collection, analysis and interpretation. In the context of this study, first, most of the instruments were adapted from previous studies (Dhlamini & Luneta, 2016; Luneta, 2015b, Porter, 2002). Second, there was rigorous use of SMP, the theoretical framework, in the instrumentation and when coding question papers and learners' responses to make the findings acceptable.

3.9 Challenges and Strengths of this Study

There were various challenges faced that I faced in the current study. Firstly, in terms of the research methods, the challenge was on the development of instruments for document analysis. Secondly, data collection was characterised by many problems. Most schools that were initially sampled could not give the current researcher access to the ANA scripts, citing reasons of scripts being lost or disposed of already. In some of these cases, it was obvious that the schools were simply refusing access. Thirdly, during data analysis some learners' responses required further information from the learners. Since the ANA was written in 2014, it was impossible to locate and interview learners who wrote the test. Lastly, the topics examined in Grade 9 mathematics ANA have not been researched which means there is dearth of literature.

The strength of the study lay in the contribution in the topic as follows; the model, adapting the SMP, reporting using proficiency levels and the methodologies used. Also, the current study adapted SMP and made them to be compatible with document analysis. Additionally, the current study reported learner achievement in Grade 9 ANA using proficiency levels instead of mere aggregated scores. Finally, the current study interpreted its findings from two mixed methods (Creswell, 2014). According to Luyt (2012) and Harrits (2011), mixed methods research is gaining popularity due to the variety of topologies it employs.

3.10 Conclusion

The research methodology, research design, data collection, sampling and research process have all been addressed in this chapter. Additionally, the chapter outlined research ethics that the researcher upheld during the research process. Finally, the quality criteria as well as challenges and strengths of the current study have been ascertained. The next chapter presents the findings of the current study and furthermore, the results are interpreted.

4. CHAPTER FOUR
PRESENTATION AND INTERPRETATION OF FINDINGS

4.1 Introduction

The previous chapter outlined the research process, design and methodology. Additionally, a detailed account of data collection and analysis was also outlined. Subsequently, this chapter presents the findings of the current study and its interpretation. The literature is used to discuss findings.

4.2 Results and Discussion for ANA Question Papers

The question papers were arranged in content areas for the purpose of exploring SMP that the questions examined. Such content areas were: 1) Numbers, operations and relations; 2) Patterns, functions and algebra; 3) Space and shape (geometry); 4) Measurement; and 5) Data handling and probability. Mathematical content is the core of all assessment, testing what the students learn (Greenleess, 2011). Subsequently, mathematical content forms the fundamental basics for what learners are taught throughout their schooling and this is often regarded as standards used by parents to determine progress of their children achievement (NCTM, 2000). In this analysis, the suggested answers in marking guidelines were used to explore SMP coherently in answers and alternative answers. The SMP that emerged from the suggested answers to ANA questions were coded (see Table 3.3) and from the resultant codes, categories of SMP were clustered to document themes of SMP that were reflective of mathematics cognitive levels. Subsequently, axial coding was employed to classify the emerging codes in strands that outline coherent SMP in each ANA question (Gibbs, 2012). To manage the coding, and abide by the notion on SMP raised by Kilpatrick et al. (2001), that SMP are intertwined, interconnected and interwoven, the codes were first categorised as procedural and conceptual. Further, classification of the codes yielded themes. Below the findings are presented for the analysis of question papers and the succinct discussion using literature and the theoretical framework, i.e. the SMP.

4.2.1 Numbers, Operations and Relations

This content area, numbers, operations and relations focuses on mathematical representation and the use of models for quantity, tallying, magnitude, order and approximation of calculations (Greenleess, 2011). As an evaluative assessment, ANA must provide useful data to gauge how curricula implementation is succeeding in this discourse (DBE, 2012a; Graven & Venkat, 2014). The results below on the analysis of ANA question papers provide an outline on basic numeracy proficiencies.

The use of SMP in the analysis of ANA points out that SMP are intertwined, interconnected and interwoven as claimed by Kilpatrick, et al. (2001). Earlier, research (Dhlamini & Luneta, 2016) pointed out that SMP are divided into three categories which are: 1) Knowledge (procedural fluency and conceptual understanding); 2) skills (strategic competence and adaptive reasoning; and 3) values (productive disposition). First, in this analysis I categorise the emerging SMP using knowledge, procedural and conceptual (Table 4.1). Second, skills required to answer the questions are coded with the relevant knowledge as pointed out by Khashan (2014) that skills are normally embedded in the knowledge. As such the codes appear in strands for example, PF1-SC1-SC3 (see Table 4.1). The code PF1 (as explained in Table 3.3) depicts procedural knowledge posed by the question and the subsequent codes (SC1 & SC3) this explain the nature of the problem and the problem solving strategy required to solve the problem. These codes were placed as chains to show that they are connected and intertwined. In following this trend in the analysis, three categories emerged in the exploration of SMP in this content area, and these were: 1) simple procedures, 2) computations, and 3) algorithms

❖ **SIMPLE PROCEDURES (SP)**

Some mathematics questions require learners to write procedures that neither demand computations nor algorithms, just recall and write the procedure (NCTM, 2000). Such procedures lack fluency. The NCTM (2000) describes fluency as the ability to execute procedures and algorithms accurately, appropriately and efficiently. Consequently, in this analysis, procedures that emerged with these qualities in this content area, were categorised as simple procedures (see Table 4.1).

Table 4.1: Codes of SMP in Numbers, Operations and Relations

ANA examination questions	**Codes of SMP in Numbers, Operations and Relations**	
	Procedural	Total
2012 ANA	2SP-SC1-SC3 1PF1-SC1-SC3 5PF1-SC1-SC2-SC3 4PF2-SC1-SC2-SC3	12
2013 ANA	2SP-SC2-SC3 3PF1-SC1-SC3 4PF2-SC1-SC2-SC3	9
2014 ANA	3SP-SC1-SC3 1SP-SC1-SC2-SC3-AR1 6PF1-SC1-SC3 7PF1-SC1-SC2-SC3	17
Totals	**48**	**48**
Percent	**100**	**100**

There were two categories of simple procedures in the ANA questions in this content area which were: 1) those coded SP-SC1-SC3; and 2) SP-SC1-SC2-SC3-AR1. An example of a simple procedure in the first category of ANA questions is question 2.1 in the 2012 ANA. The question is as follows: "*write 0.00000356 kl in scientific notation*" (DBE, 2012c: 5), one of the two coded SP-SC1-SC3 in the 2012 ANA (Table 4.1). The code 'SP' entails that this question required the use of the rule of writing numbers in scientific notation as "3.14×10^{-6}" which does not require a computation or algorithm. The two additional codes, SC1 & SC3, for the code 'SC1', the question requires familiar knowledge to Grade 9 of converting to scientific notation (DBE, 2011). Subsequently, the code 'SC3' depicts that the question tested a reproduction of number conversion to scientific notation (DBE, 2011).

An example of a question coded in the second category of simple procedure is question 8.4 in the 2014 ANA. The question is as follows;

"*Study the table below*

The length of a side of a square in *cm*	2	3	4
Area of the square in cm^3	4	9	16

Is this an example of a direct or indirect proportion? Give reasons for your answer" (DBE, 2014b: 12)

This question was coded SP-SC1-SC2-SC3-AR1 (Table 4.1), a simple procedure, which, additionally to the first category, are the codes, "SC2 and AR1". The code 'SC2' refers to the use of the table to assist in sense making. However the table does not affect the difficulty of the question. The code 'AR1' refers to inferences that need to be made regarding the proportion (a conjecture). These inferences and reasons may require approval for confirmation as true (Amir-Mofidi et al., 2012).

❖ **COMPUTATIONS (PF1)**

Computational fluency refers to accurate, efficient, and flexible consolidated mathematical computations (Russell, 2000). Subsequently, fluency in computations involves being fluent in basic operations such as addition, subtraction, multiplication and division, using rules and procedures in a particular content area (Schoenfeld, 1985). According to Kilpatrick et al. (2001), computational fluency is a category of procedural fluency.

There were two categories of computations that emerged from the analysis of ANA questions. First, there was a computation that was coded PF1-SC1-SC3 (Table 4.1). An example was question 2.2.3 in the 2014 ANA as follows: "$\frac{3\times5^9}{5^7}$", source (DBE, 2014b: 5). The question needed computation using laws of exponents (PF1), the code 'SC1', depicted that the question was familiar to Grade 9 and the code 'SC3' coded the recall of known laws of exponents.

Second, there was a category of computations that were coded PF1-SC1-SC2-SC3 for routine computations. An example of category question 6.1 was as follows; *"How long will it take to travel 432 kilometres at an average speed of 96 kilometres per hour?"* (DBE, 2013c: 10). The code 'PF1' refers to the computation of time using the given distance and speed. The code 'SCI' shows that the question is familiar in the grade and subsequently, the code 'SC3' shows that the question demands the recall of knowledge and the formula for calculating speed, then substitute given values. The additional code 'SC2' shows that the question used the context of distance, speed and time for the utility of mathematics to solve a real life problem as it allowed learners to extract mathematics from a, situation (Lee & Chen, 2015) and this did not have an effect on the complexity of the problem.

- **ALGORITHMS (PF2)**

An algorithm must be characterised by accuracy and generality (Bass, 2003). For accuracy, the algorithm must always produce the solution when correctly used. Subsequently, for generality, the question must be generic, meaning it must compute any class of the problem (Bass, 2003). There are other desirable qualities of an algorithm as follows: for the algorithm to be used by a machine it keep track with computation speed, which posits efficiency. However, since most mathematical problems are solved by humans, algorithms need to be less prone to error when used effectively (easy to use) (Russell, 2000). Additionally, they need to be transparent, meaning, the steps involved in solving the problem must advance calculations of the required solution (Bass, 2003). There was one category of algorithms coded PF2-SC1-SC3 that emerged in the 2012 and 2013 ANA questions during the analysis using SMP (Table 4.1). An example is question 6.3 which was as follows; *"Calculate simple interest on R3500 invested at 6% per annum for 3 years."* (DBE, 2013c: 10). This question requires substitution in the formula for simple interest to compute the value of the simple interest (PF2). This formula is generic such that it may compute any value of simple interest (generality). Subsequently, the problem is routine (SC1) and requires the reproduction of knowledge of calculating simple interest (SC3). As a consequence, when simple interest is computed correctly (Bass, 2003), the desired solution will always be reached (accuracy). Hence this question meets the first two qualities for it to be an algorithm and test procedural fluency (Kilpatrick et al., 2001). Additionally, other desired qualities of an algorithm, the computational speed must be fast, as is in this algorithm (efficiency), and then it may be programmed in machines (NCTM, 2000). And, since the algorithm must be used and learned by humans, (Bass, 2003) its effective use of calculating simple interest does not lead to high frequency of error (ease of use), and the steps of the problem advance calculations of simple interest (transparency).

4.2.2 Patterns, Functions and Algebra

This content area refers to functions and relations applied to routine and mathematical figures, patterns in numerical and spatial, and overall forms (procedures, formulae, tables, graphics, equations and equivalences) conveyed using words, symbols or figures (Greenleess, 2011; Harel, 2017).

Again, the analysis in this content area focused on the exploration of SMP that the ANA tests examined in three consecutive years, 2012, 2013 and 2014. Subsequently, in this content area the mathematics tested by ANA as explored, was coded in strands. The use of SMP points out that SMP are intertwined, hence the codes appear in strands (Table 4.2). Using knowledge to manage the analysis, three categories, 1) simple procedures, 2) computations and 3) algorithms, as coded, emerged from the exploration of SMP from the ANA questions.

❖ **SIMPLE PROCEDURES (SP)**

In this content area, simple procedures have emerged from the analysis of ANA questions. As stated in the previous content area, simple procedures tested only a quick recall and mention of procedures without computations or algorithms, hence again they were categorised as simple procedures. In patterns, functions and algebra, there were three categories of simple procedures that emerged from the analysis and were coded, SP-SC1-SC3, SP-SC1-SC2-SC3 and SP-SC1-SC3-AR1 (Table 4.2). For the first category, an example of that was coded SP-SC1-SC3 is question 3.1 from the 2013 ANA. The question was as follows; *"Factorise fully* $6a^3 - 12a^2 + 18a$*"*

(DBE, 2013c: 7). The question is a familiar factorisation (SC1) which required the recall of knowledge of algebraic factors (SC3). There were no computations in finding the factors or the use of an algorithm (SP).

Table 4.2: Codes of SMP in Patterns, Functions and Algebra

ANA examination questions	**Codes of SMP in Patterns, Functions and Algebra**		
	Conceptual	Procedural	Total
2012 ANA	2CU1-SC1-SC3	13SP-SC1-SC3 1SP-SC1-SC2-SC3 12PF1-SC1-SC3 1PF2-SC1-SC3	29
2013 ANA	3CU1-SC1-SC3	6SP-SC1-SC3 1SP-SC1-SC2-SC3 1SP-SC1-SC3-AR1 1SP-SC1-SC3-AR2 7PF1-SC1-SC2 2PF1-SC1-SC2-SC3 1PF2-SC1-SC3	22
2014 ANA	8CU1-SC1-SC3	14SP-SC1-SC3 1SP-SC1-SC3-AR1 2SP-SC1-SC2-SC3 8PF1-SC1-SC3 1PF1-SC1-SC2-SC3 1PF2-SC1-SC3	35
Totals	**13**	**73**	**86**
Percent	**15.1**	**84.9**	**100**

For the second category, an example that was coded SP-SC1-SC2-SC3 is question 7.1.1 which had an additional 'SC2' code and was as follows; *"Write down the coordinates of the points A,B and C in the table."* DBE, 2013c: 11). The table and the graph were given (SC2), but no computations were required, just reading from the graph (SP) and the graph was familiar to the grade (SC1) which required learners to recall of knowledge of linear functions (SC3).

For the third category, an example was question 5.2 from the 2013 ANA, coded SP-SC1-SC3-AR1 (Table 4.2), which was as follows: *"Write down the general term T_n of the above sequence."* (DBE, 2013c: 9). For learners to write the general term of the sequence, they do not need computations (SP), the question is routine (SC1) and requires recall of knowledge of sequences (SC3). However, writing the general term needs learners to base reasoning on the given sequence (AR1).

- **COMPUTATIONS (PF1)**

Again in this content area computations emerged in ANA questions, a category for procedural fluency (Kilpatrick et al., 2001). Consequently, these computations are in number patterns, algebraic expressions, equations and graphs. There were two categories of computations that emerged during the analysis of ANA questions and were coded PF1-SC1-SC3 and PF1-SC3-SC2-SC3 (Table 4.2). An example of the first category is question 5.2.2 in the 2012 ANA which was coded PF1-SC1-SC3 and the question was as follows: *"The lines intersect at T. Show by calculation that the co-ordinates of T are* $x = 1$ *and* $y = 1$ *or* $(1; -1)$*."* (DBE, 2012c: 12). This question is a follow-on from the graph drawn in question 5.2.1 (SC2). The solution requires learners to compute the co-ordinates of the point of intersection (PF1) by equating given equations (SC1) to show knowledge of intersecting lines (SC3)

An example of the second category is question 7.2.1 in the 2013 ANA which was as follows: "*Draw the graphs defined by* $y = -2x + 4$ *and* $x = 1$ *on the given set of axes. Label the graph and clearly mark the points where the lines cut the axes.*" (DBE, 2013c: 12). The solution to this problem requires learners to write the domain (x co-ordinate) for both functions then compute (PF1) the range (y-co-ordinates) of the given equations (SC1) using knowledge of linear equations (SC3). Subsequently, using those values, draw the graphs (SC2).

- **ALGORITHMS (PF2)**

In patterns, functions and algebra, there were questions that had a sequence of steps with specialised qualities of an algorithm (Bass, 2003). An example was question 4.4 in the 2012 ANA which was coded PF2-SC1-SC3. This question follows on from the conjecture in question 4.3 and using that conjecture is generic in computing any number of terms (generality). When correctly used it will always produce the desired number of terms (accuracy), then it qualifies as an algorithm (PF2). In addition, effective use of the conjecture may not lead to high frequency of error (ease of use) because it requires the reproduction of known calculations of required term (SC3). The familiar steps (SC1) advance calculations of the required term (transparency).

Hence it may be used by humans. Last, the algorithm is easy and fast to compute the required term (efficiency) then it may be programmed in machines.

❖ **CONCEPTUAL CONNECTIONS (CU1)**

In this content area, conceptual connections that emerged from the analysis refers to a connection of concepts in a single ANA question (NCTM, 2000). Concepts are discursive meanings that learners ascribes to a mathematical term (Khashan, 2014). As a consequence, for learners to be proficient in mathematical concepts, they need to exhibit quality connections and in the conceptual aspect of mathematics (Mhlolo et al., 2012; Mwakapenda, 2008). There was one category of conceptual connections, coded CU1-SC1-SC3 (Table 4.2) that emerged from the analysis of the 2012, 2013 and 2014 ANA questions in this content area and an example is question 3.3 from the 2014 ANA which was as follows: *"Simplify each of the following expressions. The denominators in the fractions are not equal to zero.* $\frac{x^2-4x}{x^2-2x-8}$*."* (DBE, 2014b: 6). The solution to this question required learners to factorise (SC3) a familiar (SC1) numerator which was a binomial, factorise a familiar denominator to Grade 9 (DBE, 2011) which is a trinomial and divide like terms, three distinct concepts comprehended (CU1) in one problem (Kilpatrick et al., 2001).

4.2.3 Space and Shape (Geometry)

This content area, geometry requires the ability to analyse features of geometric figures and make mathematical inferences about the geometric relationship, and use visualisation, spatial reasoning, and geometric modelling to solve related problems (DBE, 2011; NCTM, 2000). Geometry is a regular area of mathematics for the advancement of learners' reasoning and proof skills (Greenleess, 2011: Otten, Bleiler-Baxter & Engledowl, 2017).

In this content area, again there is the use of SMP which posits that SMP are intertwined (Kilpatrick et al., 2001) hence, the codes are represented in strands (Table 4.3). Two categories of SMP emerged: 1) simple procedures, and 2) computations. These are explored with illustrations below.

Table 4.3: Codes of SMP in Geometry

ANA examination questions	Codes of SMP in Space and Shape (Geometry)	
	Procedural	Total
2012 ANA	2SP-SC1-SC3 2SP-SC1-SC2-SC3 1SP-SC1-SC2-SC3-AR2 1SP-SC1-SC3-AR1 6SP-SC1-SC2-SC3-AR3 3PF1-SC1-SC2-SC3 7PF1-SC1-SC2-SC3-AR1	22
2013 ANA	6SP-SC1-SC2-SC3 5SP-SC1-SC2-SC3-AR1 2SP-SC1-SC2-SC3-AR3 1PF1-SC1-SC3	14
2014 ANA	5SP-SC1-SC3 4SP-SC1-SC2-SC3-AR1 4SP-SC1-SC2-SC3-AR3 3PF1-SC1-SC2-SC3-AR1	16
Totals	**52**	**52**
Percent	**100**	**100**

❖ **SIMPLE PROCEDURES (SP)**

In this content area, simple procedures have emerged from the analysis of ANA questions. As stated in the previous content areas, simple procedures tested only a quick recall and mention of procedures without computations or algorithms, hence again they were categorised as simple procedures. There were six categories of simple procedures in geometry. The first category is coded SP-SC1-SC3, and an example is question 9.1.1 from the 2014 ANA which was as follows; *"$\widehat{D}$ and $\widehat{F}$ are complementary angles if*_______" (DBE, 2014b: 13). The question requires learners to recall knowledge of complementary angles (SC1) and state that their sum is 90^0 (SC3) without computations (SP).

The second category is coded SP-SC1-SC2-SC3 An example is question 9.2 from the 2013 ANA and phrased as follows; *"Write down the coordinates of B', the image of B"* (DBE, 2013c: 17). This question required learners to read from the graph (SC2), using basic knowledge of the Cartesian plane (SC1) the coordinates of the image (SC3).

The third category is coded SP-SC1-SC3-AR1, with an example, question 6.2 from the 2012 ANA and was as follows: *"State which triangle is congruent to $\Delta ABC.$"*

(DBE, 2012c: 14). The question required learners to use their basic knowledge of congruency to state (SP), giving reasons and infer (AR1), using given triangles (SC2) that the triangles are congruent (SC3).

The fourth category is coded SP-SC1-SC2-SC3-AR1. An example is question 8.1.1 from the 2013 ANA which was phrased as follows; *"Calculate with reasons: The size of* $\hat{T}_1$*."* (DBE, 3013c: 13). The question required learners to state the size of the angle (SC1), without calculations (SP), and giving reasons recalling the knowledge (SC3) of the relationship of angles of an isosceles triangle (AR1) with the aid of the given diagram (SC2).

The fifth category of simple procedures is coded SP-SC1-SC2-SC3-AR2. An example is question 6.3.4 from the 2012 ANA, which questioned as follows; *"Hence, state the relationship between AE and BC."* (DBE, 2012c: 16). The question required learners to infer (AR2) on how two lines relate in a given diagram (SC2) based on proofs (SC3) in prior questions (SC1).

The last category of simple procedures in geometry is coded SP-SC1-SC2-SC3-AR3. An example was question 8.3 from the 2013 ANA, which was as follows; *"Prove with reasons that* $\Delta KNQ \equiv MPQ$*"* (DBE, 2013c: 15). The question required learners to use knowledge of congruency (SC3) to identify (SP) with reasons, relations of corresponding angles and sides (AR3), in a given (SC1) pair of triangles (SC2) to prove that they are congruent. This was an analogy, and analogical reasoning refers to the ability to identify similar structural commonalities of objects (Amir-Mofidi et al., 2012; Lee & Sriraman, 2011; Whitacre et al., 2017) and mostly it is regarded as moderate reasoning strategy due to the low level of rigour that is without computations. Analogical reasoning (Markovits & Doyon, 2011) is an essential recipe for bridging the gap between concrete and abstract reasoning. Proving is an essential part of mathematics to convince oneself that inferences when made when solving mathematical problems, such as proving that theorems are in fact true (Bleiler-Baxter, 2017).

❖ **COMPUTATIONS (PF1)**

In this content area, as in the previous content areas, there were three categories of computations that emerged from the analysis of the ANA questions coded PF1-SC1-SC3, PF1-SC1-SC2-SC3 and PF1-SC1-SC2-SC3-AR1 respectively.

An example for the first category is question 1.9, coded PF1-SC1-SC3 and from the 2013 ANA which was as follows: *"In the figure below, side DF of ΔEDF is produced to C. Calculate the size of $\hat{E}$ in terms of x."* (DBE, 2013c: 4). The question required learners to compute (PF1) using the given diagram (SC2) and the recall of knowledge of properties of a triangle (SC3) the value of x (SC1).

An example for the second category is question 7.3, coded PF1-SC1-SC2-SC3 and from the 2012 ANA which was as follows: "The length of each side of figure *P* is halved. Calculate the perimeter of the new figure." (DBE, 2012c: 18). The question required learners to divide each side (SC1) of a given figure (SC2) and compute (PF1) the perimeter using their knowledge of perimeter (SC3).

An example for the third category is question 9.3 coded PF1-SC1-SC2-SC3-AR1 and from the 2014 ANA was as follows; *"In $\Delta ABC, AB = AC$ and $\hat{C} = x^0$. Determine the size of $\hat{A}$ in terms of x."* (DBE, 2014b). The question required learners to compute (PF1) the size of an angle in terms of a variable, giving reasons (AR1), using a given diagram (SC2) and recall of knowledge of properties (SC1) of an isosceles triangle (SC3).

4.2.4 Measurement

This content area teaches learners qualities, units, structures, and processes of measurement and to apply techniques, tools, and formulae to determine measurements (NCTM, 2000). Subsequently, measurement can serve as a system that coherently comprehends various SMP because it affords learners the opportunity to learn and apply other areas of mathematics such as: Numbers; Space and Shape; Functions; Statistics and Probability (DBE, 2011; Greenleess, 2011).

In this content area, again the use of SMP which posits that SMP are intertwined (Kilpatrick et al., 2001) hence the codes are represented in strands (Table 4.4). Subsequently, four categories of codes emerged: 1) simple procedures, 2) computations, 3) algorithms, and 4) computational connections. These are explored with illustrations below.

Table 4.4: Codes of SMP in Measurement

ANA examination questions	**Codes of SMP in Measurement**		
	Conceptual	Procedural	Total
2012 ANA		1SP-SC1-SC3 1SP-SC1-SC2-SC3 3PF2-SC1-SC2-SC3	5
2013 ANA		4PF1-SC1-SC2-SC3 3PF1-SC1-SC2-SC3-AR1	7
2014 ANA	1CU2-SC1-SC3	1SP-SC1-SC3 5PF1-SC1-SC2-SC3	7
Totals	**1**	**18**	**19**
Percent	**5.3**	**94.7**	**100**

❖ **SIMPLE PROCEDURES (SP)**

In this content area, there were procedures that did not need computations or algorithms that emerged from the analysis of ANA question papers (see Table 4.4). There were two categories of SMP coded SP-SC1-SC3 and SP-SC1-SC2-SC3 respectively (Table 4.4). For the first category, an example is question 1.7, a multiple choice question from the 2012 ANA and the question was as follows; *"The volume of a cube with side of length 7 cm is, A) 49cm^3, B) 28cm^3, C) 343cm^3, D) 14 cm^3."* (DBE, 2012c: 3). The question does not necessarily require learners to compute the volume (SP). Learners must recall known formula (SC1) and identify the value of seven cubed (SC3).

For the second category, an example is question 8.1 from the 2012 ANA which was as follows: *"Complete the table by filling in the name of the 3-D figure, the number of faces, the number of vertices, the number of faces and the shape of the faces."* (DBE, 2012c: 19). A table was provided with the shape with 'SC2' and required learners to recall knowledge of a cylinder and use this to identify its properties (SP).

❖ COMPUTATIONS (PF1)

Again, from the analysis of question papers using SMP, there were computations that emerged and were in two categories. The first category of computations was coded, PF1-SC1-SC2-SC3. An example is question 10.2.3 from the 2013 ANA phrased as follows: *"Hence calculate the area of PQR."* (DBE, 2013c: 19). The question requires learners to calculate the area of a triangle (PF1) using already calculated values (SC3) and knowledge of calculating area (SC3) with the aid of a given triangular prism (SC2).

The second category of computations in measurement was coded PF1-SC1-SC2-SC3-AR1. An example is question 10.1.1 from the 2013 ANA which was as follows: *"Show that the area of the shaded ring is equal to* $\pi(R^2 - r^2)$*."* (DBE, 2013c: 18). The question required learners to derive (PF1) the known conjecture (SC1) with the aid of a given diagram (SC2). Subsequently, the question required learners to show that the area of the shaded is equal to the difference of the area of the outer circle and the inner circle (SC3) using the formulae (AR1).

❖ ALGORITHMS (PF2)

In this content area, there was one category of algorithms that was coded PF2-SC1-SC2-SC3. An example is question 8.2 from the 2012 ANA which was as follows: *"Calculate the total surface area of the rectangular prism with length= 7.2m breath = 5m and height = 3.32m. Give your answer correct to 2 decimal places."* (DBE, 2012c: 20). In the marking guideline, they use the formula, $2(l \times b) + 2(l \times h) + 2(b \times h)$ to substitute the given values to compute the surface area. The formula is generic and may be used to calculate the surface area for any given dimensions (SC1) of a prism (generality). Subsequently, the formula, when correctly used, yields the desired surface area (SC3) of a prism (accuracy). Additionally, the use of a diagram enhances sense making and does not affect the complexity of the question (SC2). Hence, it is an algorithm because it qualifies for generality and accuracy (PF2). Furthermore, substituting given values advances calculation of the desired surface area (transparency). Since the formula is aimed to be used by learners, its effective use of results in reduced error proneness (ease of use). Additionally, computational speed of using the algorithm is low, then it may be programmed in machines.

- **CONCEPTUAL CONNECTIONS (CU2)**

In measurement, one category of conceptual connections emerged from the analysis of ANA questions using SMP coded, CU2-SC1-SC3. An example is question 11.3 from the 2014 ANA which was as follows: *"The circumference of a circle is 52 cm. Calculate the area of the circle correct to 2 decimal places."* (DBE, 2014b). The question required learners to compute the area of a circle by first computing the radius, then use it compute the area of the circle (CU2). The question required recall of familiar knowledge (SC1) of calculating area of a circle and circumference (SC3).

4.2.5 Data Handling and Probability

In this content area, thinking statistically is necessary to educate citizens and consumers to be informed (NCTM, 2000). Subsequently, data handling and probability content standard requires learners to articulate questions and gather, organise, analyse and present pertinent data to respond to particular questions (DBE, 2011). In addition, it stresses knowledge of relevant statistical systems to analyse data, make inferences and appropriate predictions based on the data, and understand the use of fundamental concepts of probability (Greenleess, 2011).

In this content area, again there is the use of SMP which illustrates that SMP are intertwined (Kilpatrick et al., 2001) hence the codes are represented in strands (Table 4.5). Subsequently, two categories of SMP emerged: 1) simple procedures, and 2) computations. These are explored with illustrations below.

Table 4.5: Codes of SMP in Data Handling and Probability

ANA examination questions	Codes of SMP in Data Handling and Probability	
	Procedural	Total
2012 ANA	4SP-SC1-SC2-SC3 3PF1-SC1-SC2-SC3	7
2013 ANA	10SP-SC1-SC2-SC3 5PF1-SC1-SC2-SC3	15
2014 ANA		
Totals **Percent**	**22** **100**	**22** **100**

- **SIMPLE PROCEDURES (SP)**

In Probability and Data Handling, there were procedures that did not require computations nor algorithms. There was one category of simple procedures coded SP-SC1-SC2-SC3 that emerged from the analysis of ANA questions. Consequently, all simple procedures in statistics and probability had the code 'SC2' which means that contexts and diagrams were used to help in explaining statistical and probability questions. However, the use of the diagrams and contexts enhanced dispositions such as sense-making and the utility of Statistics and Probability in solving routine problems (Suurtamm, 2012) without affecting the complexity of the questions. An example is question 11.1 which was as follows: *"The histogram below illustrates the mathematics test marks, out of 10, obtained by a Grade 9 class."* (DBE, 2013c: 21). The question required learners to use a given histogram (SC2) to fill in missing values of the frequency table. This required routine knowledge (SC1) of reading values from a statistical graph (SC3).

- **COMPUTATIONS (PF1)**

In this content area, there were computations that emerged from the analysis of ANA questions using SMP. There was one category of computations coded PF1-SC1-SC2-SC3 and an example is question 11.3 from the 2013 ANA which was as follows: *"Calculate the mean test mark."* (DBE, 2013c: 22). The question required learners to calculate the mean (PF1) using routine statistical knowledge (SC1) by finding the quotient of the total data set and the total frequency (SC3), using the given marks in a histogram (SC2).

4.2.6 Themes of Strands of Mathematical Proficiency in ANA Questions

From the analysis of ANA questions, three themes explicitly emerged, simple procedures, procedural fluency and conceptual understanding. As such, these themes were explicit in that, as I explore SMP tested by ANA, they emerged from the questioning. Subsequently, from the analysis, another theme implicitly emerged, and that is productive disposition. According to Kilpatrick et al. (2001), productive disposition emerges as a result of learners' proficiency in a particular strand, which is why they were implicit. Most of the dispositions emerged from questions that tested reasoning, used context and figures to promote sense making, the utility of mathematics and valuing mathematics. The data in Table 4.6, illustrates the themes that emerged from the analysis of ANA questions using SMP. The frequencies came from categories of the themes in Table 4.1, 4.2, 4.3, 4.4 and 4.5.

Table 4.6: Themes in ANA Questions

ANA Examination	Simple Procedures Frequency (%)	Procedural Fluency Frequency (%)	Conceptual Understanding Frequency (%)	Mean Frequency (%)	Total (%)
2012	34(45.33)	39(52)	2(2.67)	25(33.33)	75(100)
2013	34(57.63)	22(37.29)	3(5.08)	19.67(33.33)	59(100)
2014	34(45.95)	31(41.89)	9(12.16)	24.67(33.33)	74(100)

The data in Table 4.6 is explored in the discussion of the themes below. I use the data in Table 4.6, the percent, to generate the chart shown in Figure 4.1. The data in Figure 4.1 is explored in the discussion of the themes in the next sections.

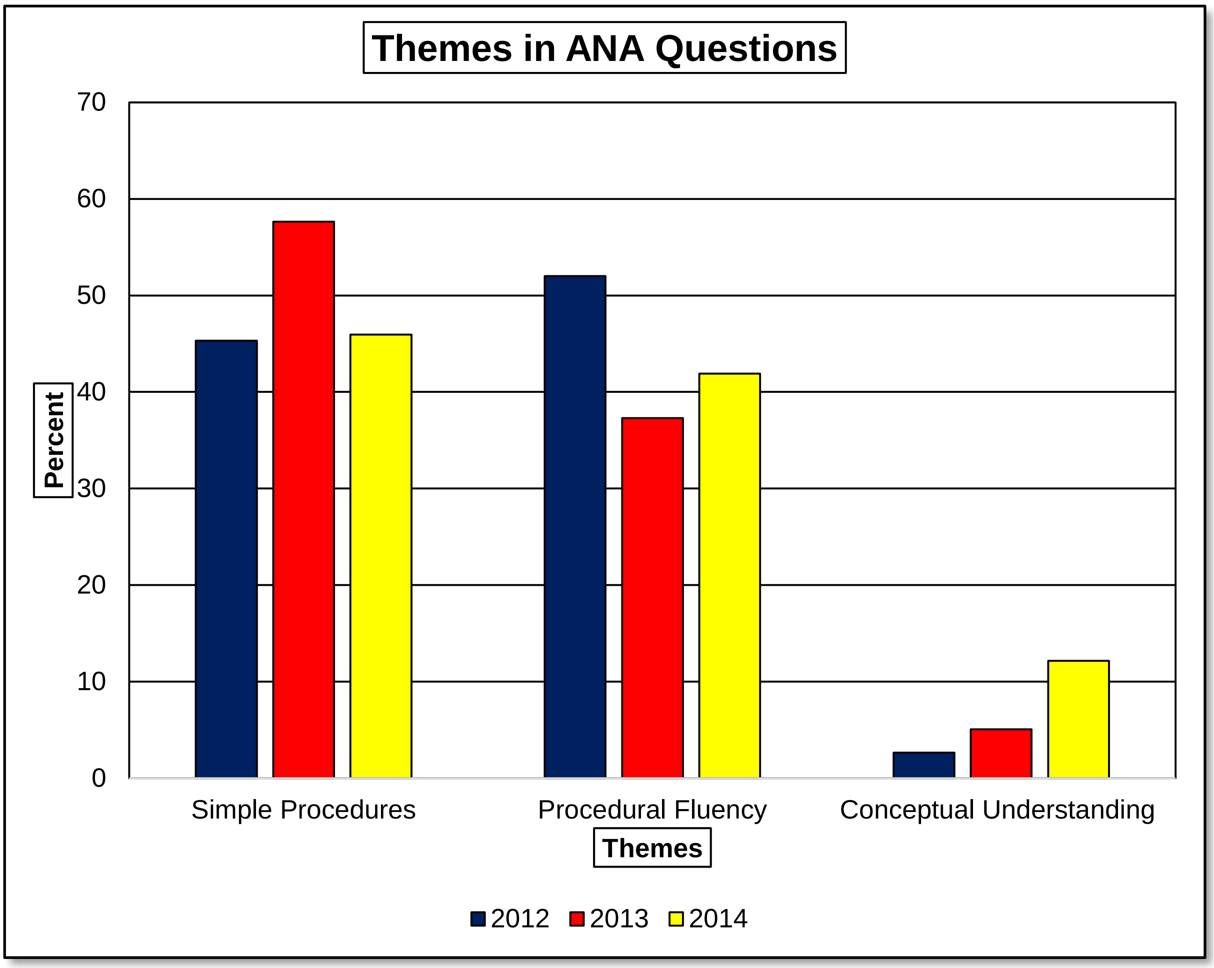


Figure 4.1: Themes in ANA questions

From the data in Table 4.6, mean deviations are generated from the frequencies and the resultant data is shown in Table 4.7. I use the data in Table 4.7 to explore the themes in the next sections.

Table 4.7: Mean discrepancies with direction

ANA Examination	Simple Procedures Frequency (%)	Procedural Fluency Frequency (%)	Conceptual Understanding Frequency (%)	Mean Frequency (%)
2012	9(12)	14(18.67)	-23(-30.67)	0(0)
2013	14.33(24.3)	2.33(3.96)	-16.66(-28.26)	0(0)
2014	9.33(12.62)	6.33(8.56)	-15.66(-21.18)	0(0)

From the frequency in the data in Table 4.7, I generate data in Figure 4.2, a synopsis of mean deviations in the ANA questions which is use in the next sections to explore the emerging themes.

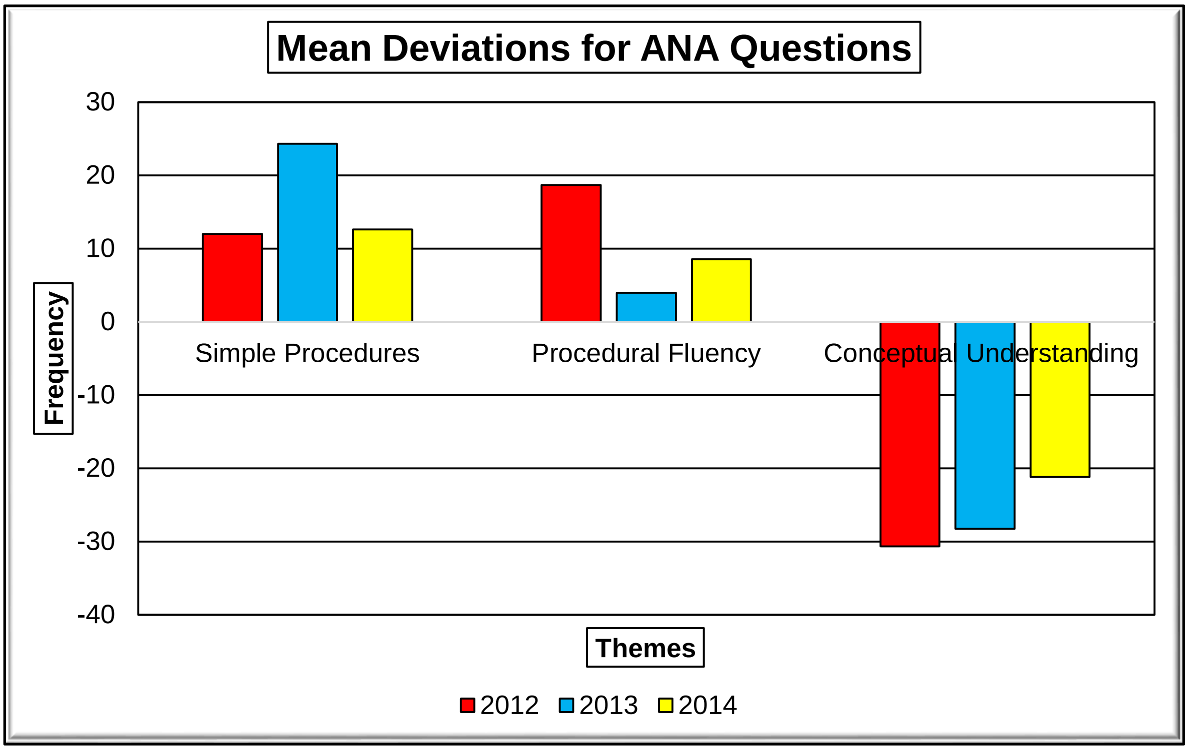


Figure 4.2: Mean Deviations for ANA Questions

THEME 1: SIMPLE PROCEDURES

The exploration of SMP in ANA questions using the knowledge strand, procedural fluency to categorise the codes, has seen various categories of procedures, one of them is named simple procedures. A simple procedure requires learners to state the procedure without computations (Schoenfeld, 1985). The data in Figure 4.1 illustrates that in the three consecutive years of ANA testing questions were characterised with simple procedures, an indication that the tests were of lower order thinking (Stein et al., 1996). Additionally, there was no consistency in the testing of these simple procedures in 2012, 2013 and 2014 (Figure 4.2), which poses reliability concerns in ANA testing (DFID, [S.a.]). National assessment with such qualities may not give a clear indicator of how the system is performing (Koretz, 2009).

There were seven categories of simple procedures that were coded as follows: 1) SP-SC1-SC3, 2) SP-SC1-SC2-SC3, 3) SP-SC1-SC3-AR1, 4) SP-SC1-SC2-SC3-AR1, 5) SP-SC1-SC3-AR2, 6) SP-SC1-SC2-SC3-AR2, 7) SP-SC1-SC2-SC3-AR3. The use of both SMP and the NAEP Taxonomy resulted in a dichotomy in matching the codes with emerging themes and mathematics cognitive levels (Berger et al., 2010; Kilpatrick et al., 2001). First, the dichotomy as observed pointed to contradictory mathematics cognitive levels when exploring further the simple procedures. Contrary to SMP, the NAEP Taxonomy stated that the use of justification meant a question may be categorised as high complexity. However, all ANA questions that required learners to justify, only demanded analogical reasoning (AR3). Research (Amir-Mofidi et al., 2012; Yopp, 2015; Zazkis, 2015) has indicated that analogical reasoning is a weaker form of proof. There are other forms of proof that were indicated when I adapted SMP, deductive and inductive reasoning (Stalvey & Vidakovic, 2015). The absence of such questions in the ANA questions deprived learners of the opportunity to use logic and evidence to make sense and reason at a higher order (Lee, 2016; Quinn et al., 2009). These involve the use of logic and empirical evidence to prove. I categorised analogical reasoning (AR3), mathematical reasoning (AR1) and conjecturing (AR2) as moderate complexity.

As a consequence, there were simple procedures that required learners to navigate through concepts, procedures and relations and give reasons for their answers (mathematical reasoning AR1) (Brodie, 2010). Additionally, there were simple procedures that required learners to make inferences (conjecturing AR2) (Aaron & Herbst, 2015). However, such simple procedures were few, with analogical reasoning emerging only from geometry. This discrepancy in questioning is worrisome, and the question is: Is geometry the only content area that uses analogies? Ironically, there are other content areas that use shapes, such as, measurement, graphs statistics that may be enhanced in their complexity through the use of proofs.

A strength is observed in simple procedures and procedural fluency, (Figure 4.2) which has questions that test recall of known procedures and reproductive thinking which are limited to low complexity and moderate complexity, as outlined by Berger, et al. (2010). A challenge for policymakers is that ANA questions did not have questions that examine metacognition and productive thinking (Roth et al., 2015). Such questions examine: (1) the effective use of knowledge of the problem to reach the required solution (Kim et al., 2013; Veloo et al., 2015); and (2) the use of inventive and creative ways to reach a solution to a problem (Guberman & Leikin, 2013). Both 1 and 2 are (high complexity). This is an indication that ANA did not focus on higher order questions, a pre-requisite for tertiary mathematics and tertiary natural sciences (Luneta, 2015b).

❖ THEME 2: PROCEDURAL FLUENCY

The analysis of ANA questions using SMP revealed two categories of procedural fluency: 1) computations; and 2) algorithms. For learners to exhibit procedural fluency, they must be proficient in either computational fluency or algorithmic fluency (Bass, 2003). During the analysis using SMP, there were various categories of procedural fluency and first, for computational fluency the emerging categories were coded as follows; 1) PF1-SC1-SC3, 2) PF1-SC1-SC2-SC3, and 3) PF1-SC1-SC2-SC3-AR1. Computational fluency refers to fluency in computation of mathematical procedures and relations (NCTM, 2000).

As explained during the explanation of codes, the code ‘SC2’ did not affect the complexity of the questions, instead, it assisted the development of dispositions such as sense making, valuing and utility of mathematics (Groves, 2012). The codes SC1 and SC3 refers to questions that required learners to recall knowledge of routine procedures (Sigley & Wilkinson, 2015). Such a structure of questions promoted lower order thinking (Land, 2017). Ironically, there was an absence of non-routine problems, those that required learners to perform lengthy steps of problem solving (Sullivan et al., 2016). Consequently, such absence of complex problems deprived learners of the opportunity to exhibit higher order reasoning such as metacognition and reproductive thinking (Guberman & Leikin, 2013). Learners may not exhibit higher order thinking justifying my assumptions, *“What You Test Is What You Get, (WYTIWUG).”* (Schoenfeld, 2007: 72). This state of ANA poses challenges to its validity, hence, ANA at its present state is not a valid form of systemic assessment (Kanjee & Moloi, 2014).

Second, for algorithmic fluency, two categories emerged that were coded as follows: PF2-SC1-SC3, and 2) PF2-SC1-SC2-SC3. (Graven & Venkat, 2014). An algorithm refers to questions that are generic and may compute any class of a particular problem (NCTM, 2000). Ironically, the computations were routine (SC3) and required productive thinking (SC3) which could not pose higher order thinking. In the second category, the additional code 'SC2' enhanced dispositions only and not the complexity of the question (Khashan, 2014). The qualities of an algorithm as shown in the explanation of codes, show that such questions are well-structured (Stein et al., 1996). Subsequently, a few algorithms emerged from the analysis of ANA questions using SMP (Table 4.1-5). A suggestion is that, ANA must test more computational and algorithmic fluency (Bass, 2003) to gauge how the system is performing (Graven & Venkat, 2014).

- **THEME 3: CONCEPTUAL UNDERSTANDING**

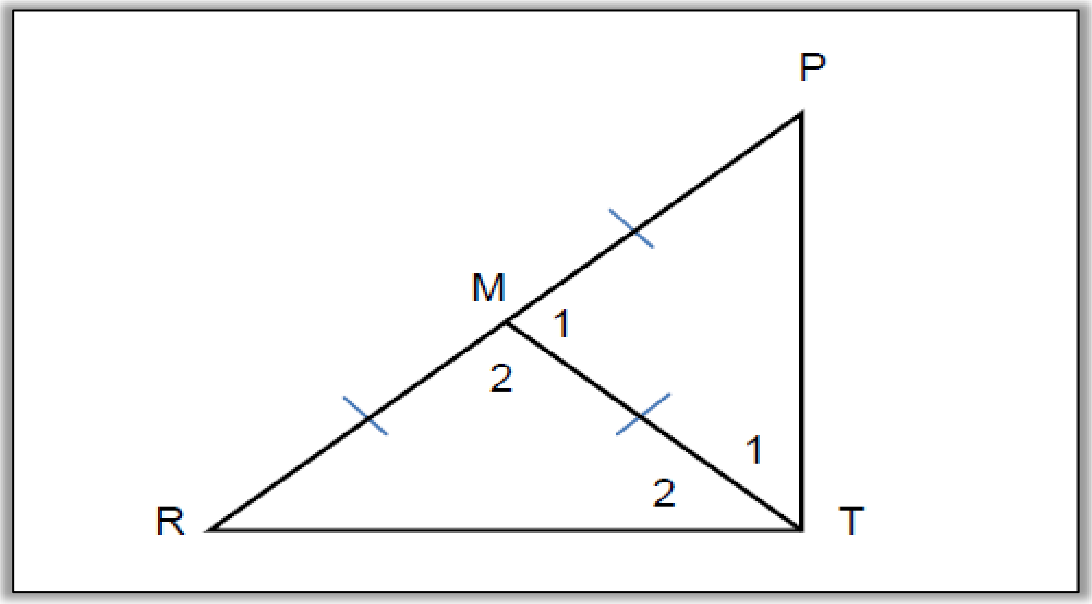
The analysis of ANA questions using SMP showed two categories of SMP as follows; 1) CU1-SC1-SC3 and 2) CU2- SC1-SC3. As explained in the previous section the codes 'CU1' and 'CU2' referred to questions that required learners to comprehend concept and computations in one solution strategy (Kilpatrick et al., 2001). Such questions afforded learners the opportunity to relate various concepts and computations meaningfully (Mwakapenda, 2008). Hence the data (Figure 4.2) justify the dearth of conceptual understanding and inconsistent testing of this theme in the three consecutive years. This presents serious reliability concerns in ANA testing, which is aimed to gauge how the system is performing (Kanjee & Moloi, 2014).

- **THEME 4: PRODUCTIVE DISPOSITION**

Productive disposition refers to the learners' ability to view mathematics as valuable, useful, sensible and worthwhile (Kilpatrick et al., 2001). To develop this SMP, the other four need to be fully developed (Groves, 2012). This means that this strand is made up of observable traits that learners exhibit as a result of being proficient in the other strands (Maharaj et al., 2015). During the analysis dispositions were captured from the code ‘SC2’ and in systemic assessment it is paramount to gauge dispositions. Most importantly, the use of SMP categorises dispositions as not being limited to learners’ feelings about mathematics (Khashan, 2014). It reveals how questions advance sense making, valuing and the utility of mathematics (Groves, 2012). During the analysis of question papers using SMP, the code ‘SC2’ emerged, depicting dispositions that learners develop as a result of engaging with such questions (Groves, 2012). Most of these were coded as strands and there were three categories of productive dispositions that emerged which were: sense making, utility of mathematics and valuing mathematics. These are explored in detail below. The dispositions that emerged were documented in Table 4.8.

First, *sense making,* involves mastery of situations, contexts and concepts by comprehending with existing knowledge (Suurtamm, 2012). Furthermore, conjecturing that is coupled with reasoning and proof enhances sense-making (Martin & Kasmer, 2010; Mueller et al., 2011). The ANA example is question 8.1.1 from the 2013 ANA which was as follows;

“In ΔPRT below, M is the midpoint of PR and MR = MT.



If $\hat{P} = 25^0$, calculate with reasons the size of $\hat{T}_1$.” (DBE, 2013c: 13)

This question uses a diagram for learners to make sense of the question being posed. Subsequently, learners need to give reasons for inferences they will make in their solution strategy. Learners develop sense making when solving questions of this nature. More than, half of the questions (Table 4.8) provided opportunities for sense-making and hence, from the results it is safe to pronounce that the ANA questions in the three consecutive years promoted sense-making. However, the inconsistency in the frequency (Figure 4.3) as observed poses reliability concerns now that ANA is used as systemic assessment (DFID, [Sa]).

Second, the utility of mathematics, as coded, referred to questions that allowed learners to solve real life problems using mathematics (Groves, 2012). Learners see mathematics as useful and applicable to everyday experiences as opposed to static and meaningful rules (Kilpatrick et al., 2001). The ANA example is question 3.3 from the 2012 ANA which was as follows; *"Bongiwe invests R12 000 in a savings account at 6.5% per annum compound interest. Calculate how much there will be in the savings account after 5 years."* (DBE, 2012c: 8). This question uses mathematics to teach learners processes of savings in banks. Learners see mathematics as a powerful tool to solve everyday problems (Schoenfeld, 2007). A smaller number of questions emerged in this category of productive disposition (Table 4.8), an indication that most questions were not connected to contexts and were perceived as static rules (Schoenfeld, 2007). Hence the effectiveness of the ANA in this instance is questionable.

Thirdly, *valuing mathematics*, when learners achieve in using mathematics to solve complex problems, they then value mathematics as powerful in problem solving (Lee & Chen, 2015). Additionally, such problems enhance perseverance that is coupled with creativity and innovation enhances well developed cognitive tools and promotes conceptual understanding (Mueller, et al., 2011). ANA example is question 11.3 from the 2014 ANA which was as follows: *"The circumference of a circle is 52 cm. Calculate the area of the circle correct to 2 decimal places."* (DBE, 2014b: 20). Learners to answer this question they need to compute the radius using the concept of the circumference. Next, they use the computed radius to calculate the area. The question is not explicit and requires learners to figure out such a solution strategy (Guberman & Leikin, 2013). The analysis has revealed a deficit in complex problems

(Table 4.8). The inconsistency observed (Figure 4.3) poses reliability challenges to ANA as a systemic assessment.

Table 4.8: Subthemes for Productive Disposition

ANA Examination	Sense Making Frequency (%)	Utility of Mathematics Frequency (%)	Valuing Mathematics Frequency (%)	Mean Frequency (%)	Total (%)
2012	23(57.5)	12(30)	5(12.5)	13.33(33.33)	40(100)
2013	20(62.5)	9(28.13)	3(9.37)	10.67(33.33)	32(100)
2014	16(72.73)	5(22.73)	1(4.54)	7.33(33.33)	22(100)

The data in Figure 4.3 was generated using the percentages in Table 4.8. The analysis of ANA questions using SMP revealed that sense-making was mostly visible in ANA questions in the three consecutive years. In contrast, the utility of mathematics and valuing mathematics were less respectively. This raised concerns in the structure of ANA (Stein et al., 1996). Ironically, the inconsistent testing (Figure 4.3) presents reliability issues in ANA as a systemic assessment (Graven & Venkat, 2014).

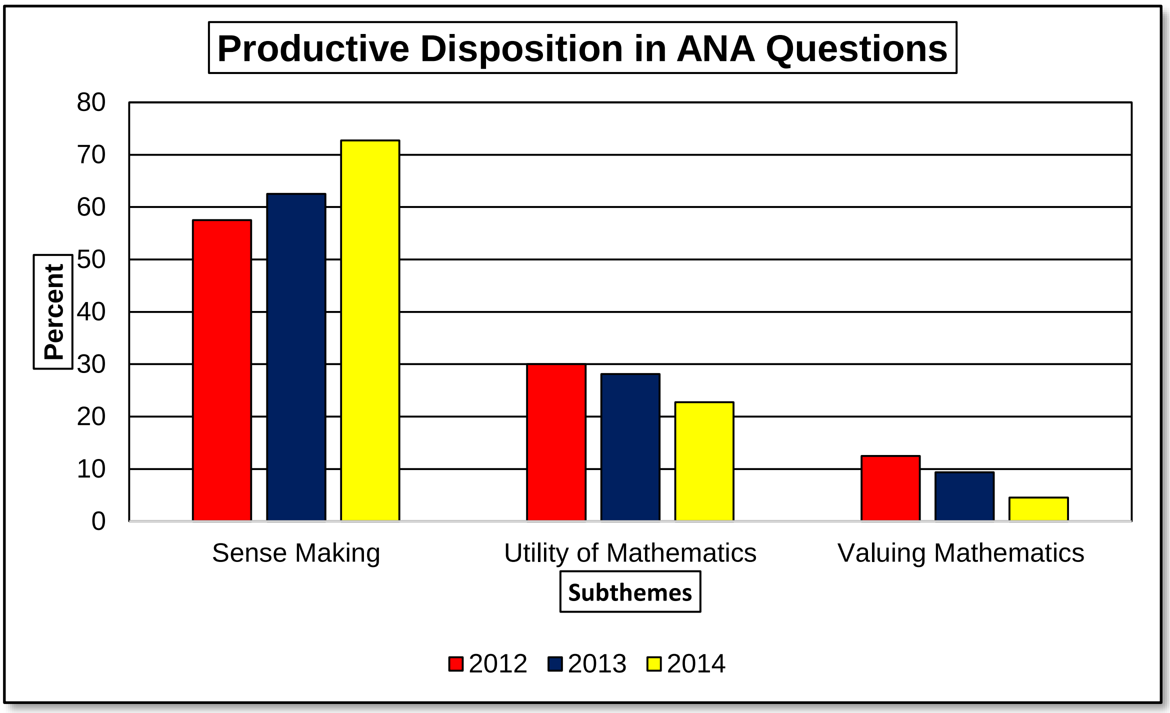


Figure 4.3: Productive Disposition in ANA Questions

Table 4.9 outlines the mean discrepancies for productive disposition by categories with. This data was generated from data in Table 4.8 which is on dispositions tested by the Grade 9 ANA mathematics tests. The direction indicates the strength of the disposition. Positive means strong dispositions while negative means weaker dispositions. Such dispositions are a result of the ANA testing of the first four SMP as tested in 2012, 2013 and 2014 (Kilpatrick et al., 2001). This is an indication that ANA tests were weak in the use of mathematics to solve real life experiences (Karakoc & Alacaci, 2015). There is no doubt that mathematics is regarded as an essential commodity to solve real life experiences (Mueller et al., 2011). This implies that the ANA deprived learners the opportunity to use mathematics to get answers of real life experiences (Lee & Chen, 2015). This is an indication that ANA in general did not support the usefulness of mathematics in solving complex problems (Mueller et al., 2011).

Table 4.9: Mean discrepancies with direction

ANA Examination	Sense Making Frequency (%)	Utility of Mathematics Frequency (%)	Valuing Mathematics Frequency (%)	Mean Frequency (%)
2012	9.66(24.16)	-1.33(-3.33)	-8.33(-20.83)	0(0)
2013	9.34(29.16)	-1.67(-5.17)	-7.67(-23.96)	0(0)
2014	8.66(39.39)	-2.33(-10.6)	-6.33(-28.79)	0(0)

The data in Figure 4.4, is a synopsis of mean deviation of productive disposition of in ANA testing. These results confirms those in Figure 4.3 which is an indication that ANA was not consistent in testing dispositions in the three consecutive years.

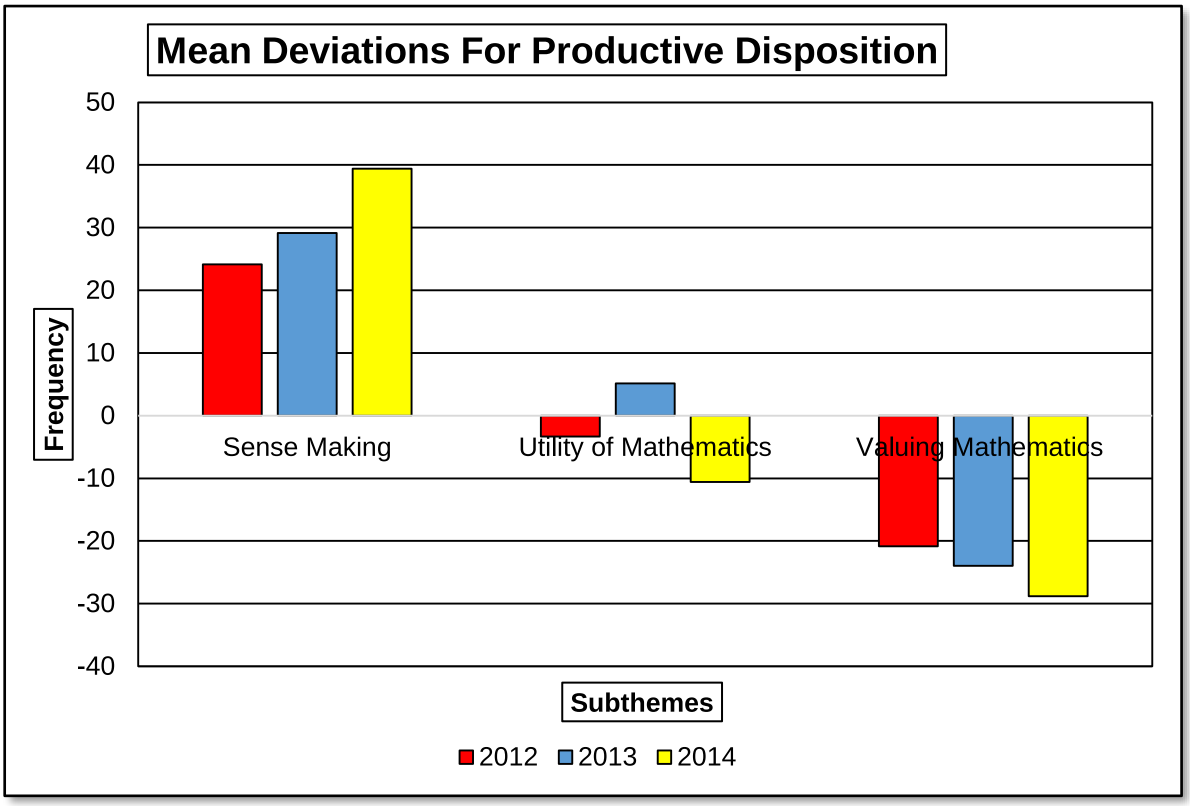


Figure 4.4: Mean Discrepancies for Productive Disposition

4.2.7 Implications of Strands Not Tested By ANA

Deductive reasoning refers to the ability to derive facts from logical inferences as well as processes involved in validating the argument and finally reaching a true conclusion (Lee 2016; Morris, 2002). The ANA does not provide these opportunities. Deductive reasoning is essential for advanced natural science courses such as advanced mathematics (Stalvey & Vidakovic, 2015; Sullivan et al., 2016) due to the high level of rigour. It implies that learners are most unlikely to succeed in advanced tertiary courses if they do not access logical reasoning in their early years of schooling.

According to Yopp (2010), inductive reasoning refers to learners' ability to reach conclusions from examples. Such reasoning enhances sense making and promotes the development of complex problem solving strategies (Koichu & Leron, 2015; Yopp, 2015). The implication in this instance is that, if systemic assessments such as ANA do not test inductive reasoning, then it justifies the suggestion made earlier that it is most likely that learners struggle in advanced courses in natural science as well as mathematics.

According to Alcock and Inglis (2008) and Cantlon (1998), conjecturing refers to learners' ability to write and conceive hypotheses that are yet to be conceived. Conjectures are powerful as they stimulate reasoning and proof (Aaron & Herbst, 2015). The results in the current study pose a serious challenge to the mathematics reasoning and they also confirm those made by Ally (2011) that there is an absence of reasoning in Grade 6 mathematics classrooms in South Africa. It implies that the ANA did not provide opportunities for learners to conjecture.

It seems that the absence of reasoning in earlier schooling years like Grade 9 and Grade 6 that was identified by Ally (2011) has a serious consequence for learners' success in advanced mathematics and natural sciences in tertiary. The need to introduce mathematical reasoning and proof in earlier grades of schooling is essential in mathematics education in South Africa. Some studies justify the importance of engaging students with mathematical reasoning at lower grades such as: Grade 5 learners using connected tasks (Richardson et al., 2010); study on inductive reasoning with primary school students (Tomic & Klauer, 1996); and a study with first graders on inductive reasoning (Kagan et al., 1966). Hence, it is safe to say that mathematical reasoning is essential in the Grade 9 mathematics ANA and its absence may pose serious challenges for learners in advanced mathematics.

Additionally, the absence of problem solving questions deprived learners of the opportunity to solve higher thinking such as productive thinking and metacognition (Roth et al., 2015). For productive thinking, questions require learners to use their knowledge with innovation and perseverance to reconstruct solution strategies (Guberman & Leikin, 2013) .As such, learners become metacognitive, which involves assessing and reconstructing their solution strategies (Veloo et al., 2015). Such questions did not emerge from the analysis of ANA using SMP. From the results it is also safe to say that the ANA 2012, 2013 and 2014 questions were low in demanding high cognitive demand which Sullivan, et al. (2016) regarded as key to advanced mathematics. It implies that the ANA questions did not support high cognitive thinking, and so these results confirm findings from studies like that of Luneta (2015b) who found that learners were operating at low cognitive thinking even in situations where the Grade 12 examinations were requiring them to operate at high cognitive thinking.

4.2.8 Synopsis: Mathematics Cognitive Levels in ANA Questions

The theoretical framework for this study, SMP identified five SMP and these were: procedural fluency, conceptual understanding, strategic competence, adaptive reasoning, and productive disposition. Subsequently, procedural fluency and conceptual understanding are forms of mathematical knowledge while strategic competence and adaptive reasoning are mathematical skills (Dhlamini & Luneta, 2016). While productive disposition cannot be classified as mathematical knowledge or skill, this strand represents dispositions that learners display as they become proficient in the first four strands (Kilpatrick et al., 2001). Identifying cognitive levels in the ANA that are reflective of SMP, I use the themes that emerge from the analysis of ANA using SMP, simple procedures, procedural fluency and conceptual understanding and adaptive reasoning and not productive disposition. Most specifically, I use the codes as their composition, especially, AR1-3, had an effect on the complexity of the ANA questions. The use of the NAEP Taxonomy has also been justified in this study, which implies that for mathematics cognitive levels in the ANA, I will focus on three aspects, low complexity, moderate complexity and high complexity (Berger et al., 2010). Table 4.10 outlines the summary of the NAEP Taxonomy which is used to classify codes in the three categories of complexity.

Table 4.10: NAEP Taxonomy (Berger et al., 2010: 40)

Levels of Complexity	NAEP TAXONOMY
Low Complexity	- Students to recall or recognise concepts or procedures. - Items typically specify what the student is to do, which is often to carry out some procedure that can be performed mechanically. - The student does not need to use an original method or to demonstrate a line of reasoning.
Moderate Complexity	-Items involve more flexibility of thinking and choice among alternatives than do those in the low complexity category. -Students are expected to decide what to do and how to do it, bringing together concepts and processes from various domains. For example, student may need to represent a situation in more than one way, to draw a geometric figure that satisfies multiple conditions, or to solve a problem involving multiple unspecified operations. -Students might be asked to show explain their work, but would not be expected to justify it mathematically.
High Complexity	- Students are expected to use reasoning, planning, analysis, judgment, and creative thought. - Students may be expected to justify mathematical statements or construct a mathematical argument. Items may require students to generalise from specific examples. - Items at this level take more time than those at other levels due to the demands of the task, not due to the number of parts or steps.

From the results from Table 4.1, 4.2, 4.3, 4.4 and 4.5, I generate Table 4.11 which outlines the levels of complexity in the 2012, 2013 and 2014 ANA tests. The questions from the sub-themes are classified as low complexity, moderate complexity and high complexity. This classification is informed by the NAEP Taxonomy and SMP. This gives an indication of the complexity of the Grade 9 mathematics ANA in 2012, 2013 and 2014 respectively. Subsequently, there exists some dichotomy in the classification. The two notions, NAEP Taxonomy and SMP were contradictory, for example, the SMP classified analogical proof as a weaker form of reasoning (Berger et al., 2010). In contrast, the NAEP Taxonomy pointed out that the use of proofs depicts high complexity. Hence this study integrated both conceptions and regarded the additional code 'AR3' as an extension of cognition and categorised all questions with this code as moderate complexity. Other subsequent codes were categorised in complexities as per the NAEP Taxonomy.

Table 4.11: Complexity of 2012, 2013 and 2014 ANA questions

Complexity of ANA questions			
ANA Examinations	***Low complexity code (Frequency)***	***Moderate complexity code (Frequency)***	***High complexity code (Frequency)***
2012 ANA	SP-SC1-SC3 (18) SP-SC1-SC2-SC3 (8)	SP-SC1-SC3-AR1 (1) SP-SC1-SC2-SC3-AR2 (1) SP-SC1-SC2-SC3-AR3 (6) PF1-SC1-SC3 (13) PF1-SC1-SC2-SC3 (11) PF1-SC1-SC2-SC3-AR1 (7)	CU1-SC1-SC3 (2) PF2-SC1-SC3 (1) PF2-SC1-SC2-SC3 (7)
	Total (26)	**Total (39)**	**Total (10)**
2013 ANA	SP-SC1-SC3 (8) SP-SC1-SC2-SC3 (17)	SP-SC1-SC3-AR1 (1) SP-SC1-SC2-SC3-AR1 (5) SP-SC1-SC3-AR2 (1) SP-SC1-SC2-SC3-AR3 (2) PF1-SC1-SC3 (11) PF1-SC1-SC2-SC3 (7) PF1-SC1-SC2-SC3-AR1 (3)	CU1-SC1-SC3 (3) PF2-SC1-SC3 (1) PF2-SC1-SC2-SC3 (4)
	Total (25)	**Total (30)**	**Total (8)**
2014 ANA	SP-SC1-SC3 (23) SP-SC1-SC2-SC3 (2)	SP-SC1-SC3-AR1 (2) SP-SC1-SC2-SC3-AR1 (5) SP-SC1-SC2-SC3-AR3 (4) PF1-SC1-SC3 (14) PF1-SC1-SC2-SC3 (13) PF1-SC1-SC2-SC3-AR1 (3)	CU1-SC1-SC3 (8) Cu2-SC1-SC3 (1) PF2-SC1-SC3 (1)
	Total (25)	**Total (41)**	**Total (10)**

Figure 4.5 is a synopsis of the complexity of the 2012, 2013 and 2014 Grade 9 mathematics ANA. The longer bars signify that in 2012, 2013 and 2014 most of the questions were of moderate complexity in the three consecutive years, lower in low complexity and low in high complexity. Mathematics assessment tasks does not have to be too weak or too strong (Hsu & Silver, 2014). In the case of the ANA tests in the three consecutive years, most of the test items are moderate and less in high complexity which is contrary to assessment protocols (Gysling, 2016). For example, the SAGM for mathematics in South Africa suggested that questions in mathematics assessments must be as follows: knowledge (25%), routine procedure (45%), complex procedures (20%) and problem solving (10%) (DBE, 2011). Surprisingly, ANA tests were not compliant with these demands, where there was absence of complex procedures and problem solving. Hence this condition of ANA pose validity issues as systemic assessment which are as follows; how can ANA test if the system is producing learners who think at a higher level, if it does not test at that level? The assumption I made earlier which is *"What You Test Is What You Get, (WYTIWUG)."* (Schoenfeld, 2007: 72) is true for ANA testing.

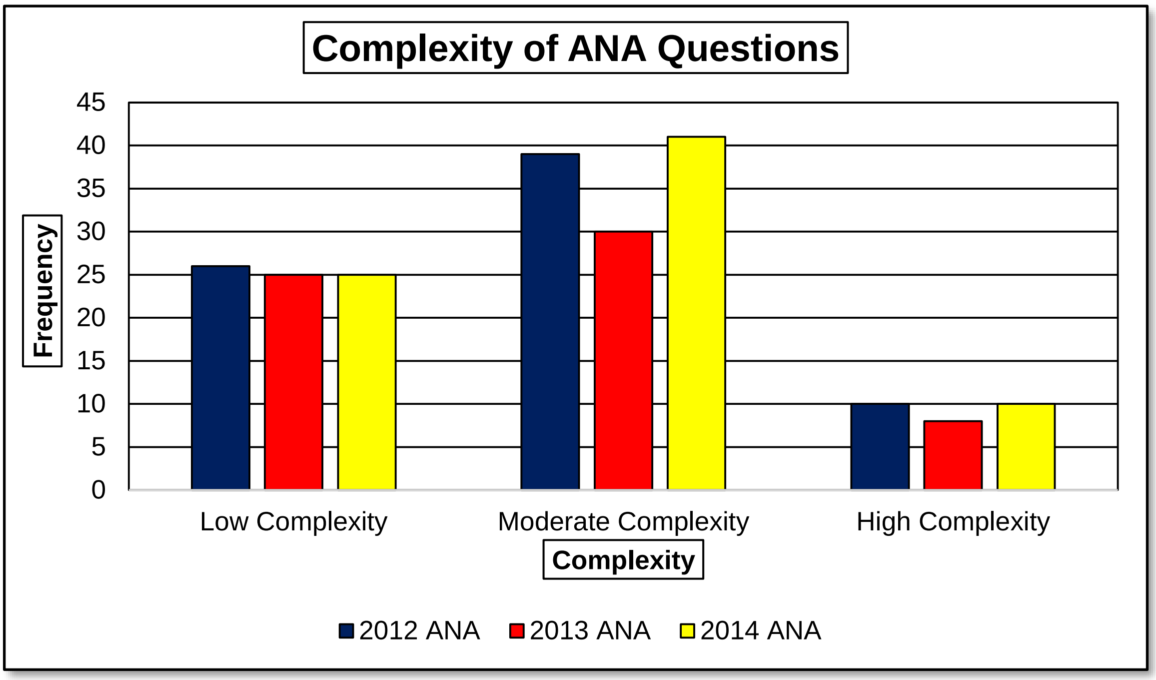


Figure 4.5: Complexity of ANA Questions

4.3 Results and Discussion for Learners' Responses to ANA Questions

In Chapter Three, I explained and justified the three questions in the learners' responses that were sampled and analysed from the 2014 Grade 9 mathematics ANA. Below, I explore the learners' responses to the sampled questions.

4.3.1 Learners' Responses to Question Three

In question three, the focus is on algebra and algebraic fractions. Learners struggle with algorithms of fractions due to the conceptual nature of fractions (Boyce & Norton, 2016). Learners in Grade 9 often confuse algorithms of fractions with those of natural numbers (Dhlamini & Kibirige, 2014). Hence, they perform a range of errors and misconceptions which may lead to difficulty in performing procedures and relations that involve all four basic operations on fractions (Ramful, 2014).

Table 4.12 is a synopsis of learners' responses to question 3 in seven schools. The analysis classified learner's responses into four categories: correctly answered, partially answered, incorrect response and no response. Detail on how these categories reflect to SMP that were exhibited by learners in response to the questions has been described in chapter three. These categories depicted the proficiency levels of learners for a variety of questions in various schools. Table 4.12 is a synopsis of how in seven schools' learners responded to question 3

The third column in Table 4.12 shows results for School A, with the highest number of correctly answered responses in all questions except question 3.3; which was 6.51% for question 3.1, followed by 26.05% for question 3.2, then 3.72% for question 3.3, further 9.3% for question 3.4 and lastly 11.63% for question 3.5. Slight increase in numbers of learners who were classified as partially answered in question 3.1, 3.2 and 3.4, with 31.16%, then 20.47% and 25.58% respectively with low numbers for question 3.3 and 3.5 which was 3.26% and 7.91%. A bulk of learners for school A were in the category incorrectly answered, 61.4 % for question 3.1, followed by 46.97% for question 3.2, then 87.44% for question 3.3, a further 55.35% for question 3.4 and lastly 69.76% for question 3.5. There were lower numbers in the category no response. A total of 0.93% for question 3.1, followed by 6.51% for question 3.2 then 5.58% for question 3.4, a further 9.77% for question 3.4 and last 10.7% for question 3.5.

Column four of Table 4.12 illustrates the results for School B. Significantly low number of learners were classified in the category correctly answered. No learners answered correctly for question 3.1, 3.3 and 3.4. Only 0.94% for question 3.2 and (0.5%) for question 3.5. A low number of learners partially answered the questions. There were 1.4% for question 3.1, followed by 0.47% for question 3.2 and 0.9% for question 3.5. No learners partially answered question 3.3.and 3.4. Most learners incorrectly answered all the questions. There were 93.9% for question 3.1, followed by 91.51% for question 3.2, a further 92% for question 3.3, then 87.7% for question 3.4 and last, 87.7% for question 3.5. There were lower numbers for the category no response. There were 4.7% for question 3.1, then 7.08% for question 3.2, followed by 8% for question 3.3, a further 12.3% for question 3.4 and last 10.9% for question 3.5.

In column five of Table 4.12 are results for School C for question three. For the category, correctly answered, there were no learners for question 3.1, 3.2, 3.3 and 3.4. Only two learners correctly answered question 3.5. For the category, partially answered, no learners partially answered question 3.3 and 3.5. A small number of learners, only 2% for question 3.1, with 10% for question 3.2, and 2% for question 3.4. A bulk of learners incorrectly answered all the questions for school C. There were 98% for question 3.1, followed by 90% for question 3.2, then 100% for question 3.3, a further 98% for question 3.4 and last, 96% for question 3.5. There were no learners that fell in the category no response for all the questions.

Column six of Table 4.12 summarises learners' results for School D for question three. In the category, correctly answered, there were no learners who correctly answered question 3.1, 3.3 and 3.5. For question 3.2 and 3.4 there were 0.53% and 0.5% learners who correctly answered respectively. For the category partially answered, no learners partially answered question 3.2 and 3.5. There were low numbers for the other questions, 0.5%, 1.05% and 2.1% for question 3.1, 3.3 and 3.4 respectively. Most learners incorrectly answered all the questions. There was 97.4% for question 3.1, then 94.8% for question 3.2, plus 96.3% for 3.3, further 95.3% for question 3.4 and last 93.1% for question 3.5. There were low numbers in the category no response. For question 3.1 there were no learners, question 3.2 had 0.53%, question 3.3 had 1.05%, then question 3.4 was 2.1% and last, question 3.5 with 1.1%.

Column seven in Table 4.12 outlines results for School E for question three. For the category correctly answered, no learners correctly answered question 3.1, 3.3, 3.4 and 3.5. Only 0.6% correctly answered question 3.2. In the category, partially answered, a few learners partially answered all questions. For question 3.1 there was 1.7%, question 3.2 had 2.3%, question 3.4 there was 0.57% and last question 3.5 there was 1.2%. No learners partially answered question 3.3. A bulk of learners incorrectly answered all the questions. There was 97.4% for question 3.1, followed by 94.8% for question 3.2, then 96.3% for question 3.3, further 94.27% question 3.4 and last 93.1% for question 3.5. There were low numbers for the category no response. Only 0.9% for question 3.1, then 2.3% for question 3.2, followed by 3.7% for question 3.3, a further 5.16% for question 3.4 and last, and 5.52% for question 3.5.

In column eight on Table 4.12 are results of School F for question three. In the category correctly answered, no learners correctly answered question 3.1, then 4.91% for question 3.2 and 3.3, further 3.68% for question 3.4, last, 5.52% for question 3.5. For the category partially answered, there was 1.84% for question 3.1 and 3.4, followed by 2.45% for question 3.2 and 3.3, last 5.52% for question 3.5. In the category incorrectly answered, the bulk of learners were in this category, 95.09% for question 3.1, and 87.12% for question 3.3 and 3.3, followed by 82.82% for question 3.4 and last 85.89% for question 3.5.

Column nine on Table 4.12, shows results for School G for question three. In the category correctly answered, there were 2.78% for question 3.1, followed by 5.55% for question 3.2, and no learners correctly answered question 3.3, then 4.17% for question 3.4 and last, 2.78% for question 3.5. For the category partially answered, 11.11% for question 3.1, followed by 15.28% for question 3.2, then 4.17% for question 3.3, further 12.5% for question 3.4 and last 6.94% for question 3.5. In the category incorrectly answered, 86.11% for question 3.1, followed by 79.17% for question 3.2, then 95.83% for question 3.3, then 83.33% for question 3.4 and last, 90.28% for question 3.5. For the category no response, there were no learners who had no responses.

In column ten, Table 4.12, shows means and standard deviations for learners' responses to the five levels of question three in four categories. The standard deviation is useful in instances where there is a need to measure variance of data from a comparable point (Lathrop, 1961). The range in standard deviation (Table 4.12) shows the degree of variance from the mean which may be positive or negative or zero (Saary, 2008). A zero range means an item is placed at the mean, followed by a negative range which depicts that an item is below the mean (Gorard, 2005). The last range, a positive range depicts that the item is above the mean (Lathrop, 1961). As such, for question 3.1, category 'correctly answered', the mean is 1.33 and the standard deviation is 2.51. Schools B, C, D, E and F are in the range,$-1\sigma < x < 0$, school A is in the range, $+2\sigma < x < +3\sigma$ and school G is in the range, $0 < x < +1\sigma$. In the category 'partially answered', the mean is 7.1 and the standard deviation is

11.21. Schools B, C, D, E and F are in the range, $-1\sigma < x < 0$, school A is in the range, $+2\sigma < x < +3\sigma$, and school G is in the range, $0 < x < +1\sigma$. School B, C, D, E and F are in the range, $0 < x < +1\sigma$, school A is in the range,$-3\sigma < x < -2\sigma$, and School G is in the range,$-1\sigma < x < 0$. In the category 'incorrectly answered', the mean is 90.2 and the standard deviation is 13.43. In the category 'no response', the mean is 1.37 and the standard deviation is 1.83. Schools B and F are in the range,$+1\sigma < x < +2\sigma$, and Schools A, C, D, E and G are in the range,$-1\sigma < x < 0$.

For question 3.2, in the category 'correctly answered', the mean is 6.65 and the standard deviation is 9.72. Schools B, C, D, E, F and G are in the range, $-1\sigma < x < 0$, and school SA is in the range, $+1\sigma < x < +2\sigma$. In the category 'partially answered', the mean is 4.99 and the standard deviation is 8.16. Schools B, D, E and F are in the range, $-1\sigma < x < 0$, school A is in the range, $+1\sigma < x < +2\sigma$, and School B and G are in the range,$0 < x < +1\sigma$. In the category 'incorrectly answered', the mean is 83.11 and the standard deviation is 13.43. Schools B, C, D, E and F are in the range, $0 < x < +1\sigma$, school A is in the range,$-3\sigma < x < -2\sigma$, and school G is in the range, $-1\sigma < x < 0$. In the category 'no response', the mean is 3.04 and the standard deviation is 3.09. School C, D, E and G are in the range, $-1\sigma < x < 0$, School A and B are in the range, $+1\sigma < x < +2\sigma$, and schools F is in the range, $0 < x < +1\sigma$.

For question 3.3, in the category 'correctly answered', the mean is 1.23 and the standard deviation is 2.13. Schools B, C, D, E and G are in the range,$-1\sigma < x < 0$, schools SA and SF are in the range,$+1\sigma < x < +2\sigma$. In the category 'partially answered', the mean is 1.56 and the standard deviation is 1.73. Schools B, C, D and E are in the range, $-1\sigma < x < 0$, schools SA, and SF are in the range,$0 < x < +1\sigma$, and school G is in the range, $+1\sigma < x < +2\sigma$. In the category 'incorrectly answered', the mean is 93.8 and the standard deviation is 5.07. School SA and SF are in the range, $-2\sigma < x < -1\sigma$, school B is in the range,$-1\sigma < x < 0$, school C is in the range,$+1\sigma < x < +2\sigma$, and Schools D, E and G are in the range, $0 < x < +1\sigma$. In the category 'no response', the mean is 3.41 and the standard deviation is 3.14. School A, E and F are in the range, $0 < x < +1\sigma$, school B is in the range, $+1\sigma < x < +2\sigma$, Schools C and G are in the range, $-2\sigma < x < -1\sigma$, and school D is in the range, $-1\sigma < x < 0$.

For question 3.4, in the category ‘correctly answered’, the mean is 2.52 and the standard deviation is 3.49. School A is in the range, $+1\sigma < x < +2\sigma$, schools B, C, D and E are in the range,$-1\sigma < x < 0$, and schools F and G are in the range,$0 < x < +1\sigma$. In the category ‘partially answered’, the mean is 6.37 and the standard deviation is 9.48. School A is in the range,$+2\sigma < x < +3\sigma$, schools A, B, D, E and F are in the range,$-1\sigma < x < 0$, and school G is in the range,$0 < x < +1\sigma$. In the category ‘incorrectly answered’, the mean is 85.25 and the standard deviation is 14.46. School A is in the range,$-3\sigma < x < -2\sigma$, and schools B, C, D and E are in the range,$0 < x < +1\sigma$, and schools, F and G are in the range,$-1\sigma < x < 0$. In the category ‘no response’, the mean is 5.86 and the standard deviation is 5.38. School A is in the range,$0 < x < +1\sigma$, schools B and F are in the range,$+1\sigma < x < +2\sigma$, schools C and G is in the range,$-2\sigma < x < -1\sigma$, D and E are in the range,$-1\sigma < x < 0$.

For question 3.5, in the category ‘correctly answered’, the mean is 3.49 and the standard deviation is 4.17. School A is in the range,$+1\sigma < x < +2\sigma$, schools B, D, SE and G are in the range,$-1\sigma < x < 0$, and schools C and F are in the range,$0 < x < +1\sigma$. In the category ‘partially answered’, the mean is 2.86 and the standard deviation is 3.3. Schools A and G are in the range,$+1\sigma < x < +2\sigma$, schools B, C, D and E are in the range,$-1\sigma < x < 0$, and school F is in the range,$0 < x < +1\sigma$. In the category ‘incorrectly answered’, the mean is 88.8 and the standard deviation is 9.55. School A is in the range,$-2\sigma < x < -1\sigma$, schools B and F are in the range,$-1\sigma < x < 0$, schools C, E and G are in the range,$0 < x < +1\sigma$, and school D is in the range,$+1\sigma < x < +2\sigma$. In the category ‘no response’, the mean is 4.85 and the standard deviation is 4.71. Schools A and B are in the range,$+1\sigma < x < +2\sigma$, schools C and G are in the range,$-2\sigma < x < -1\sigma$, school D is in the range,$-1\sigma < x < 0$, and schools E and F are in the range,$0 < x < +1\sigma$.

Table 4.12: Learners' responses to question 3

ANA 2014 GRADE 9 MATHEMATICS LEARNERS' RESPONSES IN VARIOUS SCHOOLS									
Question 3 Algebra and algebraic fractions	*Learners' responses*	*SA* *(F) %*	*SB* *(F) %*	*SC* *(F) %*	*SD* *(F) %*	*SE* *(F) %*	*SF* *(F) %*	*SG* *(F) %*	*Standard Deviations & Means* $(\bar{x})\sigma$
3.1 Simplify	Correctly answered	(14)6.51	(0)0	(0)0	(0)0	(0)0	(0)0	(2)2.78	(1.33)2.51
$2(x+2)^2-(2x-2)(x+2)$	Partially answered	(67)31.16	(3)1.4	(1)2	(1)0.5	(6)1.7	(3)1.84	(8)11.11	(7.10)11.21
(n=1250)	Incorrectly answered	(132)61.4	(199)93.9	(49)98	(188)99.5	(340)97.4	(155)95.09	(62)86.11	(90.2)13.43
	No response	(2)0.93	(10)4.7	(0)0	(0)0	(3)0.9	(5)3.07	(0)0	(1.37)1.83
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
3.2 Simplify	Correctly answered	(56)26.05	(2)0.94	(0)0	(1)0.53	(2)0.6	(8)4.91	(4)5.55	(6.65)9.72
$\frac{15x^2y^3-9x^2y^3}{8x^2y^3}$	Partially answered	(44)20.47	(1)0.47	(5)10	(0)0	(8)2.3	(4)2.45	(11)15.28	(4.99)8.16
	Incorrectly answered	(101)46.97	(194)91.51	(45)90	(187)98.94	(331)94.8	(142)87.12	(57)79.17	(83.11)17.49
(n=1250)	No response	(14)6.51	(15)7.08	(0)0	(1)0.53	(8)2.3	(9)5.52	(0)0	(3.04)3.09
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
3.3 Simplify	Correctly answered	(8)3.72	(0)0	(0)0	(0)0	(0)0	(8)4.91	(0)0	(1.23)2.13
$\frac{x^2-4x}{x^2-2x-8}$	Partially answered	(7)3.26	(0)0	(0)0	(2)1.05	(0)0	(4)2.45	(3)4.17	(1.56)1.73
	Incorrectly answered	(188)87.44	(195)92	(50)100	(185)97.9	(336)96.3	(142)87.12	(69)95.83	(93.80)5.07
(n=1250)	No response	(12)5.58	(17)8	(0)0	(2)1.05	(13)3.7	(9)5.52	(0)0	(3.41)3.14
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
3.4 Simplify	Correctly answered	(20)9.3	(0)0	(0)0	(1)0.5	(0)0	(6)3.68	(3)4.17	(2.52)3.49
$\frac{x^2}{2}+\frac{2x^2}{3}-\frac{7x^2}{6}$	Partially answered	(55)25.58	(0)0	(1)2	(4)2.1	(2)0.57	(3)1.84	(9)12.5	(6.37)9.48
	Incorrectly answered	(119)55.35	(186)87.7	(49)98	(180)95.3	(329)94.2	(135)82.82	(60)83.33	(85.25)14.46
(n=1250)	No response	(21)9.77	(26)12.3	(0)0	(4)2.1	7	(19)11.66	(0)0	(5.86)5.38
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	(18)5.16	**(163)**	**72**	(178.57)99.77
						(349)			
3.5 Simplify	Correctly answered	(25)11.63	(1)0.5	(2)4	(0)0	(0)0	(9)5.52	(2)2.78	(3.49)4.17
$\frac{6x^2}{7xy}\times\frac{3y^3}{2x}$	Partially answered	(17)7.91	(2)0.9	(0)0	(0)0	(4)1.2	(5)3.07	(5)6.94	(2.86)3.30
	Incorrectly answered	(150)69.76	(186)87.7	(48)96	(187)98.9	(325)93.1	(140)85.89	(65)90.28	(88.80)9.55
(n=1250)	No response	(23)10.7	(23)10.9	(0)	(2)1.1	(20)5.7	(9)5.52	(0)0	(4.85)4.71
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77

Key: SA, school A; SB, school B; SC, school C; SD, school D; SE, school E; SF, school F; SG, school G; %, percent; F, frequency.

Table 4.13 is an outline of mathematical activities that were examined by question 3. There are also codes of SMP that these questions demanded. For more detail on the coding of learners' SMP see Table 3.3. Both Table 4.13 and 4.14 show codes of SMP that were exhibited by learners in response to question 3.

Table 4.13: SMP required by question 3

SMP in question 3		
Question	*Mathematical activity*	*Codes*
3.1 Simplify $2(x+2)^2-(2x-2)(x+2)$	Apply BODMAS rule, evaluate a square of a binomial, multiply binomial, add and subtract like terms.	PF1-SC1-SC3
3.2 Simplify $\frac{15x^2y^3-9x^2y^3}{8x^2y^3}$	Adding like terms and dividing through.	CU1-SC1-SC3
3.3 Simplify $\frac{x^2-4x}{x^2-2x-8}$	Factorisation of numerator and denominator, dividing through.	CU1-SC1-SC3
3.4 Simplify $\frac{x^2}{2}+\frac{2x^2}{3}-\frac{7x^2}{6}$	Multiply each term by $\frac{6}{6}$, to make every denominator 6 and add numerators.	CU1-SC1SC3
3.5 Simplify $\frac{\boldsymbol{6x^2}}{\boldsymbol{7xy}}\times\frac{\boldsymbol{3y^3}}{\boldsymbol{2x}}$	Divide through and multiply like terms.	CU1-SC1-SC3

Table 4.14 explains with codes of SMP that are likely to be exhibited by learners as they respond to the five parts of question three. The codes are divided into four categories for each part of the question. These codes are key to the analysis and categorising of learners' responses to question three. The codes were derived from the general codes of SMP shown in Table 3.3 and the suggested answers for the ANA test.

Table 4.14: Explanation of learners' SMP to the five parts of question 3

SMP analysis key on Question 3			
SMP3.1A No response	SMP3.1B Incorrect procedures for squaring a binomial, multiplication of brackets and simplification	SMP3.1C One or two of these were incorrect, squaring a binomial or multiplication of brackets or simplification	SMP3.1D Correct squaring of binomial, multiplication of brackets and simplification
SMP3.2A No response	SMP3.2B Incorrect procedures for adding like terms and dividing like terms	SMP3.2C Either incorrect addition of like terms or division of like terms	SMP3.2D Correct addition of like terms and division of like terms
SMP3.3A No response	SMP3.3B Incorrect procedures for factoring the numerator, denominator and division by $'x-4'$	SMP3.3C Either incorrect factoring numerator or factoring trinomial or division by $'x-4'$	SMP3.3D Correct factoring of common factor in numerator, factorisation of a trinomial and dividing by $'x-4'$
SMP3.4A No response	SMP3.4B Applying incorrect procedure e.g. that of equation	SMP3.4C Either incorrect use of LCD or subtraction or division	SMP3.4D Correct use of LCD, addition and subtraction
SMP3.5A No response	SMP3.5B Incorrect procedures of multiplication and division of like terms	SMP3.5C Either, incorrect multiplication or division of like terms	SMP3.5D Correct multiplying and division of like terms

❖ **LEARNERS' RESPONSES TO QUESTION 3.1**

In this question, learners were required to simplify an algebraic expression. The simplification required learners to apply the BODMAS rule, multiply binomials, and finally add and subtract like terms, (Grade 8 work) or factorise common factors (Grade 9 work) as shown in the curriculum (DBE, 2011). Question 3.1 examined SMP according to the codes that emerged as shown in Table 3.3, and the analysis, as shown in Table 4.13 summarises the proficiency levels in the question.

Figure 4.6 summarises the trend of learners' responses to question 3.1 using the radar charts. The use of radar charts has been prevalent in presenting observational data (Saary, 2008). Subsequently, the radar charts measure the area of a polygon in multivariate data (Feldman, 2013). A regular polygon depicts maximum frequency in all variables (Nurse et al., [Sa]). In contrast, when a vertex of the polygon is centred towards zero, shows that the variable is minute (Saary, 2008). In the current study, the use of seven schools make the data multivariate. Subsequently, the use of four levels of proficiency (incorrectly answered, no response, moderate proficient and partially answered) confirms the multivariate nature of the data. For the category, correctly answered, the distribution is almost centred for all schools. As such it shows that a small number of learners answered question 3.1 correctly. For the category incorrectly answered, the distribution is almost a regular heptagon showing that most learners incorrectly answered question 3.1. The obvious irregular in the polygon is in School A. This shows the 61.4% of learners who incorrectly answered question 3.1. For the category partially answered, there is an irregular heptagon with only school A (31.16%) and School G (11.11%). The remainder of the schools were almost centred in the distribution showing that the numbers were either too low or zero. The category no response, is almost centred, showing that the values are too small or zero.

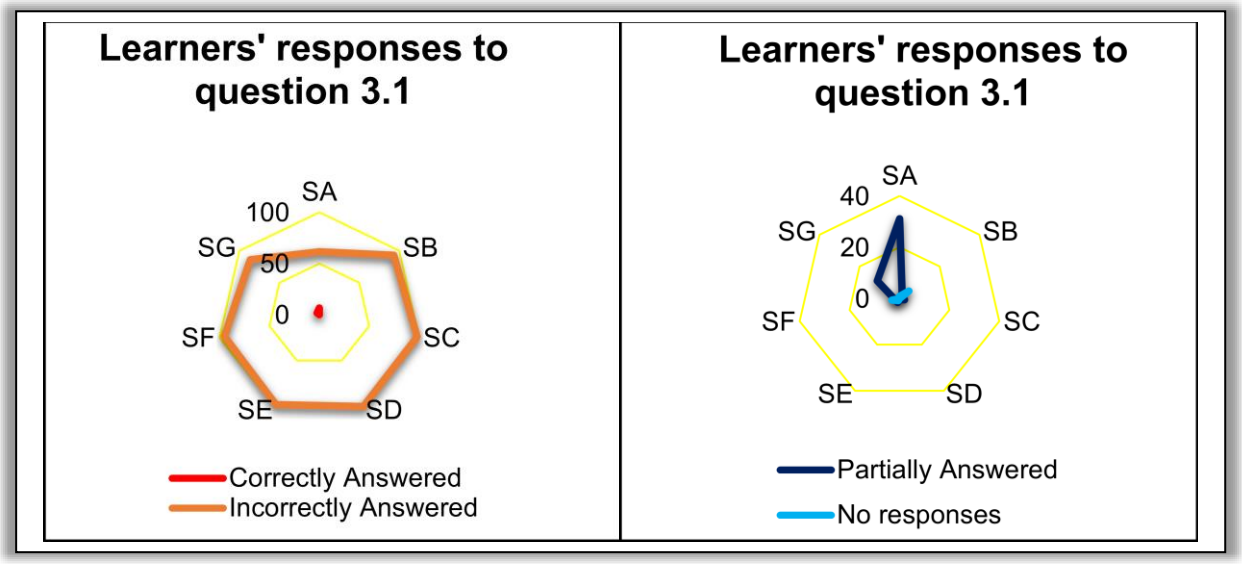
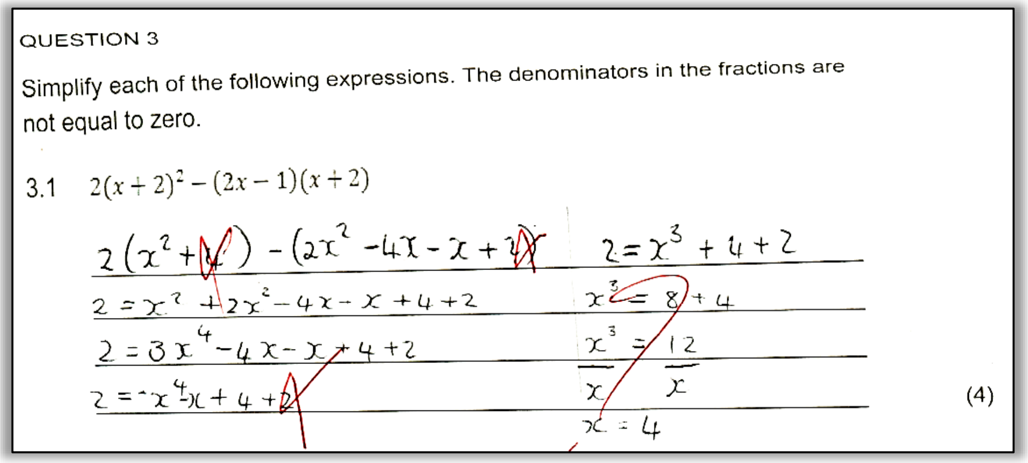
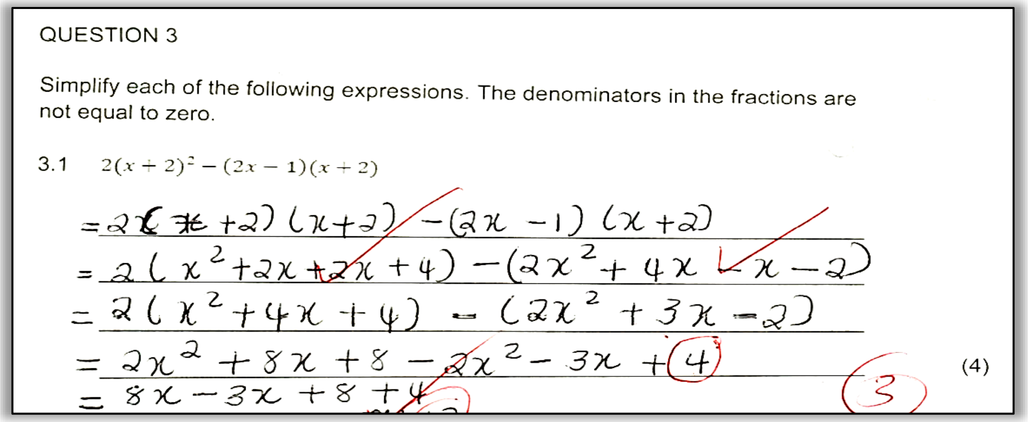


Figure 4.6: Trend in learners' responses to question 3.1

Figure 4.7 are vignettes of two learners' responses to question 3.1. Earlier, in the analysis of ANA questions, question 3.1 was coded PF1-SC1-SC3 which indicated that this was a routine procedure for Grade 9, for simplification of an algebraic expression. This question was a computation that requires procedural fluency, described by Kilpatrick et al. (2001) as the skill of performing procedures f*lexibly*, *accurately*, *efficiently* and *appropriately*.



Learner A



Learner B

Figure 4.7: Learners' responses to question 3.1

Part one of the procedure, *flexibility* refers to knowing various procedures and relative efficiencies of the procedures (Star, 2004). Here, Learner A and Learner B were limited to one procedure, which is to simplify by expanding brackets, then group like terms and finally make the solution simple. According to DBE (2011) this is a Grade 8 procedure which test the use of BODMAS to simplify algebraic expressions and in Grade 9 learners must factorise to simplify algebraic expressions. Ironically the two learners were not wrong to apply the Grade 8 procedure as the question did not specify a procedure and the question that provided flexibility. It could be said that these learners were operating procedurally at Grade 8 level.

Part two, *accuracy* often refers to the correct use of signs to carry out basic operations that involve addition, subtraction, multiplication and division to reach consistently the required answer (Bass, 2003). Learner A computes $(x + 2)^2$ into be $x^2 + 4$ which shows incorrect use of an algorithm of exponents, which applies to monomial terms and not binomials. The second expansion, $(2x - 1)(x + 2)$ results to $2x^2 - 4x - x + 2$ which has wrong signs for '-4' and '$+2$' respectively. This justifies that the learner lacked accuracy. In contrast, Learner B, expanded and grouped like terms correctly; however there was a mistake in writing '$(+4)$'. This mistake could have been caused by lack of flexibility. The use of another procedure such as factorisation could have avoided too many signs which make learners less error prone.

Part three, *efficiency* is often visible in learners who conduct procedures certainly, and use intermediate outcomes to execute the problem (NCTM, 2000). Since the problem allowed flexibility, and the learners opted for BODMAS, a grade eight outcome, Learner B showed efficiency in this procedure and it is worrisome that the marker penalised for a slip when procedurally the learner seemed fluent. This is an indication that the marking focused on the product at the expense of the procedure. Learner A incorrectly introduced the equal sign, an algorithm for solving equations not simplification. Hence the marking correctly penalised the learner.

Part four, *appropriation* refers to learners who are conscious of the right time of applying a procedure (Schoenfeld, 1985). Learner A seemed to lack appropriation by using two irrelevant algorithms in this procedure. In Contrast, Learner B correctly applied the procedure in all steps only to be let down by a slip. In both cases the cause was lack of flexibility, which was, choice of procedure that was error prone. The response by Learner A was coded (SMP3.1B, see Table 4.14) which was common, learners who were *not proficient* amongst the seven schools (90.2%). Subsequently, the response by Learner B was coded SMP3.1C, one of the 7.7% learners who were *moderately proficient* to question 3.1.

❖ LEARNERS' RESPONSES TO QUESTION 3.2

In Figure 4.8 shows the distribution of learners' responses to question 3.2. The category incorrect responses is almost a regular heptagon, except in school A. This irregular distribution in School A shows 46.97% learners who incorrectly answered question 3.1 in school A. Other schools that distort the regular heptagon are School G (79.17% incorrectly answered), School F (95.09% incorrectly answered), School C (98% incorrectly answered) and School B (91.51% incorrectly answered). For the category partially answered is almost centred for most of the schools except for School SA (20.47% partially answered). School SC (2% partially answered) and School SG (15.28% partially answered). This shows schools with low numbers or zero in this category. In the category correctly answered the distribution shows an irregular shape that is almost centred in some of the schools. The distribution shows that shows that for correctly answered, there were very low number of learners and in some schools it was zero. The exception was in School A, (26.05% correctly answered), School F (4.95% correctly answered) and School G (5.55% correctly answered). The category no response is an irregular distribution with some school centred. The exception was in three School A (6.51% no response), School B (7.08% no response) and School F (5.52% no response).

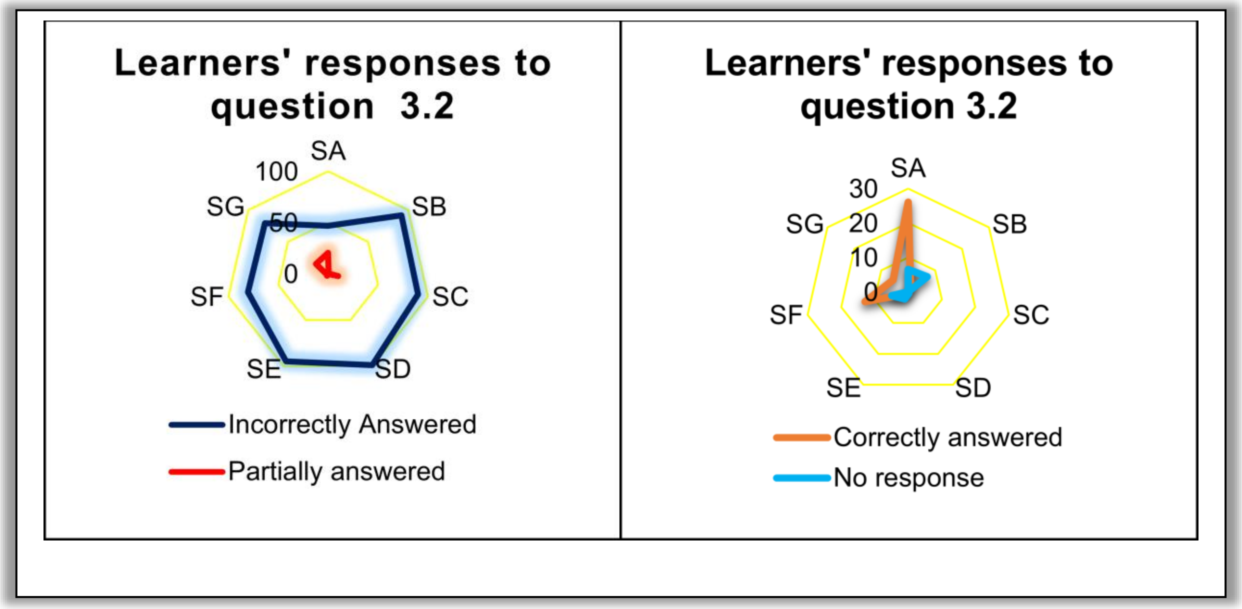
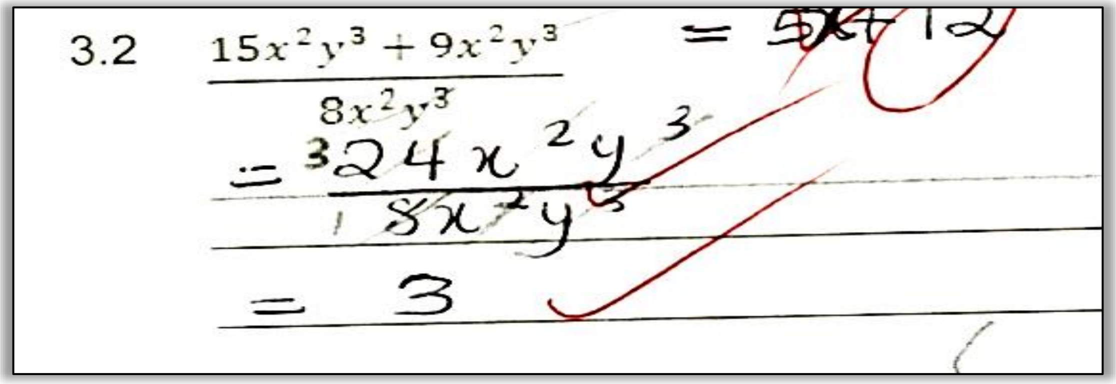


Figure 4.8: Trend in learners' responses to question 3.2

Figure 4.9 sets out learners’ responses to question 3.2. This question was earlier coded CU1-SC1-SC3 (see Table 4.13) which shows that this question was routine conceptual connections (see Table 3.3 for explanation of coding). Kilpatrick et al. (2001) pointed out that conceptual understanding is visible when learners comprehend mathematical concepts, operations and relations. Similarly, conceptual understanding is knowing how and why (McCormick, 1997). Ironically, the use of the word ‘comprehend’ by Kilpatrick et al. (2001) allows categorising questions that connects mathematical ideas, concepts and relations to be conceptual. Additionally, another assertion made is that conceptual understanding and procedural fluency are inseparable (Schneider et al., 2011). These inferences highlights that the *‘how’* part is procedural knowledge and the extension to *‘how and why’* makes a question examines conceptual understanding (Star, 2004). Additionally, another extension of the procedure comprehending a variety of concepts in one question. In question 3.2, concepts that were connected were addition of algebraic and division of algebraic expressions.

Learner A



Learner B

Figure 4.9: Learners’ responses to question 3.2

I view the learners' responses from the way the question was coded, conceptual understanding, which comprehends two concepts addition and division, and factorisation and division. First, *flexibility*: Learner A and Learner B chose addition in the numerator, a procedure for Grade 8. The question did not state the procedure to follow, meaning it was flexible. However, simplifying through factorisation is a Grade 9 procedure (DBE, 2011). Second, *accuracy*: adding the numerators both learners followed the procedure correctly to get $24x^2y^3$. In the other procedure, division, Learner A, incorrectly divided both the numerals, by writing the inverse and the variables by applying an algorithm of addition and subtraction, which says keep the variables unchanged. Mistakenly, Learner A missed x^2 which confirms that this learner lacked accuracy. The marking focused on the answer and disregarded the procedure by not awarding any mark to Learner A. On the other hand, Learner B accurately performed division to confirm that the learner exhibited accuracy. Third, *efficiency*: Learner A was not efficient due to the two observed mistakes and yet, Learner B was efficient by the consistency in both concepts, addition and division. As such Learner A was coded SMP3.2B for being *not proficient* which was common in 90.2% on learners' responses, amongst the seven sampled schools. Subsequently, Learner B was coded SMP 3.2D, 5.51% for learners' who were *proficient* in question 3.2 in the sampled schools. Fourth: for *appropriation*, although the question was flexible, both learners were operating at Grade 8 level (DBE, 2011) and could not use factorisation an outcome for Grade 9 (DBE, 2011). It is safe to infer that Learner A did not show appropriation to pre-knowledge and Learner B only showed appropriation in pre-knowledge, Grade 8.

❖ LEARNERS' RESPONSES TO QUESTION 3.3

In Figure 4.10 shows the trend of learners' responses to question 3.3. The category no response is centred which shows that in most of the schools there were low or zero learners that did not respond to question 3.3. For the category incorrectly answered, the distribution is almost a regular heptagon. This shows that most of the learners incorrectly answered question 3.3 in almost all the schools. In the category correctly answered is a distribution that is irregular, School A (3.72% correctly answered), School F (4.91% correctly answered) and the other schools have no learners who correctly answered question 3.3. The category partially answered has an irregular shape which is centred in three schools (zero partially achieved) and a few learners for School A (3.26% partially achieved), School D (1.05% partially achieved) School F (2.45% partially achieved) and School G (4.17% partially achieved).

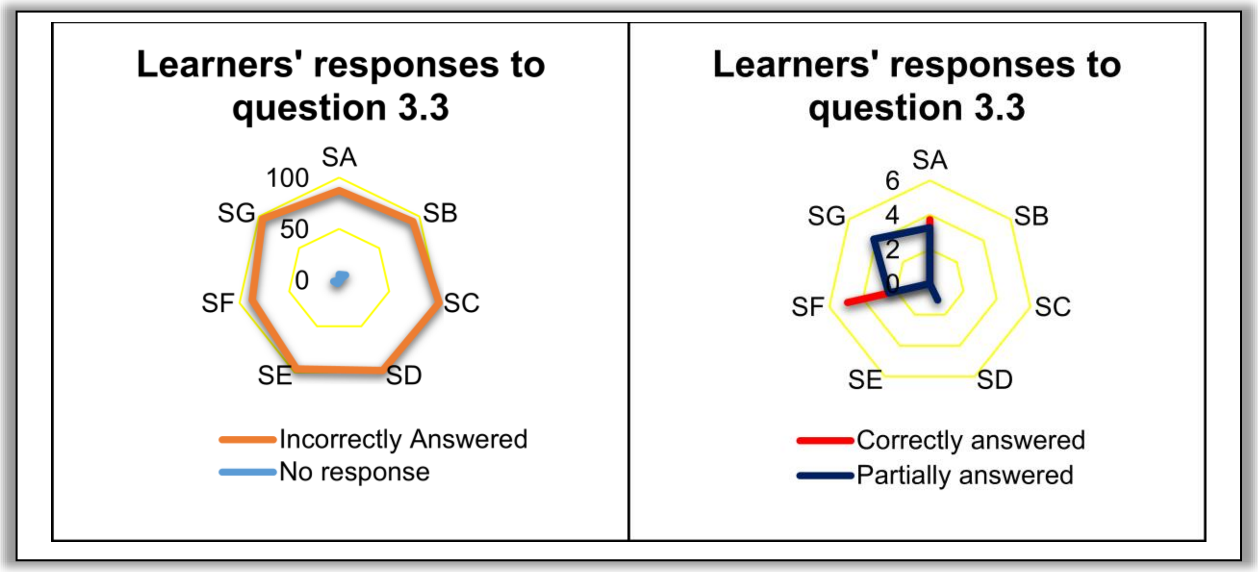
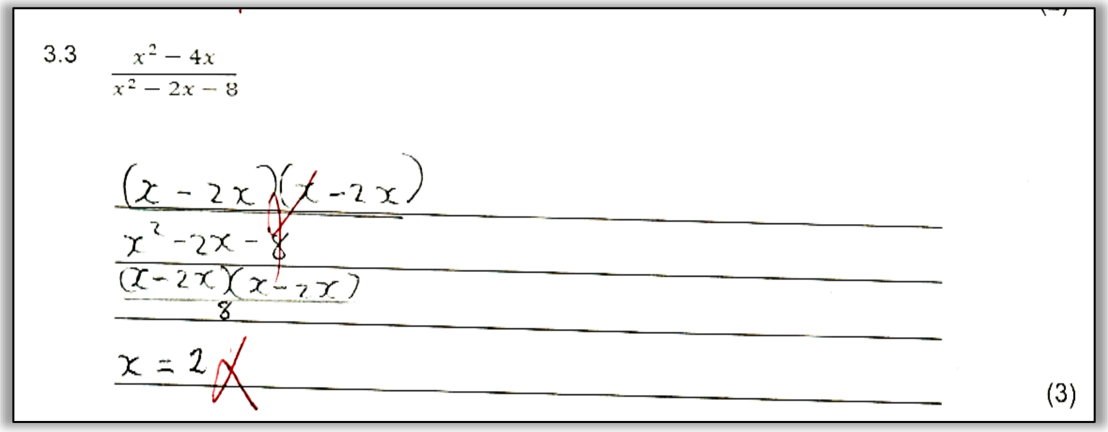
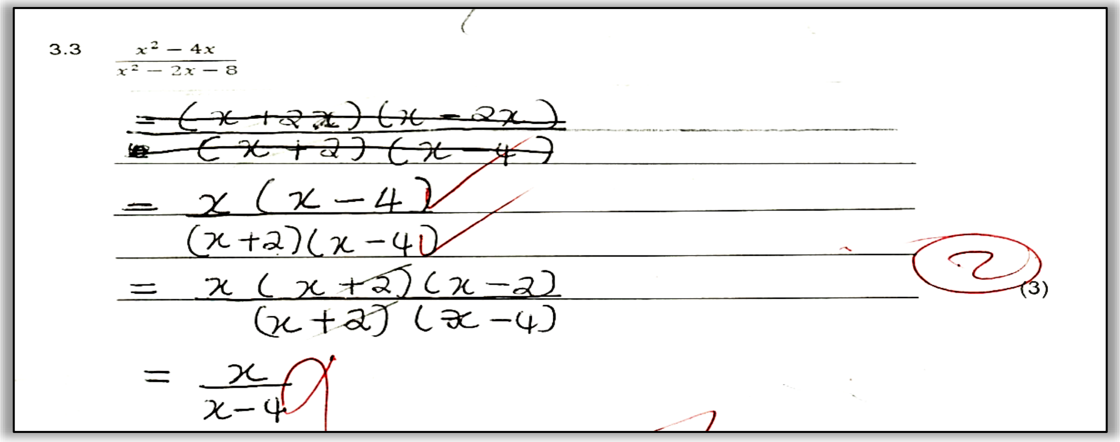


Figure 4.10: Trend in learners' responses to question 3.3

The vignettes in Figure 4.11 are learners' responses to question 3.3. This question was earlier coded CU1-SC1-SC3 (see Table 4.13) which shows that this question required routine conceptual connections. The concepts that were comprehended in this question were: (1) factorisation of a binomial (the numerator), (2) factorisation of a trinomial (the denominator) and division to make the expression simple. Phase one, *flexibility*, Learner A and Learner B chose to factorise the numerator which was the only procedure meaning that the question did not posit transparency. In Phase two, *accuracy,* in the numerator, Learner A used the wrong algorithm, which is difference of two squares instead of factoring a common factor. By contrast Learner B correctly factored the common factor. In the denominator, Learner A reduced the trinomial to 8 and it is not clear how that was computed to show lack of accuracy. By contrary, Learner B correctly factorised the denominator to confirm accuracy.



Learner A



Learner B

Figure 4.11: Learners' responses to question 3.3

In Phase three, *efficiency*, in Learner A, the factorisation is wrong which shows the learner has not yet achieved the learning outcome for Grade 9. Similarly, Learner B, incorrectly factored $x - 4$ to $(x + 2)(x - 2)$. Here the learner applied the algorithm of the difference of two squares. Ironically the factors are; $x - 4 = (x^{\frac{1}{2}} + 2)(x^{\frac{1}{2}} - 2)$. If the learner could have reached this step, the question could have not factorised unless the same was done to the denominator. Here the learner was not efficient by the failure to divide $x - 4$. In Phase four, *appropriation* the response by Learner A posits some deficiencies in the procedures, and the conclusion '$x = 2$, is for equations not algebraic expressions, and justifies lack of appropriation. By contrast, learner B failed to realise common factors that divide, and this posits lack of appropriation. As such, Learner A was one of 93.8% for learners who were *not proficient* in question

3.3 and coded SMP3.3B. Subsequently, Learner B was coded SMP3.1C, partially answered category which had 7.7% of learners' responses amongst the sampled seven schools.

❖ LEARNERS' RESPONSES TO QUESTION 3.4

In Figure 4.12 is a distribution for learners' responses to question 3.4. In the category correctly answered, the distribution is centred which indicate that there were either low numbers or no learners who correctly answered question 3.4. For school A, there were 9.3%, school F had 3.68%, school D had 0.5% and school G had 4.17%. The other schools, which are school B, school C and school E had no learners who correctly answered question 3.4. In the category incorrectly answered, the distribution is almost a regular heptagon, showing that most learners incorrectly answered question 3.4, with an exception of School A, (55.35% incorrectly answered). For the category partially answered, the distribution shows an irregular shape centred for most of the schools, showing low numbers for this category in most schools except for School A (25.58% partially answered) and School G (6.94% partially answered). The last category, no response, shows a distribution that is irregular with some schools centred to show that they had low numbers of learners who had no responses to question 3.4. There is an exception in four schools, which are School A (9.77% no response), School B (12.3% no response), School E (5.16% no response) and School F (11.66% no response).

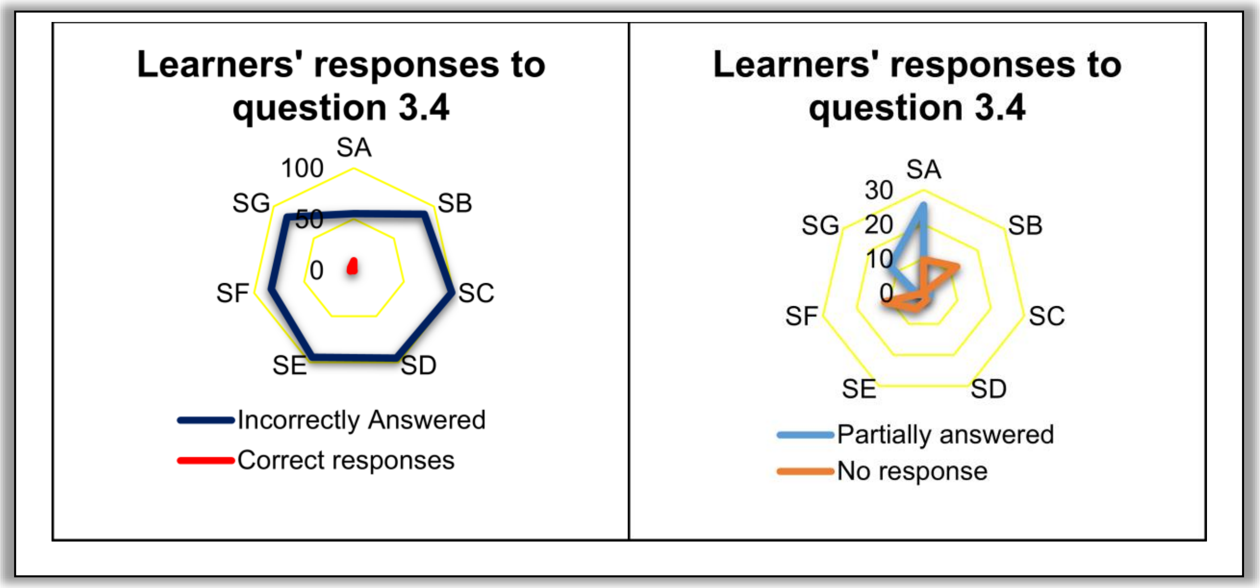
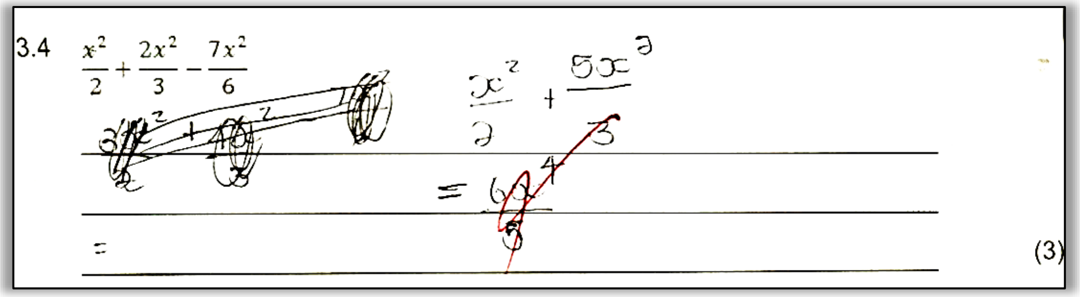


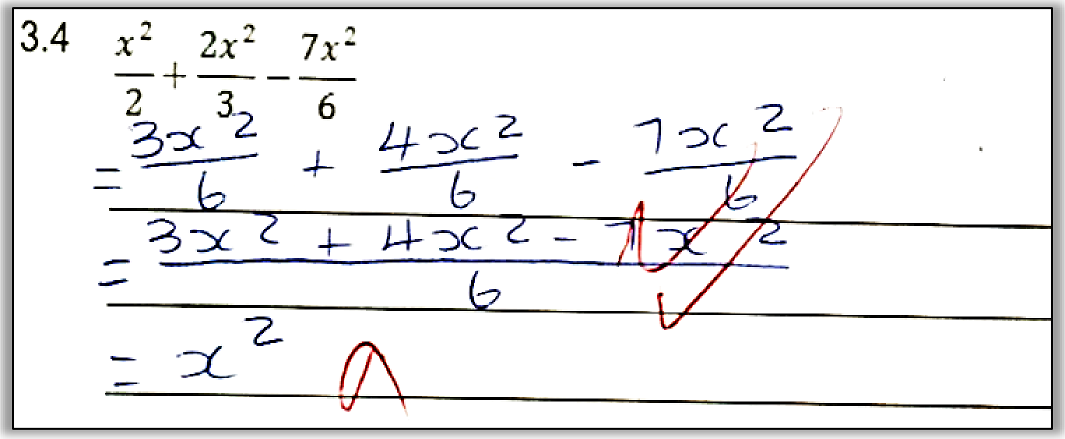
Figure 4.12: Trend in learners' responses to question 3.4

The vignettes Figure 4.13 are two samples of learners' responses to question 3.4.. This question was earlier coded CU1-SC1-SC3 (see Table 4.13) which shows that this question required routine conceptual connections. The concepts that were comprehended in this question were: 1) addition, 2) subtraction, 3) BODMAS, 4) division and 5) Equivalence.

Phase one, *flexibility*: Learner A first subtracted the second and third term then added the terms. By contrast, Learner B first expressed the terms as equivalent fractions then simplified. The question provided flexibility by not stating the procedure and allowing learners to use the following: 1) make terms equivalent then compute (BOBMAS); and 2) express terms in a common denominator, then compute to simplify. Learner A did not use any of these procedures, hence procedurally the learner was wrong. Learner B used the first procedure. Phase two, *accuracy*, Learner A could not observe BODMAS by subtracting first and the resulting sign was wrong $(+5x^2)$. Subsequently the learner added the numerators and the denominators, first correct, '$1+5=6$' and '$2+3=5$', but in the context of fractions this was wrong. The learner applied an algorithm for integers in fractions and incorrectly regarded the numerators as separate numbers. Learner B computed incorrectly in the simplification.



Learner A



Learner B

Figure 4.13: Learners' responses to question 3.4

Phase three, *efficiency*: for Learner A there was no certainty in conducting the procedure and this could have been caused by the learner being procedurally wrong. By contrast, Learner B was procedurally correct however, the learner failed to conduct the procedure certainly because of lacking accuracy in computations when simplifying. Phase four, *appropriation*: Learner A was not conscious of BODMAS and the correct algorithms to apply. Learner B was conscious of the procedure to use but was not certain in conducting the procedure during the simplification. As such, Learner A was coded SMP3.4B, one of 85.25%, for learners' not proficient in question 3.4 in the seven sampled schools. Learner B was one of 6.37% learners' responses coded SMP3.1C, who partially answered question 3.4.

❖ LEARNERS’ RESPONSES TO QUESTION 3.5

Figure 4.14 illustrates learners’ responses to question 3.5. The category no response, is a distribution that is almost centred. This shows that there were few or no learners who were in this category in the schools. In the category, incorrectly answered, the distribution is almost a regular heptagon which means that most learners incorrectly answered question 3.5 with the exception of School A (69.76% incorrectly answered). For the category partially answered, the distribution is irregular, showing a small number of learners and zero in two schools that partially answered question 3.5. Last, the category correctly answered is also irregular with a low number of learners in the majority of the schools except for School A (11.63% correctly answered), School C (4% correctly answered) and School F (5.52% correctly answered).

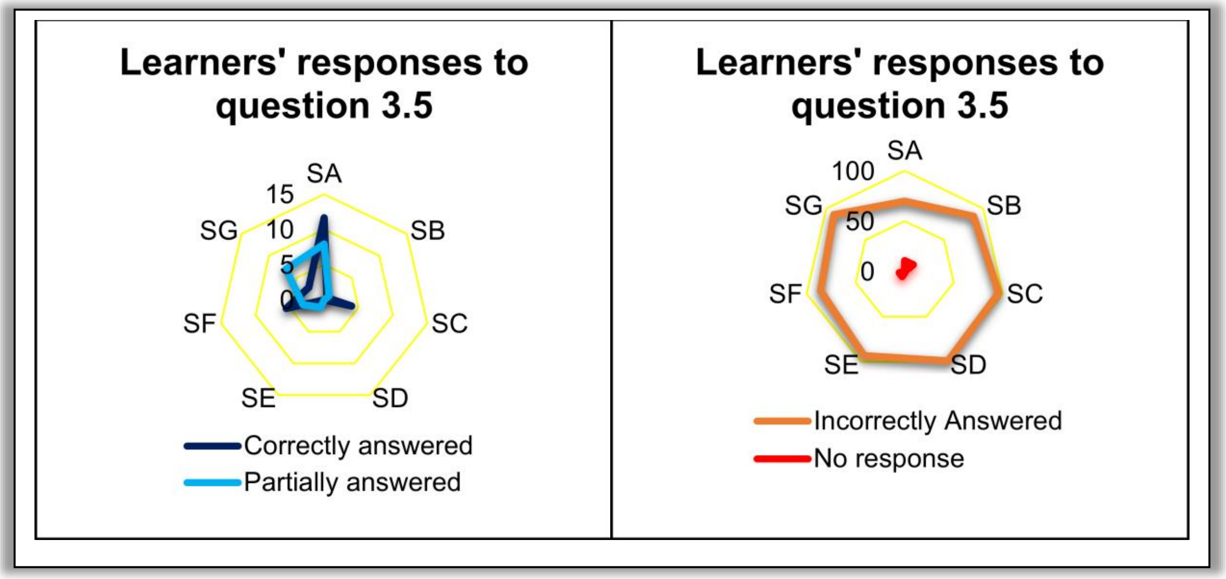
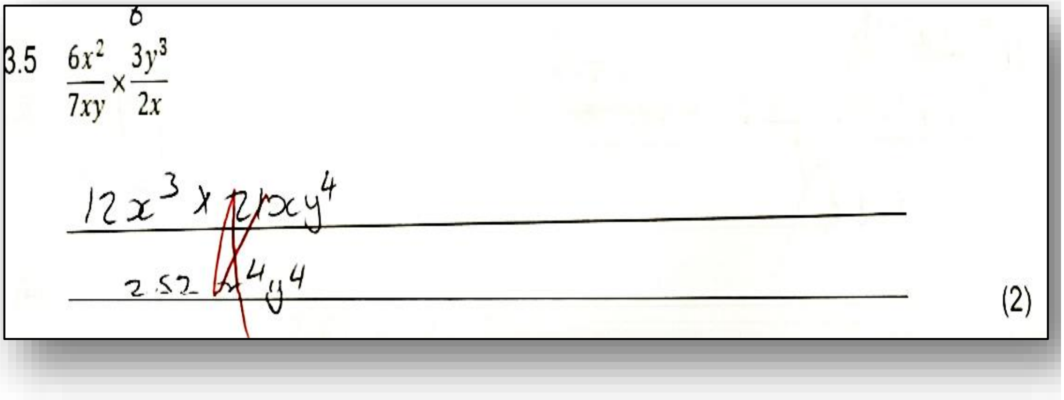
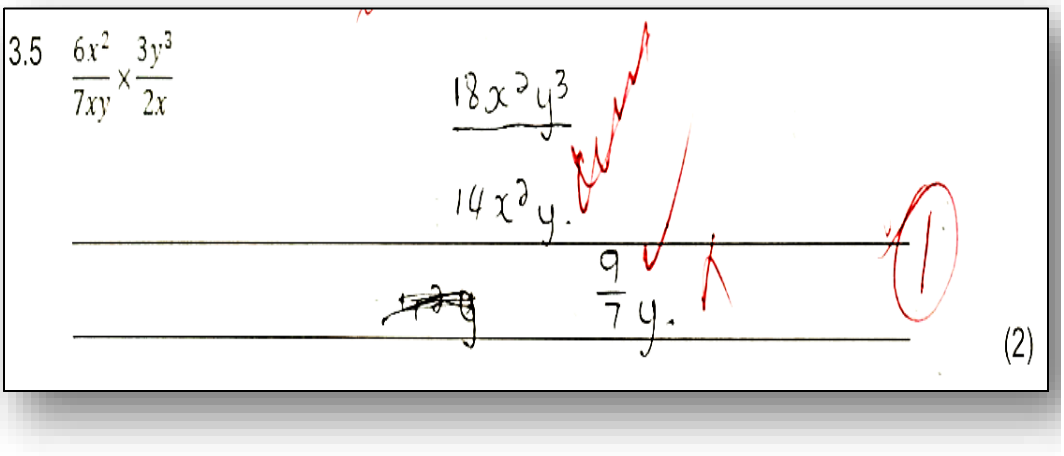


Figure 4.14: Trend in learners’ responses to question 3.5

Figure 4.15 posits on learners’ responses to question 3.5. This question was earlier coded CU1-SC1-SC3 (see Table 4.13) which shows that this question contained routine conceptual connections. The concepts that were comprehended in this question were: 1) multiplication and relative law of exponents; 2) division and relative law of exponents; and 3) factorisation.



Learner A



Learner B

Figure 4.15: Learners' responses to question 3.5

Phase one: the question provided f*lexibility*, by not specifying the procedure out of the following: 1) multiply first, then divide to simplify; 2) divide first, then multiply for simpler version; and 3) Factorise first a common factor for both numerator and denominator, then divide the common factor to remain with the simpler version. Learner A chose a procedure of multiplying across, hence the learner was wrong procedurally. Contrary, Learner B multiplied, then divided for the simplified version.

Phase two, *accuracy*, cross multiplied the terms in the expression for an example, $6x^2 \times 2x^2 = 12x^3 \; and \; 7xy \times 3y^2 \; and \; got \; 21xy^4$, Subsequently the learner multiplied $12x^3 \; and \; 21xy^2$ to get $252x^4y^4$. This indicates an incorrect procedure irrespective of the resultant multiplication. By contrast, Learner B, multiplied correct and the subsequent division was correct. Unfortunately, the marking indicated that the learners' final simplification was wrong when it was actually correct. Phase three, *efficiency*: Learner A is efficient with multiplication as exhibited in the response, but in the context of the question, fractions, the learner was not efficient by choosing a wrong procedure. By contrast, Learner B was efficient with the procedure.

Phase four, *appropriation*, for Learner A, the learner did not exhibit appropriation with the wrong procedure and Learner B exhibited appropriation. Learner A was coded SMP3.5B, one of 88.8% learners who were not proficient in question 3.5 in the seven sampled schools. Contrary, Learner B was one of 3.49% on learners' responses coded SMP3.5D who correctly answered question 3.5.

- **LEVELS OF MATHEMATICAL PROFICIENCY TO QUESTION 3**

Table 4.15, presents a summary of the levels of mathematical proficiency in question three. From the table it is evident that most learners (92.66%) were not proficient in Algebra and Algebraic Fractions questions. A very small proportion of learners (4.56% and 2.78%) were moderate and proficient in algebra and algebraic fractions questions respectively. Algebra and Algebraic Fractions are regarded as abstract content that assist learners to handle higher order questions (Dhlamini & Kibirige, 2014). This deficiency in proficiencies that learners exhibit in evaluative assessment such as ANA justifies urgent attention in the monitoring of curriculum reform.

Table 4.15: summary of learners' levels of mathematical proficiency to question 3

ANA 2012 Question 3 (Algebra and algebraic fractions)	**Codes for levels of mathematical proficiency**			
ANA questions	*Not Proficient*	*Moderate Proficient*	*Proficient*	*Total*
3.1 Simplify $2(x+2)^2-(2x-2)(x+2)$	(20)SMP3.1A (1125)SMP3.1B	(89)SMP3.1C	(16)SMP3.1D	1250
3.2 Simplify $\frac{15x^2y^3-9x^2y^3}{8x^2y^3}$	(47)SMP3.2A (1057)SMP3.2B	(73)SMP3.2C	(73)SMP3.2D	1250
3.3 Simplify $\frac{x^2-4x}{x^2-2x-8}$	(53)SMP3.3A (1165)SMP3.3B	(16)SMP3.3C	(16)SMP3.3D	1250
3.4 Simplify $\frac{x^2}{2}+\frac{2x^2}{3}-\frac{7x^2}{6}$	(88)SMP3.4A (1058)SMP3.4B	(74)SMP3.4C	(30)SMP3.4D	1250
3.5 Simplify $\frac{6x^2}{7xy}\times\frac{3y^3}{2x}$	(77)SMP3.5A (1101)SMP3.5B	(33)SMP3.5C	(39)SMP3.5D	1250
Totals	**5791**	**285**	**174**	**6250**
Percent	**92.66**	**4.56**	**2.78**	**100**

4.3.2 Learners' Responses to Question Six

In question 6, the question were on number patterns, a sub-topic of numbers and operations. The questions present solving number problems in exponential form (DBE, 2011). Learners require extensive computational fluencies when performing algorithms of numbers (Hecht & Vagi, 2012; Siegler, Thompson & Schneider, 2011). These are visible in their responses to questions that demand these proficiencies as in question 6. Table 4.16 is a synopsis of learners' responses to question 6 in 7 schools. The analysis classified learners' responses into four categories, correctly answered, partially answered, incorrect response and no response. These categories depicted the proficiency levels of learners for a variety of questions in 7 schools.

The third column in Table 4.16 shows results for learners in question 6 in school A. In the category correctly answered, 42.33% learners correctly answered question 6.1, followed by 28.84% for question 6.2 and last, 29.3% for question 6.3. For the category partially answered, 1.4% for question 6.1, followed by 0.93% for question 6.2 and 2.8% for question 6.3. In the category incorrectly answered, 41.86% answered question 6.1 wrong, 61.86% incorrectly answered question 6.2 and 53.02%

for question 6.3. In the category no response, there were 14.41% for question 6.1, then 8.37% learners for question 6.2 and 14.88% for question 6.3.

Column four of Table 4.16 shows the results for learners to question 6 in school B. The category correctly answered, 10.4% learners correctly answered question 6.1, followed by 3.3% for question 6.2 and 4.7% for question 6.3. In the category partially answered, no learners partially answered question 6.1, 6.2 and 6.3. In the category incorrectly answered, 82.1% learners answered question 6.1 wrong, followed by 85.4% for question 6.2 and last 83.5% for question 6.3. For the category no response, 7.5% for question 6.1, then 11.3% for question 6.3 and 11.8% for question 6.3.

In column five, Table 4.16 are results for school C for question six. In the category correctly answered, 8% learners correctly answered question 6.1, then 6% for question 6.2 and 4% for question 6.3. For the category partially answered, no learners partially answered question 6.1, 6.2 and 6.3. In the category incorrectly answered, 90% of learners answered question 6.1 wrong, 92% for question 6.2 and 94% for question 6.3. The category no response, there were 2% of learners for question 6.1, 6.2 and 6.3.

In column six, Table 4.16, are the results for school D for question six. In the category correctly answered, 6.3% learners correctly answered question 6.1, then 1.1% for question 6.2 and 0.53% for question 6.3. For the category partially answered, 2.1% for question 6.1, then no learners partially answered question 6.2 and 6.3. The category incorrectly answered, 88.4% learners answered question 6.1 wrong, then 96.8% learners for question 6.2 and 95.24% for question 6.3. In the category no response, there were 3.2% learners who had no responses for question 6.1, followed by 2.1% for question 6.2 and 4.23% for question 6.3.

Column seven in Table 4.16, outlines results for school E in question six. In the category correctly answered, there were 9.2% learners who correctly answered question 6.1, then 1.44% for question 6.2 and 0.3% for question 6.3. The category partially answered, 0.6% learners partially answered question 6.1, no learners for question 6.2 and 1.7% for question 6.3. For the category incorrectly answered, 84.8%

“What You Test Is What You Get, (WYTIWUG).” (Schoenfeld, 2007: 72) learners answered question 6.1 wrong, then 90.54% for question 6.2 and 87.4% for question 6.3. The category no response, 5.4% for question 6.1, then (8.02%) for question 6.2 and 10.6% for question 6.3.

In column eight, Table 4.16, are the results for school F for question 6. In the category correctly answered, 17.79% of learners correctly answered question 6.1, then 7.97% for question 6.2 and 7.36% for question 6.3. In the category partially answered, 1.23% partially answered question 6.1, then no learner for question 6.2 and 0.61% for question 6.3. For the category incorrectly answered, 66.87% for question 6.1, then 71.78% for question 6.2 and 64.42% for question 6.3. In the category no response, 14.11% had no responses for question 6.3, then 20.25% for question 6.2 and 27.61% for question 6.3.

In column nine, Table 4.16, are the results for school G for question 6. The category correctly answered, 13.89% learners correctly answered question 6.1, then 2.78% for question 6.2 and 4.17% for question 6.3. For the category partially answered, no learners partially answered question 6.1, 6.2 and 6.3. In the category incorrectly answered, 86.11% learners answered question 6.1 wrong, then 97.22% for question 6.2 and 95.83% for question 6.3. In the category no response, there were no learners did not respond to question 6.1, 6.2 and 6.3.

In column ten, Table 4.16, shows means and standard deviations for learners' responses to the three levels of question six in four categories. For question 6.1, in the category ‘correctly answered’, the mean is 15.42 and the standard deviation is 12.48. School A is in the range,$+2\sigma < x < +3\sigma$, schools B, C, D, E and G are in the range,$-1\sigma < x < 0$, and school F is in the range,$0 < x < +1\sigma$. In the category ‘partially answered’, the mean is 0.76 and the standard deviation is 0.84. Schools A, E and F are in the range,$0 < x < +1\sigma$, schools B, C and G are in the range,$-1\sigma < x < 0$, and school D is in the range,$+1\sigma < x < +2\sigma$. The category ‘incorrectly answered’, the mean is 77.16 and the standard deviation is 17.35. School A is in the range,$-2\sigma < x < -3\sigma$, schools B, C, D, E and G are in the range,$0 < x < +1\sigma$, and school F is in the range,$-1\sigma < x < 0$. In the category ‘no response’, the mean is 6.66

and the standard deviation is 5.71. Schools A and F are in the range, $+1\sigma < x < +2\sigma$, schools C, D and E are in the range, $-1\sigma < x < 0$, school B is in the range, $0 < x < +1\sigma$, and school G is in the range, $-2\sigma < x < -1\sigma$.

For question 6.2, in the category ‘correctly answered’, the mean is 7.35 and the standard deviation is 9.79. In the category ‘partially answered’, the mean is 0.13 and the standard deviation is 0.35. in the category ‘incorrectly answered’, the mean is 85.09 and the standard deviation is 13.41. In the category ‘no response’, the mean is 8.67 and the standard deviation is 6.7.

For question 6.3, in the category ‘correctly answered’, the mean is 7.19 and the standard deviation is 10.05. In the category ‘partially answered’, the mean is 0.73 and the standard deviation is 1.11. In the category ‘incorrectly answered’, the mean is 81.92 and the standard deviation is 16.79. In the category ‘no response’, the mean is 10.16 and the standard deviation is 9.44.

Table 4.16: Learners' responses to question 6

ANA 2014 GRADE 9 MATHEMATICS LEARNERS' RESPONSES IN VARIOUS SCHOOLS									
Question 6 Number Patterns	*Learners' responses*	*SA* (F) %	*SB* (F) %	*SC* (F) %	*SD* (F) %	*SE* (F) %	*SF* (F) %	*SG* (F) %	*Standard Deviations & Means* $(\bar{x})\sigma$
6.1	Correctly answered	(91)42.33	(22)10.4	(4)8	(12)6.3	(32)9.2	(29)17.79	(10)13.89	(15.42)12.48
Complete the table	Partially answered	(3)1.4	(0)0	(0)0	(4)2.1	(2)0.6	(2)1.23	(0)0	(0.76)0.84
below.	Incorrectly answered	(90)41.86	(174)82.1	(45)90	(167)88.4	(296)84.8	(109)66.87	(62)86.11	(77.16)17.35
(n=1250)	No response	(31)14.41	(16)7.5	(1)2	(6)3.2	(19)5.4	(23)14.11	(0)0	(6.66)5.71
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
6.2	Correctly answered	(62)28.84	(7)3.3	(3)6	(2)1.1	(5)1.44	(13)7.97	(2)2.78	(7.35)9.79
Write down the	Partially answered	(2)0.93	(0)0	(0)0	(0)0	(0)0	(0)0	(0)0	(0.13)0.35
general term T_n of the	Incorrectly answered	(133)61.86	(181)85.4	(46)92	(183)96.8	(316)90.54	(117)71.78	(70)97.22	(85.09)13.41
above number pattern.	No response	(18)8.37	(24)11.3	(1)2	(4)2.1	(28)8.02	(33)20.25	(0)0	(8.67)6.70
(n=1250)	**Total scripts**	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
6.3	Correctly answered	(63)29.3	(10)4.7	(2)4	(1)0.53	(1)0.3	(12)7.36	(3)4.17	(7.19)10.05
If $T_n = 512$, determine	Partially answered	(6)2.8	(0)0	(0)0	(0)0	(6)1.7	(1)0.61	(0)0	(0.73)1.11
the value of n.	Incorrectly answered	(114)53.02	(177)83.5	(47)94	(180)95.24	(305)87.4	(105)64.42	(69)95.83	(81.92)16.79
(n=1250)	No response	(32)14.88	(25)11.8	(1)2	(8)4.23	(37)10.6	(45)27.61	(0)0	(10.16)9.44
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77

Table 4.17 presents an outline of mathematical activities that were examined by question 6. There are also codes of SMP that these questions demanded. For more detail on the coding of SMP see Table 3.3.

Table 4.17: SMP demanded by question 6

Strands of mathematical proficiency in question 6		
Question	*Mathematical activity*	*Codes*
6.1 Complete the table below.	Evaluating consecutive terms of a pattern, multiplication.	SP-SC1-SC3
6.2 Write down the general term T_n of the above number pattern.	Finding the general formula of a pattern, write using variables.	SP-SC1-SC3-AR2
6.3 If $T_n = 512$, determine the value of n.	Substitution in the formula, finding the cube.	PF2-SC1-SC3

Table 4.18 explain with codes SMP that were required by ANA and are likely to be exhibited by learners as they respond to the three parts of question six. These codes explain the proficiency levels tested by ANA and are key in the analysis and categorising of learners' responses to question 6. The codes were derived from the generic of SMP as shown in Table 3.3 and the suggested answers for the ANA test.

Table 4.18: Explanation of learners' mathematical proficiencies in the three parts of question 6

Strands of mathematical proficiency analysis key on Question 6			
SMP6.1A No response	SMP6.1A Incorrect calculation of values in the sequence	SMP6.1C One of the values is incorrect	SMP6.1D Correct calculation of values of the sequence
SMP6.2A No response	SMP6.2B Incorrect formula	SMP6.2C Not simplified formula	SMP6.2D Correct formula
SMP6.3A No response	SPM6.3B Wrong statement, substitution and simplification	SMP6.3C Either, wrong statement, substitution or simplification	SMP6.3D Correct statement, substitution and simplification

❖ LEARNERS' RESPONSES TO QUESTION 6.1

Figure 4.16 illustrates the trend in learners' responses to question 6.1. The category no response is centred showing that in the distribution, there were a few or no learners in the schools who had no responses to question 6.1. In the category partially answered, the distribution is centred which also show a few or no learners in the schools who partially answered question 6.1. For the category incorrectly answered, the distribution is almost a regular heptagon with the exception of School A (42.33% incorrectly answered). This indicates that most learners incorrectly answered question 6.1. Last, for the category correctly answered, the distribution is irregular and shows that most of the schools had less than 10%, correctly answered with an exception of School A (44.23% correctly answered). The vignettes that I used were from three learners because the current researcher wanted to explore the consistency in question 6.1, 6.2 and 6.3 which were follow-on questions.

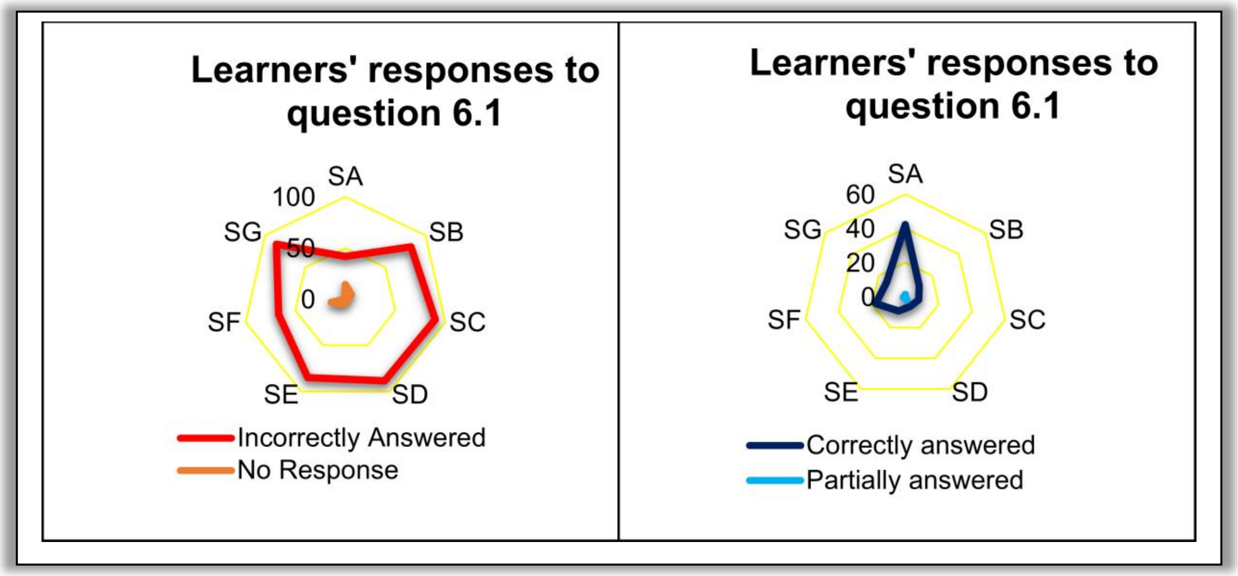
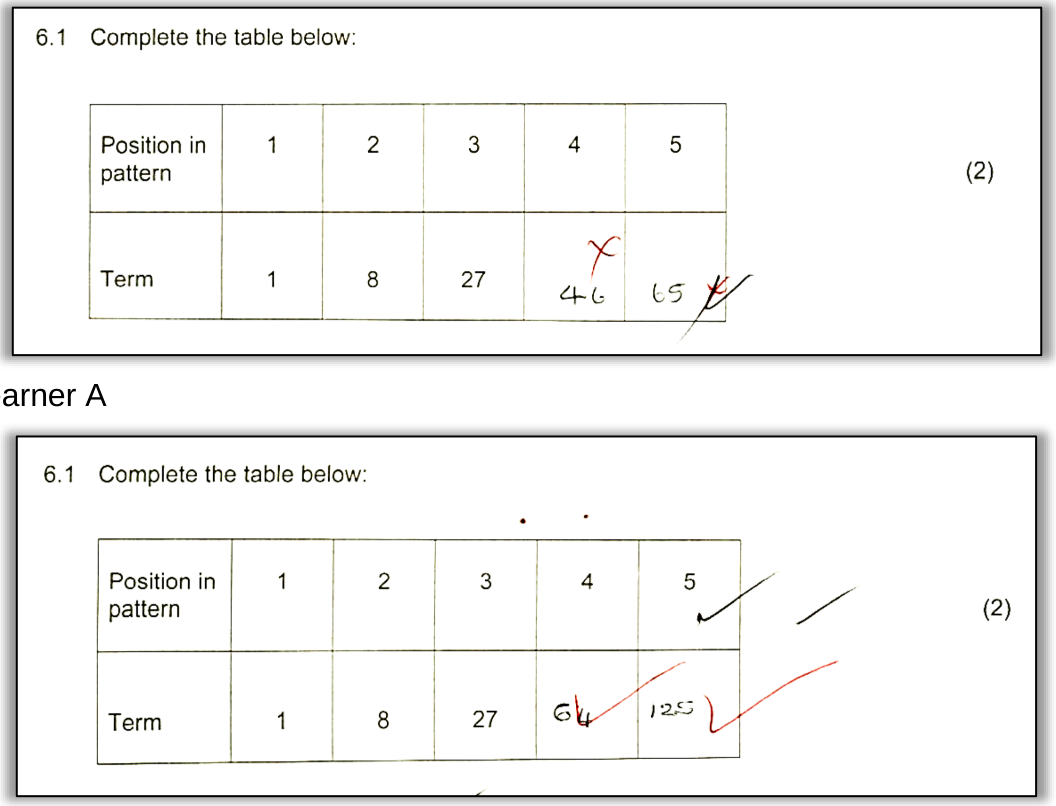


Figure 4.16: Trend in learners' responses to question 6.1

The vignettes in Figure 4.17 are learners' responses to question 6.1. This question was earlier coded SP-SC1-SC3 (see Table 4.17) which shows that this question required routine simple procedure. The procedure required knowledge of Number Patterns. Specifically, in Grade 8 and 9, learners are expected to use their own judgement to generalise Sequences.



6.1 Complete the table below:

Position in pattern	1	2	3	4	5
Term	1	8	27	46	65

(2)

Learner A

6.1 Complete the table below:

Position in pattern	1	2	3	4	5
Term	1	8	27	64	125

(2)

Learner B

Figure 4.17: Learners' responses to question 6.1

Phase one, *flexibility*: the question does not specify a procedure to follow to generate the subsequent terms of the sequence, hence the question is flexible. Learner A generates the subsequent terms by adding the common difference. This procedure is applicable to linear sequences and not cubic sequences. Learners in Grade 9 are expected to use their discretion (not in a systematic way) to complete numeric and geometric sequences (DBE, 2011). Only in Grade 10 learners are expected to use systematic formulae for linear and geometric sequences (DBE, 2011). As such, this learner is applying a Grade 10 procedure. By contrast, Learner B used the procedure of looking at how the term is generated from the position in the pattern. In Phase two, *accuracy*, Learner A, computed the terms of the sequence as follows; $27 - 8 = 19$ to generate the subsequent terms $46\ and\ 65$, correct computations for linear sequences, but wrong procedure for this question, hence the answers are incorrect. Learner B, computed as follows; $2 \times 2 \times 2 = 8\ and\ 3 \times 3 \times 3 = 27$ then $4 \times 4 \times 4 = 64\ and\ 5 \times 5 \times 5 = 125$, correct procedure and correct computations. Phase three, *efficiency*: Learner B showed efficiency but by contrast Learner A was not efficient due to wrong procedure followed. Phase four, *appropriation*: Learner A did not exhibit appropriation due to misapplication of procedure. By contrast, Learner B exhibited appropriation by choosing the correct procedure and applied it at the right time. As such, Learner A was one of 81.92% coded SMP6.1B, learners who were not proficient in question 6.1. Learner B was one of 15.42% learners coded SMP6.1D who were proficient in question 6.1.

❖ LEARNER'S RESPONSES TO QUESTION 6.2

Figure 4.18 shows the trend in learners' responses to question 6.2. The category partially answered is centred, which shows that there were very few or no learners who partially answered question 6.2. In the category correctly answered, the distribution is irregular, in schools E and D it is centred meaning no learners correctly answered question 6.2 in these Schools. In school A, the results were (28.84% correctly answered), school B (3.3% correctly answered), school C (6% correctly answered), school F (7.97% correctly answered) and school G (2.78% correctly answered). In the category incorrectly answered, the distribution is almost a regular heptagon except for schools A and F. This is an indication that in most schools majority of learners incorrectly answered question 6.2. Last, the category no response is irregular and almost centred. This indicates that there were a few or no learners in the schools that had no responses to question 6.2.

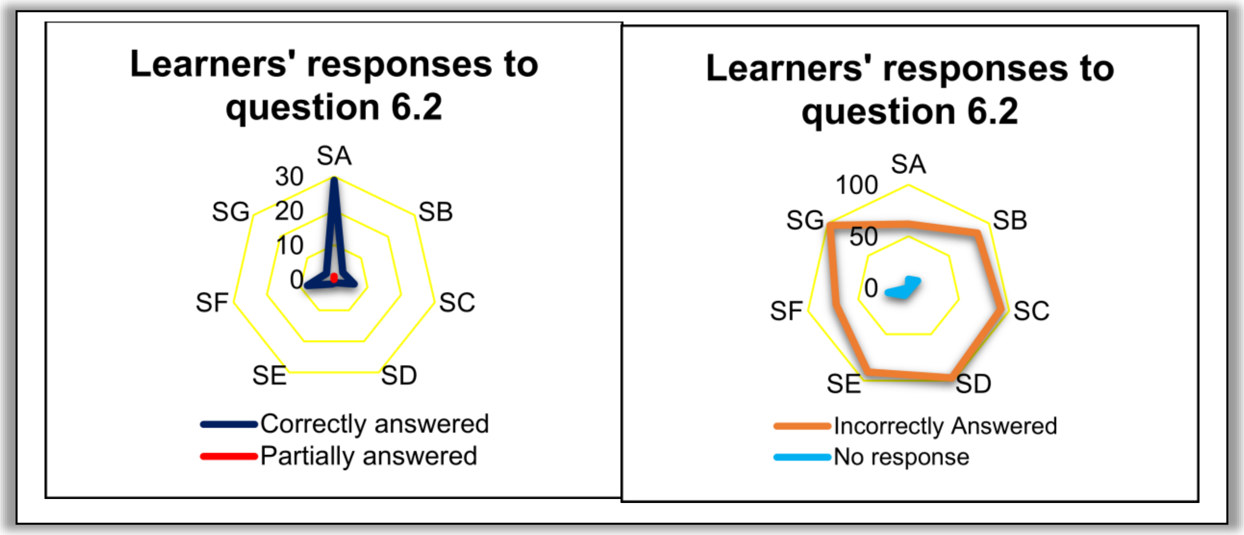


Figure 4.18: Trend in learners' responses to question 6.2

The vignettes in Figure 4.19 are learners' responses to question 6.2. These responses are from the same learners in the previous question (6.1) and I also use them in question 6.3 because these were follow-on questions. This question was earlier coded SP-SC1-SC3-AR2 (see Table 4.17) which shows that this question involved routine simple procedure which allowed learners to conjecture. The procedure required knowledge of Number Patterns, especially the ability to generalise number patterns.

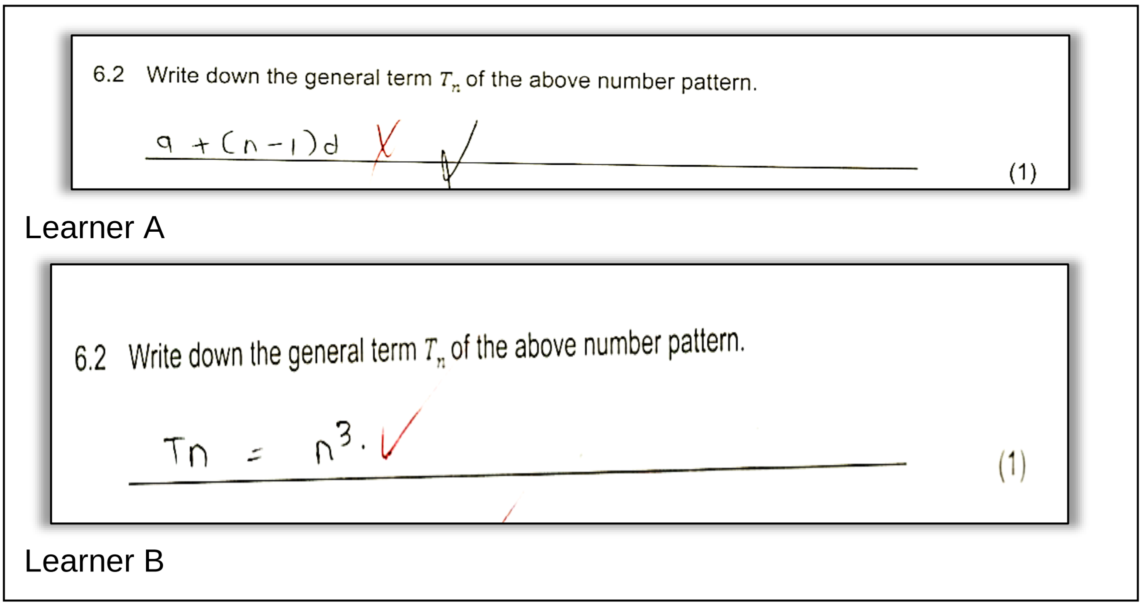


Figure 4.19: Learners' responses to question 6.2

Phase one, *flexibility*: this question followed-on from the previous one and learners' responses were dependant on answers for question 3.1. Hence, the question was only flexible to the discursive learners' responses exhibited in the previous question. Learner A misapplied the procedure for linear sequence in cubic sequence and Learner B applied procedure for cubic sequences. Phase two, *accuracy*: Learner A followed a procedure of finding the general formula a linear sequence which was not consistent with the given pattern (cubic sequence). Learner B correctly wrote the general term of the cubic sequence which was a generalisation from the response in the previous question (6.1). Phase three, *efficiency*: Learner A was not efficient due to the wrong choice of procedure and Learner B was efficient by exhibiting correct generalisation of the cubic sequence. Phase four, *appropriation*: Learner A did not exhibit appropriation due to an incorrect procedure. Learner B exhibited appropriation by knowing the correct procedure, and when and how to use it. As such, Learner A was one of 85.09% learners' responses coded SMP6.2B those not proficient in question 6.2. Subsequently, Learner B was one of 7.35% learners' responses coded SMP6.2D for learners who were proficient in question 6.2.

❖ LEARNERS' RESPONSES TO QUESTION 6.3

Figure 4.20 illustrates the trend in learners' responses to question 6.3. The category partially answered is almost centred showing that very few or no learners partially answered question 6.3. The category correctly answered is irregular and shows values less than 10% in most schools except for school A (29.3% correctly answered). The category incorrectly answered is almost a regular heptagon except for school A, (53.02% incorrectly answered) and school F (64.42% incorrectly answered). Last, the category no response, is irregular and is centred for some schools. This is an indication that there were very few or no learners who had no responses to question 6.3.

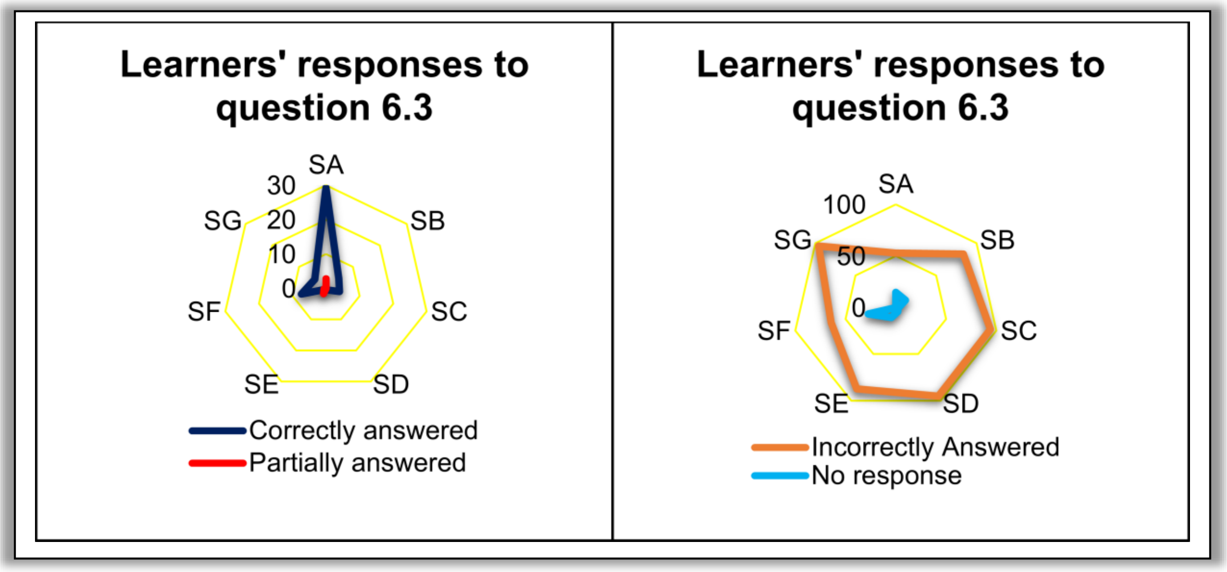


Figure 4.20: Trend in learners' responses to question 6.3

The vignettes in Figure 4.21 are responses by Learner A and Learner B to question 6.3. This question was earlier coded PF2-SC1-SC3 (see Table 4.17) which shows that this question was a routine algorithm.

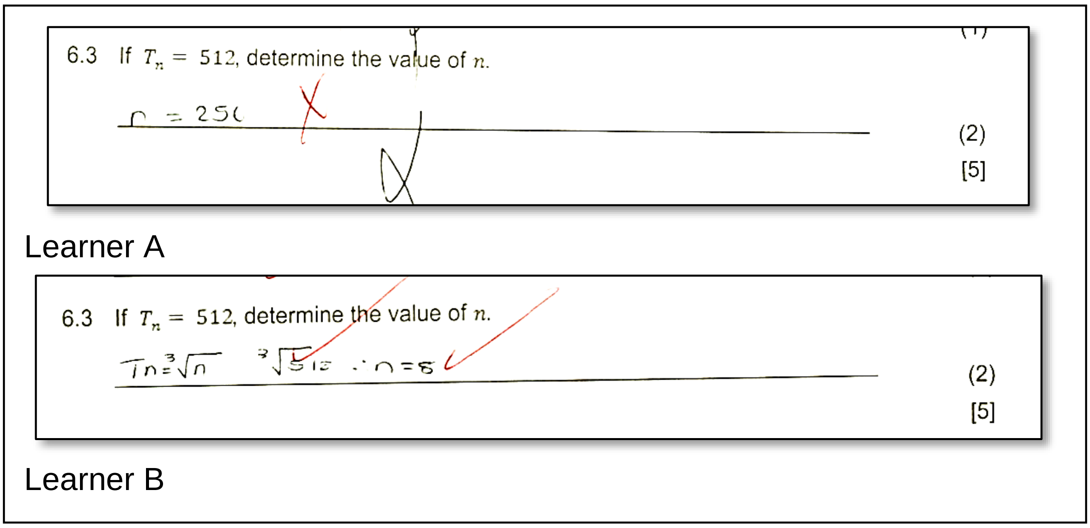


Figure 4.21: Learners' response to question 6.3, SMP6.3B

Phase one, *flexibility*: the question is flexible because it does not specify the procedure to be followed. The answer written by Learner A is not clear on the procedure used. By contrast, Learner B used the formula that was computed in question 6.2. Phase two, *accuracy*: if it is assumed that Learner A followed on the previous answers (Question 6.1 & 6.2), and substituting the values, the answer does not yield the answer written by the learner here. Hence Learner A did not exhibit accuracy. By contrast, Learner B, started with a wrong statement, $T_n = \sqrt[3]{n}$, since '512' is T_n and not 'n'. This statement was supposed to read as $T_n = n^3 \therefore n = \sqrt[3]{T_n}\ then\ n = \sqrt[3]{512}$. Fortunately the second statement $\sqrt[3]{512}$ is accurate, then Learner B exhibited accuracy. Phase three, *efficiency*: Learner A was not efficient by not consistently applying the procedure from the previous follow-on questions. By contrast, Learner B, was efficient irrespective of the incorrect statement in the solution strategy. Phase four, *appropriation*: Learner A did not exhibit appropriation by not being consistent in the procedure used and by contrast Learner B exhibited appropriation due to the consistency in the procedure to achieve the outcome. As such, Learner A was one of 81.92% coded SMP6.3B for learners who were not proficient in question 6.3. Subsequently during the coding, Learner B was one of 81.92% one of 7.19% coded SMP6.3D, for learners that were proficient in question 6.3.

❖ **LEVELS OF MATHEMATICAL PROFICIENCY TO QUESTION 6**

In Table 4.19, is a summary of the levels of mathematical proficiency to question 6. From the table it is evident that most learners (89.01%) were not proficient in number patterns' questions. A very small proportion of learners (0.7% and 10.29%) were moderate and proficient in number patterns questions respectively.

Table 4.19: Summary of learners' levels of mathematical proficiency to question 6

ANA 2012 Question 6 (Number patterns)	**Codes for levels of mathematical proficiency**			
ANA questions	*Not Proficient*	*Moderately Proficient*	*Proficient*	*Total*
6.1 Complete the table below.	(96)SMP6.1A (943)SMP6.1B	(11)SMP6.1C	(200)SMP6.1D	1250
6.2 Write down the general term T_n of the above number pattern	(108)SMP6.2A (1046)SMP6.2B	(2)SMP6.2C	(94)SMP6.2D	1250
6.3 If $T_n = 512$, determine the value of n.	(148)SMP6.3A (997)SMP6.3B	(13)SMP6.3C	(92)SMP6.3D	1250
Totals	**3338**	**26**	**386**	**3750**
Percent	**89.01**	**0.70**	**10.29**	**100**

4.3.3 Learners' Responses to Question Ten

For question ten, the learners responded to Geometry questions that examined analogical reasoning. Geometry is difficult to learn and teach because of its abstract mathematical proofs that learners must read, write and understand (Luneta, 2015b). Learners often perform errors and misconceptions in Geometry problems because they have to deal with Three-dimensional problems on a Three-dimensional plane (Palha et al., 2013; Soto-Johnson & Troup, 2014). Analogical reasoning entails reasoning about commonalities among mathematical relations (Amir-Mofidi et al., 2012). These are visible in learners' responses to questions that demand this proficiency as in question 10. Table 4.20 is a synopsis of learners' responses to question ten in seven schools. The analysis classified learners' responses into four categories, correctly answered, partially answered, incorrect response and no response. These categories depicted the proficiency levels of learners for a variety of questions in seven schools.

In column three, Table 4.20 are learners' responses to question t10 in school A. In the category correctly answered, 4.65% of learners correctly answered question 10.2, then 4.19% for question 10.3.1, followed by 15.81% for question 10.3.2 and 4.19% for question 10.4.1. For the category partially answered, 25.58% learners partially answered question 10.2, then 14.88% for question 10.3.1, followed by (38.61%) for question 10.3.2 and 24.65% for question 10.4.1. In the category, incorrectly answered, 50.7% learners answered question 10.2 wrong, 67.91% for question 10.3.1, then 30.23% for question 10.3.2 and 49.3% for question 10.4.1. In the category no response, there were 19.07% learners who did not respond to question 10.2, then 13.02% for question 10.3.1, followed by 15.35% for question 10.3.2 and 21.86% for question 10.4.1.

Column four, Table 4.20 shows results for learners' responses to question ten for School B. There were no learners who correctly answered question 10.2, 10.3.1, 10.3.2 and 10.4.1. For the category partially answered, 1.4% of learners partially answered question 10.2, then 2.8% learners for question 10.3.1, and 0.9% for both question 10.3.2.and question 10.4.1. In the category incorrectly answered, 67% learners incorrectly answered question 10.2, then 62.3% for question 10.3.1, followed by (53.8%) for question 10.3.2 and 55.2% for question 10.4.1. The category no response, 31.6% learners did not respond to question 10.2, then 34.9% for question 10.3.1, followed by 45.3% for question 10.3.2 and 43.9% for question 10.4.1.

In column five, Table 4.20 there are results for learners' responses to question ten for School C. In the category correctly answered, no learners correctly answered question 10.2, 10.3.1, 10.3.2 and 10.4.1. For the category partially answered, 8% learners partially answered question 10.2, then 16% for question 10.3.1, followed by 8% for question 10.3.2, and 20% for question 10.4.1. In the category incorrectly answered, 90% of learners incorrectly answered question 10.2, then 82% for question 10.3.1 and question 10.3.2, and 68% for question 10.4.1. In the category no response, 2% of learners did not answer question 10.2 and question 10.3.1, followed by 10% for question 10.3.2 and 12% for question 10.4.1.

In column six, Table 4.20 are the results for learners' responses to question 10 for school D. For the category correctly answered, no learners correctly answered question 10.2, 10.3.1, 10.3.2 and 10.4.1. In the category partially answered, 12.2% learners partially answered question 10.2, then 9.5% for question 10.3.1, followed by 16.9% for question 10.3.2 and 6.3% for question 10.4.1. The category incorrectly answered, 78.3% of learners answered incorrectly question 10.2, followed by 85.7% for question 10.3.1, then 74.1% for question 10.3.2 and 85.2% for question 10.4.1. In the category no response, 9.5% of learners did not answer question 10.2, followed by 4.8% who did not answer question 10.3.1, then 9% for question 10.3.2, and 8.5% for question 10.4.1.

Column seven, Table 4.20, shows the results for learners' responses to question 10 for school E. In the category correctly answered, there were no learners who answered question 10.2 correct, then 0.9% for question 10.3.1, followed by 0.3% for question 10.3.2 and 1.15% for question 10.4.1. For the category partially answered, 8.3% partially answered question 10.2, followed by 3.2% for question 10.3.1, then 8.6% for question 10.3.2, and 5.73% for question 10.4.1. In the category incorrectly answered, 84% of learners got question 10.2 wrong, then 86.2% for question 10.3.1, followed by 74.8% for question 10.3.2 and 79.08% for question 10.4.1. In the category no response, 7.7% of learners did not respond to question 10.2, followed by 9.7% for question 10.3.1, then 16.3% for question 10.3.2 and 14.04% for question 10.4.1.

In column eight, Table 4.20 are the results for learners' responses to question 10 for school F. In the category correctly answered, there were no learners who correctly answered question 10.2 and 10.4.1, then 2.45% for question 10.3.1 and 0.61% for question 10.3.2. For the category partially answered, 9.82% partially answered question 10.2, followed by 6.75% for question 10.3.1, then 14.11% for question 10.3.2 and 7.36% for question 10.4.1. In the category incorrectly answered, 75.46% of learners incorrectly answered question 10.2, then 73.62% for question 10.3.1, followed by 68.1% for question 10.3.2, and 68.71% for question for question 10.4.1. In the category no response, 14.72% of learners did not answer question 10.2, then 17.18% for question 10.3.1 and question 10.3.2 and last 23.93% for question 10.4.1.

In the column nine, Table 4.20 are the results for learners' responses to question 10 for school G. For the category correctly answered, 1.39% of learners correctly answered question 10.2, 10.31, 10.32 and 10.4.1. In the category partially answered, 26.39% of learners partially answered question 10.2, then 15.28% for question 10.3.1, followed by 27.78% for question 10.3.2 and 12.5% for question 10.4.1. For the category incorrectly answered, 72.22% of learners incorrectly answered question 10.2, followed by 83.33% for question 10.3.1, then 70.83% for question 10.3.2 and 86.11% for question 10.4.1. The category no response, there were no learners who did not respond to question 10.2, 10.3.1, 10.3.2 and 10.4.1.

Column ten, Table 4.20, shows means and standard deviations for learners' responses to the four levels of question ten in four categories. For question 10.2, category correctly answered, the mean is 0.86 and the standard deviation is 1.75. For the category partially answered, the mean is 13.1 and the standard deviation is 9.4. For the category incorrectly answered, the mean is 73.95 and the standard deviation is 12.73. For the category no response, the mean is 12.08 and the standard deviation is 10.88. For question 10.3.1, in the category correctly answered, the mean is 1.28 and the standard deviation is 1.58. In the category partially answered, the mean is 9.77 and the standard deviation is 5.72. In the category incorrectly answered, the mean is 77.29 and the standard deviation is 9.44. In the category no response, the mean is 11.66 and the standard deviation is 11.92.

For question 10.3.2, in the category correctly answered, the mean is 2.59 and the standard deviation is 5.85. In the category partially answered, the mean is 16.41 and the standard deviation is 12.91. In the category incorrectly answered, the mean is 64.84 and the standard deviation is 17.54. In the category no response, the mean is 16.16 and the standard deviation is 14.15. For question 10.4.1, in the category correctly answered, the mean is 0.96 and the standard deviation is 1.54. For the category partially answered, the mean is 11.06 and the standard deviation is 8.51. In the category incorrectly answered, the mean is 70.23 and the standard deviation is 14.29. In the category no response, the mean is 17.75 and the standard deviation is 14.06.

Table 4.20: Learners' responses to question 10

ANA 2014 GRADE 9 MATHEMATICS LEARNERS' RESPONSES IN VARIOUS SCHOOLS									
Question 10 Concurrency and Similarity	*Learners' responses*	*SA (F) %*	*SB (F) %*	*SC (F) %*	*SD (F) %*	*SE (F) %*	*SF (F) %*	*SG (F) %*	*Standard Deviations & Means* $(\bar{x})\sigma$
10.2	Correctly answered	(10)4.65	(0)0	(0)0	(0)0	(0)0	(0)0	(1)1.39	(0.86)1.75
Proving that two sides are	Partially answered	(55)25.58	(3)1.4	(4)8	(23)12.2	(29)8.3	(16)9.82	(19)26.39	(13.10)9.40
equal	Incorrectly answered	(109)50.7	(142)67	(45)90	(148)78.3	(293)84	(123)75.46	(52)72.22	(73.95)12.73
(n=1250	No response	(41)19.07	(67)31.6	(1)2	(18)9.5	(27)7.7	(24)14.72	(0)0	(12.08)10.88
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
10.3.1	Correctly answered	(9)4.19	(0)0	(0)0	(0)0	(3)0.9	(4)2.45	(1)1.39	(1.28)1.58
Proving that two lines are	Partially answered	(32)14.88	(6)2.8	(8)16	(18)9.5	(11)3.2	(11)6.75	(11)15.28	(9.77)5.72
equal	Incorrectly answered	(146)67.91	(132)62.3	(41)82	(162)85.7	(301)86.2	(120)73.62	(60)83.33	(77.29)9.44
(n=1250)	No response	(28)13.02	(74)34.9	(1)2	(9)4.8	(34)9.7	(28)17.18	(0)0	(11.66)11.92
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
10.3.2	Correctly answered	(34)15.81	(0)0	(0)0	(0)0	(1)0.3	(1)0.61	(1)1.39	(2.59)5.85
Proving that two triangles	Partially answered	(83)38.61	(2)0.9	(4)8	(32)16.9	(30)8.6	(23)14.11	(20)27.78	(16.41)12.91
are congruent (n=1250)	Incorrectly answered	(65)30.23	(114)53.8	(41)82	(140)74.1	(261)74.8	(111)68.1	(51)70.83	(64.84)17.54
	No response	(33)15.35	(96)45.3	(5)10	(17)9	(57)16.3	(28)17.18	(0)0	(16.16)14.15
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77
10.4.1	Correctly answered	(9)4.19	(0)0	(0)0	(0)0	(4)1.15	(0)0	(1)1.39	(0.96)1.54
Proving that two triangles	Partially answered	(53)24.65	(2)0.9	(10)20	(12)6.3	(20)5.73	(12)7.36	(9)12.5	(11.06)8.51
are similar	Incorrectly answered	(106)49.3	(117)55.2	(34)68	(161)85.2	(276)79.08	(112)68.71	(62)86.11	(70.23)14.29
(n=1250)	No response	(47)21.86	(93)43.9	(6)12	(16)8.5	(49)14.04	(39)23.93	(0)0	(17.75)14.06
	Total scripts	**(215)**	**(212)**	**(50)**	**(189)**	**(349)**	**(163)**	**(72)**	(178.57)99.77

Table 4.21 is an outline of mathematical activities that were examined by question 10. There are also codes of SMP that these questions demanded. For more detail on the coding of SMP see Table 3.3.

Table 4.21: SMP examined by question 10

Strands of mathematical proficiency in question 10		
Question	*Mathematical activity*	*Codes*
10.2 Proving that two sides are equal	Show equal radii, showing a common side, showing two right angles in perpendicular lines, give valid reasons.	SP-SC1-SC2-SC3-AR3
10.3.1 Proving that two lines are equal	Showing equal given sides, showing that BF + FC = CE + FC, give valid reasons.	SP-SC1-SC2-SC3-AR3
10.3.2 Proving that two triangles are similar	Showing with valid reasons three pairs of equal sides.	SP-SC1-SC2SC3-AR3
10.4.1 Proving that two triangles are similar	Showing with valid reasons three pairs of equal sides.	SP-SC1-SC2-SC3-AR3

Table 4.22 explains with codes SMP that are likely to be exhibited by learners as they respond to the four parts of question 10. The codes are divided into four categories for each part of the question. These codes are key to the analysis and categorising of learners' responses to question ten. The codes were derived from the general codes of SMP shown in Table 3.3 and the suggested answers for the ANA test.

Table 4.22: Explanation of learners' SMP to the four parts of question 10

SMP analysis key on Question 10			
SMP10.2A No response	SMP10.2B Incorrect procedure of proof.	SMP10.2C Either did not show with reasons, equal radii, common side, equal right angles or conclusion.	SMP10.2D Correctly show with reasons equal radii, common side, equal right angles and conclusion.
SMP10.3.1A No response	SMP10.3.1B Incorrect procedure of proof.	SMP10.3.1C Either did not show $BF = CE$ or $BF + FC = CE + FE$ and cannot show $BC = EF$	SMP10.3.2D Correctly show $BF = CE$ as given and $BF + FC = CE + FE$ resulting in $BC = EF$
SMP10.3.2A No response	SMP10.3.2B Incorrect procedure of proof.	SMP10.3.2C Either did not show 2 pairs of equal sides or that $BC = EF$ as proved	SMP10.3.2D Correctly show 2 pairs of equal given sides, and show $BC = EF$ as proved.
SMP10.4.1A No response	SMP10.4.1B Incorrect procedure of proof.	SMP10.4.1C Either, incorrect reasons, or pair of equal angles.	SMP10.4.1D Showing with reasons common angle, 2 equal pair of angles.

❖ **LEARNERS' RESPONSES TO QUESTION 10.2**

Figure 4.22 is an illustration of learners' responses to question 10.2. The category correctly answered is centred, showing a distribution with very low or zero learners who correctly answered question 10.2. The category incorrectly answered shows a distribution that is an irregular heptagon. This shows, that a high number of learners incorrectly answered question 10.2 with the exception of school A, (50.7% incorrectly answered) and school B (67% incorrectly answered). In the category partially answered, the distribution is irregular, which shows less or no learners who partially answered question. An exception is in school A, (25.58% partially answered) and school G (26.39% partially answered). Finally, in the category no response, the distribution is also irregular and shows less or no learners who did not respond to question 10.2. An exception was in school A, (19.07% no response) and school B (31.6% no response).

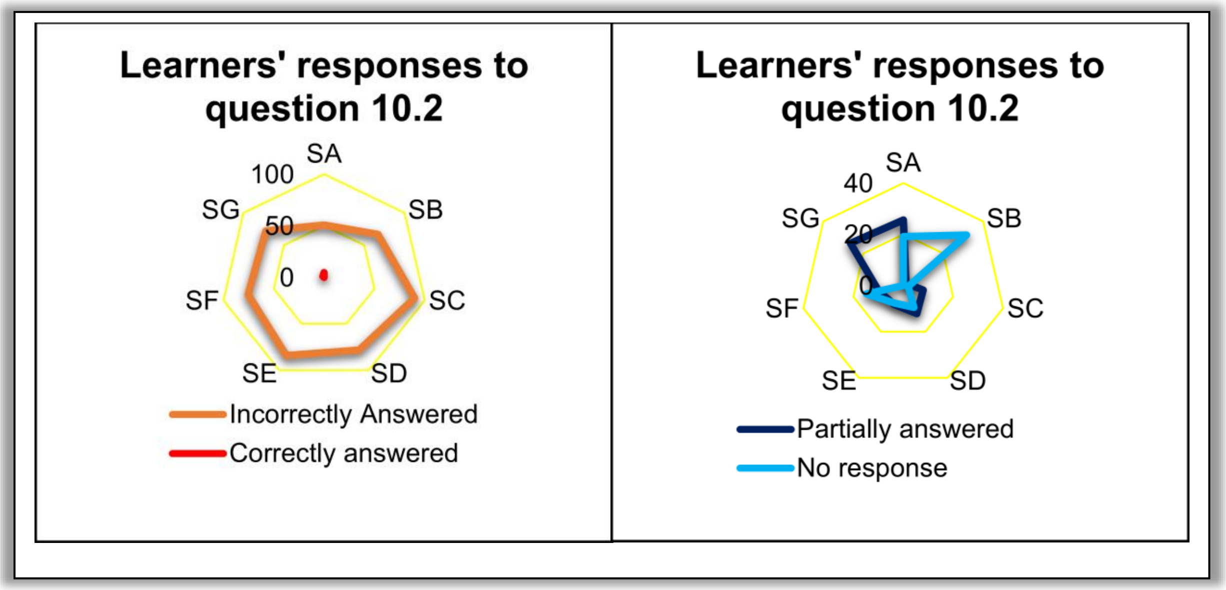
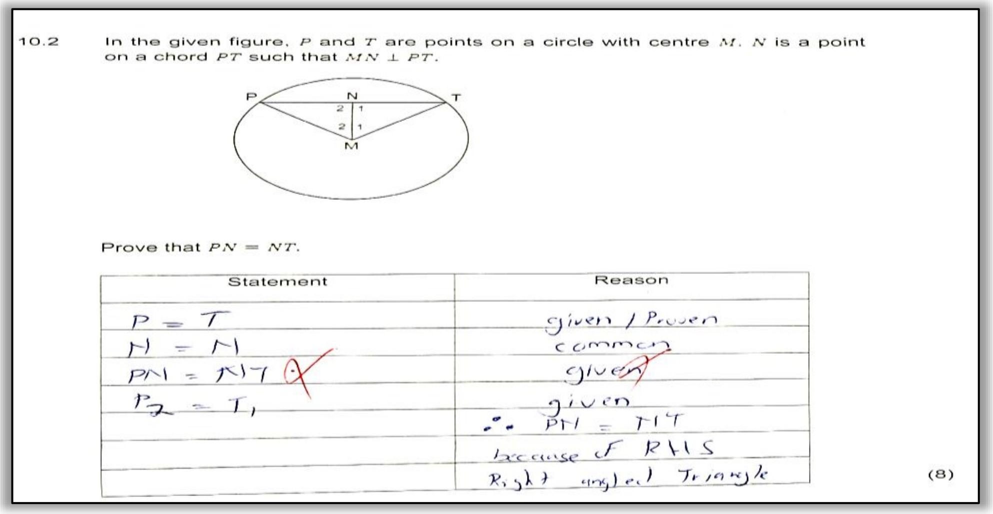


Figure 4.22: Trend in learners' responses to question 10.2

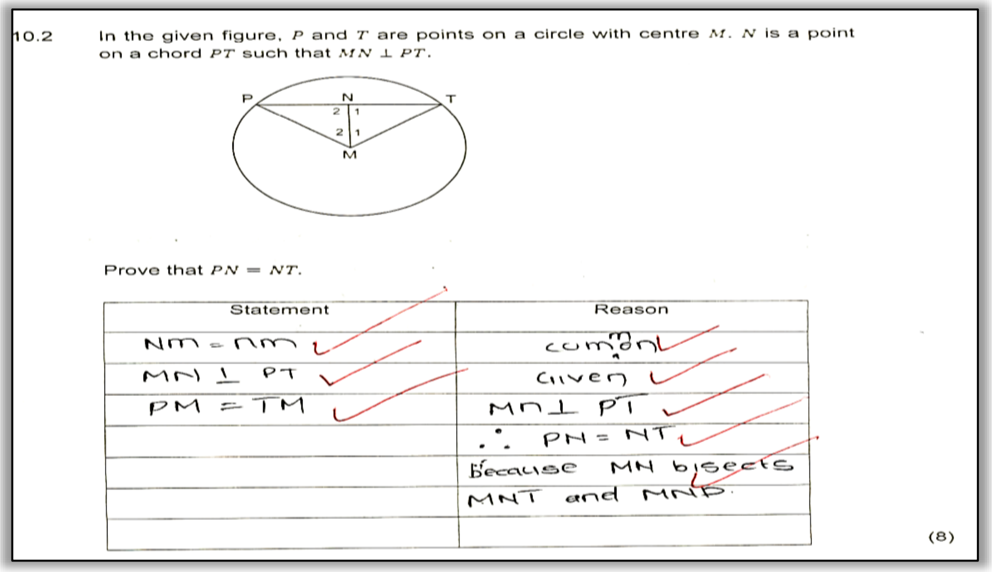
The vignettes in Figure 4.23 are learners' responses to question 10.2. This question was earlier coded SP-SC1-SC2-SC3-AR3 (see Table 4.21) which shows that this question involved routine simple procedure which allowed learners to prove an analogy.

Phase one, *flexibility*, the question did not specify the procedure to use to prove the analogy, to prove that '$PN = NT$' by showing that $\Delta PMN \equiv \Delta TMN$ by using two possible conditions for congruence. Learner A followed a procedure for proving congruence, and the same for Learner B. Phase two, *accuracy*: Learner A, the first step '$P = T$' is not a condition for congruence (P and T are points), '$PN = NT$'needs to be proved and at this stage cannot be pronounced to be equal. The last statement '$P_2 = T_1$'is invalid because the learner failed to mark such angles on the diagram. In contrast, Learner B, proves for congruency by showing a condition for congruency. However the learner did not specify the congruent triangles.

Phase three, *efficiency*: Learner A was not efficient in the procedure due to lack of accuracy and Learner B was efficient in the procedure for congruency due to the accuracy observed. Phase four, *appropriation*: Learner A was not appropriate due to the lack of both accuracy and appropriation. As such Learner A is one of 73.95% for learners' responses who were coded SMP10.2B those who were not proficient in question 10.2. Subsequently, Learner B was one of 0.86% for learners' responses coded SMP10.2D, those who were proficient in question 10.2.



Learner A



Learner B

Figure 4.23: Learners' responses to question 10.2

❖ LEARNERS' RESPONSES TO QUESTION 10.3.1

Figure 4.24 illustrates learners' responses to question 10.3.1. In the category correctly answered, the distribution is almost centred which shows that a very small number of learners correctly answered question 10.3.1. An interesting observation was that, even in school A, a small number of learners (4.19%) correctly answered question 10.3.1. The category partially answered shows a distribution that is an irregular heptagon, school A, (14.88% correctly answered), school B (2.8% correctly answered), school C (16% correctly answered), school D (9.5% correctly answered), school E (3.2% correctly answered), school F (6.75% correctly answered) and school G (15.25% correctly answered).

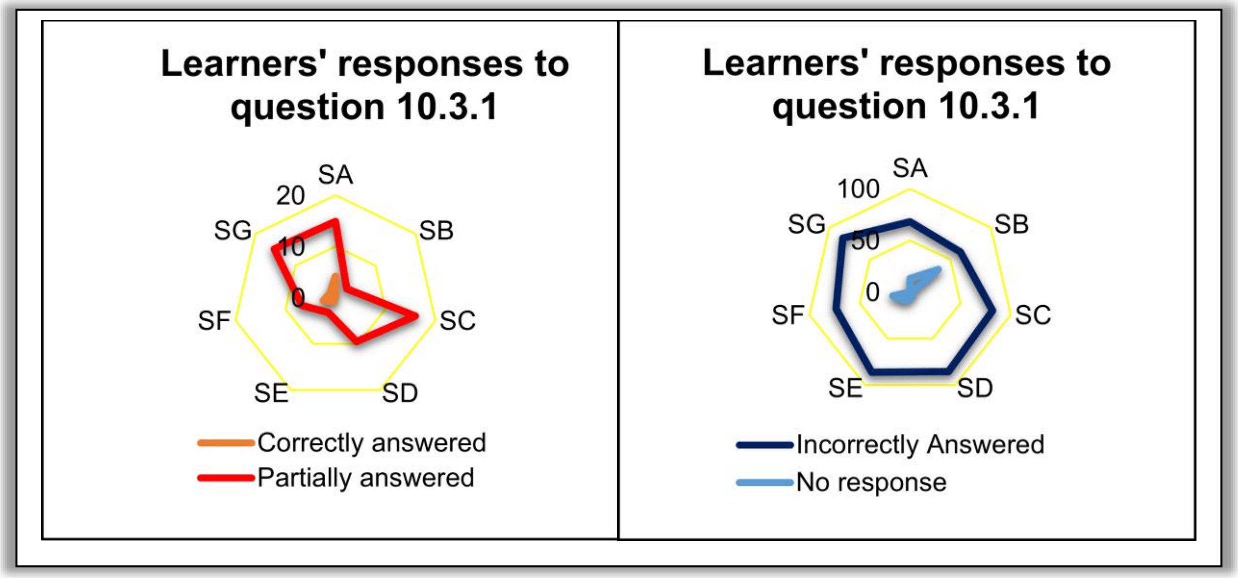
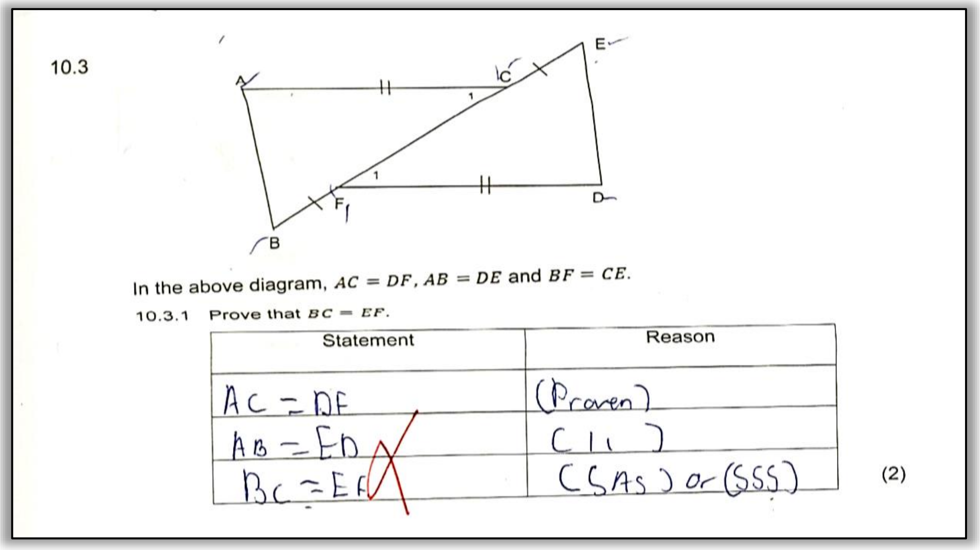


Figure 4.24: Trend in learners' responses to question 10.3.1

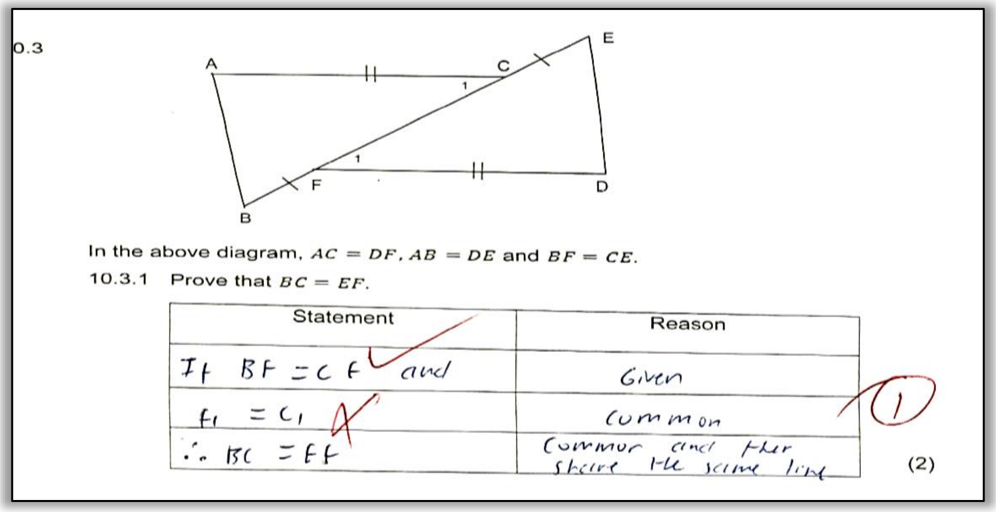
The vignettes in Figure 4.25 are learners' responses to question 10.3.1. This question was earlier coded SP-SC1-SC2-SC3-AR3 (see Table 4.21) which shows that this question required routine simple procedure which allowed learners to prove an analogy.

Phase one, *flexibility*: the question is not flexible because there only one procedure, using sides, that can be used to prove congruency and $BC \text{ and } EF$, the sides that need to be proved are in the triangles. Both Learner A and B used the procedure for congruency. Phase two: *accuracy*, Learner A was not accurate by giving invalid statements and reasons as follows: $\mathrm{AC} = \mathrm{DF}\ (\mathrm{Proven})$, $\mathrm{AB} = \mathrm{ED}\ (\mathrm{proven})$ and $\mathrm{BC} = \mathrm{EF}\ (\mathrm{SAS})\mathrm{or}\ (\mathrm{SSS})$. Learner B showed only one valid statement; $BF = CF\ (\mathrm{given})$, and the other statements were invalid; $\mathrm{F}_1 = \mathrm{C}_1\ (\mathrm{common})\ and\ BC =$

EF (common and they share the same line). Therefore, both learners were not accurate.



Learner A



Learner B

Figure 4.25: Learners' responses to question 10.3.1

Phase three, *efficiency*: both Learner A and B were not efficient due to failure to exhibit accuracy in the procedures. Phase four, *appropriation*: due to the failure by both learners to exhibit accuracy and efficiency in congruency, Learner A was one of 77.29% coded SMP10.3.1B who were not proficient in question 10.3.1. Subsequently in the coding, Learner B, due to one correct step of the procedure for congruency, was coded SMP10.3.1C, one of 9.77% for learners who partially answered question 10.3.1.

❖ LEARNERS' RESPONSES TO QUESTION 10.3.2

Figure 4.26 illustrates the trend in learners' responses to question 10.3.2. The category correctly answered is almost centred which shows very low or no learners who correctly answered question 10.3.2. It is worth noting that even school A had 15.81% of learners who correctly answered question 10.3.2. The category incorrectly answered shows a distribution that is an irregular heptagon with high number of learners who incorrectly answered question 10.3.2. The exception was in school A, (67.91% incorrectly answered) and school B (62.3% incorrectly answered). The category partially answered is irregular which shows a very low or no learners who partially answered question 10.3.2 in the schools. An exception is in School A, (38.61% correctly answered) and school G (27.78% correctly answered). Last, the category no response is also irregular showing small number of learners who did not respond to question 10.3.2. An exception was in school A, (15.35% no response) and school B (45.3% no response).

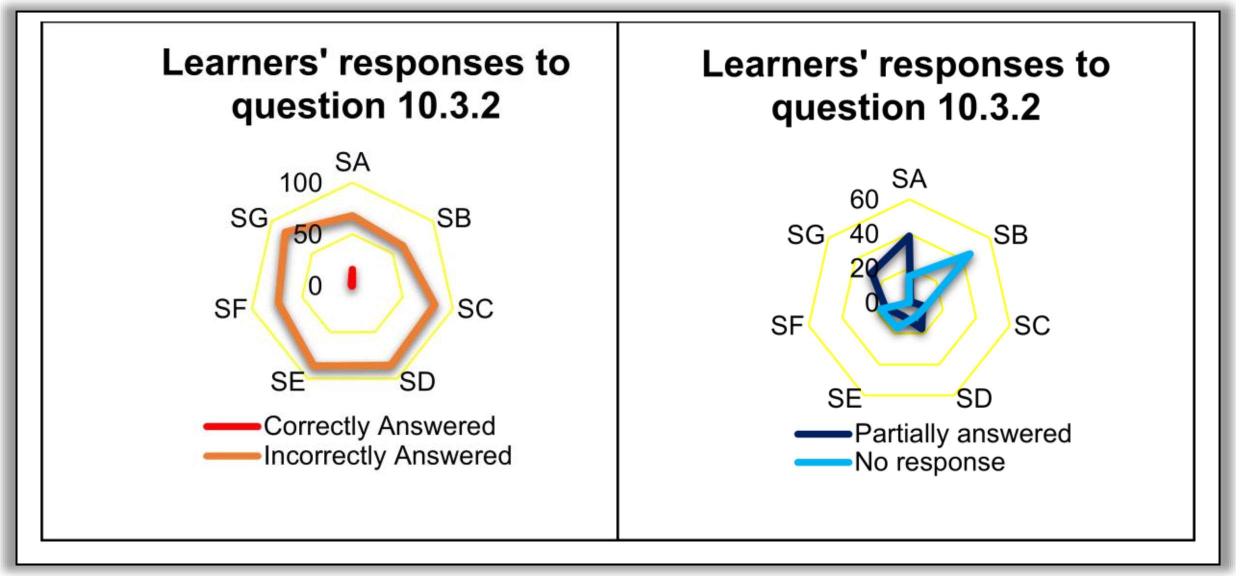
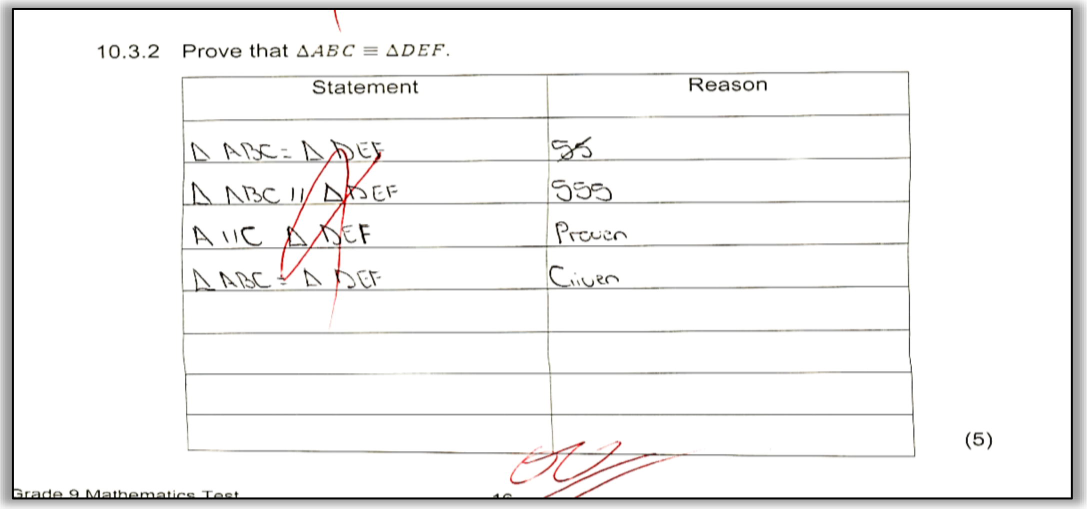
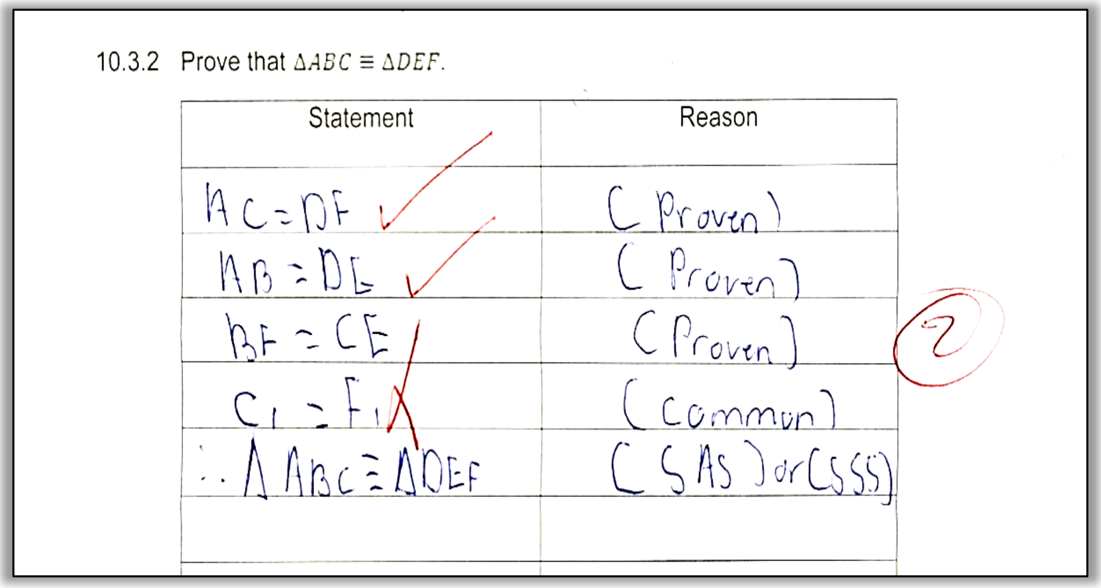


Figure 4.26: Trend in learners' responses to question 10.3.2

The vignettes in Figure 4.27 are learners' responses to question 10.3.2. This question was earlier coded SP-SC1-SC2-SC3-AR3 (see Table 4.21) which shows that this question was routine simple procedure which allowed learners to prove an analogy.



Learner A



Learner B

Figure 4.27: Learners' responses to question 10.3.2

Phase one, *flexibility*: the question is a follow-on from question 10.3.1 and the procedure that learners used was dependant on the previous responses. Hence the question is flexible concerning various procedures that learners exhibited in the previous question. Learner A and B followed a procedure for congruency. Phase two, *accuracy*: Learner A gave invalid statements and reasons which are not consistent with how the learner responded to the previous question and not accurate for a proof for congruency. Learner B correctly showed three statements for congruency, but two reasons were incorrect as only one '$BF = CE$' has been proved while the other two were not proven, but given in the original question. Hence Learner A and B did not exhibit accuracy in this question.

Phase three, *efficiency*: Learner A and B were not efficient in this question due to their inconsistencies in executing procedures that were earlier exhibited in the previous question. Phase four, *appropriation*: Learner A and B failed to execute accuracy and efficiency, as such they gave invalid statements that were justified with false reasons that showed their failure to apply rules of congruency. As such, Learner A was one of 64.84% for learners' responses coded SMP10.3.2B for being not proficient in question 10.3.2. Subsequent in the coding, Learner B was classified as SMP10.3.2C, one of 16.41% for learners who partially answered question 10.3.2.

- **LEARNER'S RESPONSES TO QUESTION 10.4.1**

Figure 4.28 illustrates the trend in learners' responses to question 10.4.1. The category correctly answered is centred showing a very low number of learners who correctly answered question 10.4.1 in all the schools. The category partially answered is an irregular heptagon showing a low number of learners who partially answered question 10.4.1. The exception was in school A, (24.65% partially answered), school C (20% partially answered) and school G (12.5% partially answered). The category incorrectly answered shows a distribution that is an irregular heptagon. This is an indication that there was a very high number of learners who incorrectly answered question 10.4.1 in most schools. An exception was in school A, (49.3% incorrectly answered) and school B (55.2% correctly answered). Last, the category no response is irregular with low numbers for learners who did not respond to question 10.4.1. An exception was in school A, (21.86% no response) and school B (43.9% no response).

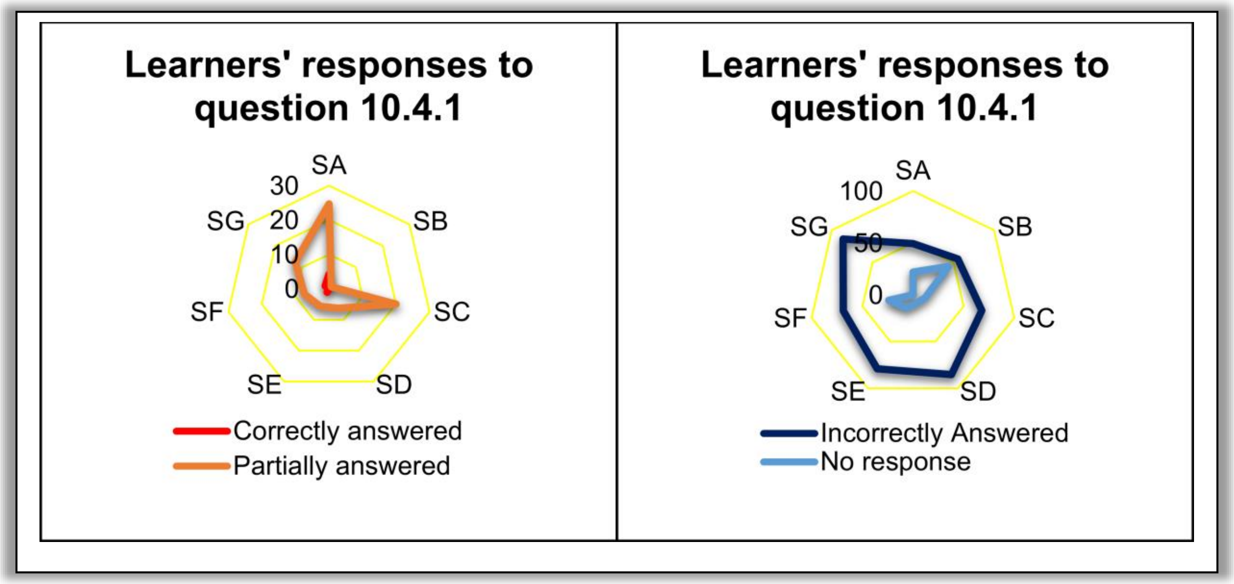
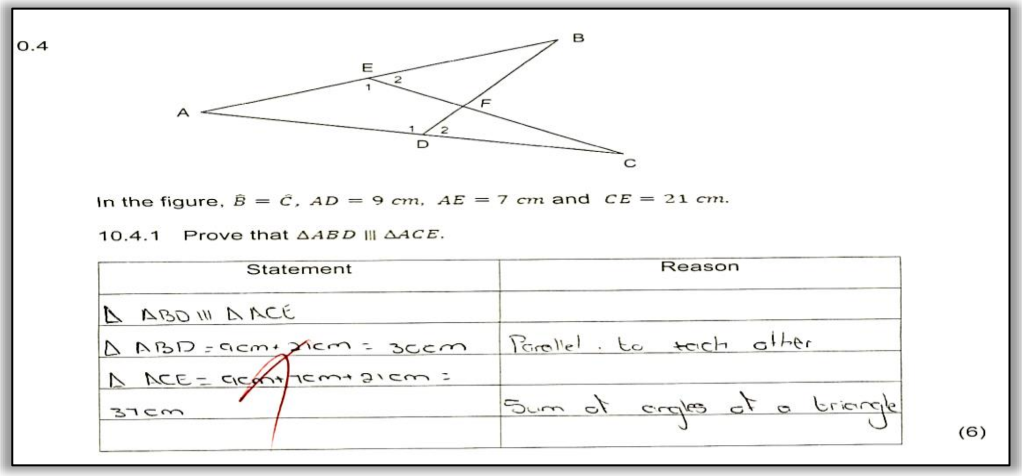


Figure 4.28: Trend in learners' responses to question 10.4.1

The vignette in Figure 4.29 are learners' responses to question 10.4.1. This question was earlier coded SP-SC1-SC2-SC3-AR3 (see Table 4.21) which shows that this question was routine simple procedure that allowed learners to prove an analogy.

Phase one, *flexibility*, the question allowed one condition for congruency, the use of only angles to prove the analogy, then the question was not flexible. Learner A used a wrong procedure, which is, the given lengths and is not a procedure to prove congruency. Learner B used a procedure for proving congruency.

Phase two, *accuracy*, due to the choice of a wrong procedure and that this question was not flexible, accuracy by Learner A cannot be considered. In contrast, Learner B, correctly showed two given conditions, '$\hat{B} = \hat{C}$ and $\hat{A} = \hat{A}$' in the question as conditions for congruency. However, the learner could not identify that '$\widehat{D_1} = \widehat{E_1}$' after showing that other corresponding angles of the two triangles are equal. Instead, the learner wrote an incorrect statement and reasons; '$E_1 = E_2 = D_1 = D_2, common.$'



0.4

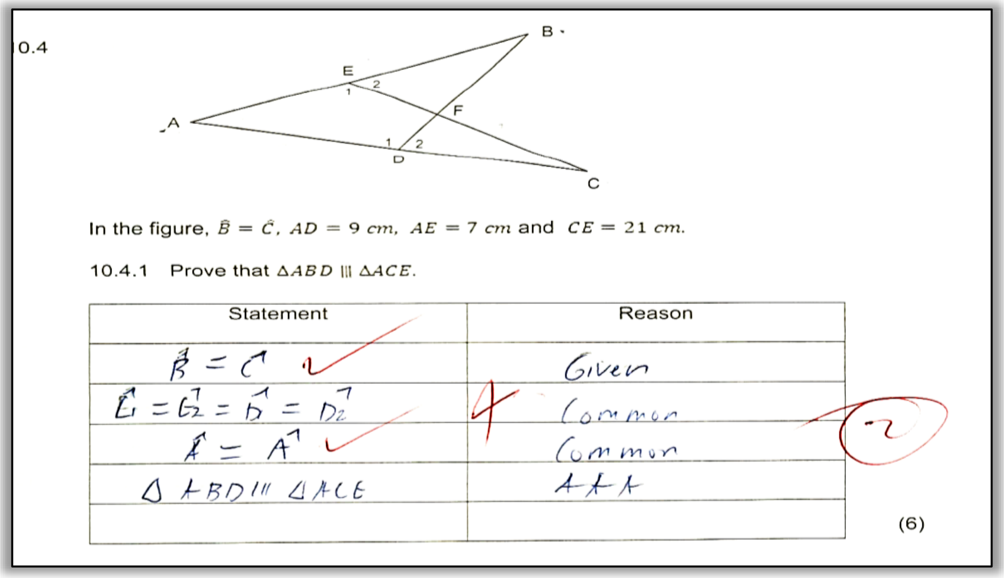
In the figure, $\hat{B} = \hat{C}$, $AD = 9\ cm$, $AE = 7\ cm$ and $CE = 21\ cm$.

10.4.1 Prove that $\Delta ABD\ |||\ \Delta ACE$.

Statement	Reason
Δ ABD ||| Δ ACE	
Δ ABD = 9cm + 21cm = 30cm	Parallel to each other
Δ ACE = 9cm + 7cm + 21cm =	
37cm	Sum of angles of a triangle

(6)

Learner A



0.4

In the figure, $\hat{B} = \hat{C}$, $AD = 9\ cm$, $AE = 7\ cm$ and $CE = 21\ cm$.

10.4.1 Prove that $\Delta ABD\ |||\ \Delta ACE$.

Statement	Reason
$\hat{B} = \hat{C}$	Given
$\hat{E}_1 = \hat{E}_2 = \hat{D}_1 = \hat{D}_2$	Common
$\hat{A} = \hat{A}$	Common
Δ ABD ||| Δ ACE	AAA

(6)

Learner B

Figure 4.29: Learners' responses to question 10.4.1.

Phase three, *efficiency*, Learner A was not efficient due to the choice of a wrong procedure in a question that was not flexible. Questions that are not flexible are essential to foster achievement of particular outcomes (Star, 2004). In contrast, Learner B was not fully efficient due to failure to exhibit the third condition for congruency. Phase four, *appropriation*, failure by Learner A to use the relevant procedure justified lack of appropriation. Learner B could not be consistent in the proof for congruency (not fully efficient), then this learner did not fully exhibit appropriation. Hence Learner A was coded SMP10.4.1B, one of (70.23%) learners who were not proficient in question 10.4.1. Subsequent in the coding, Learner B was coded SMP10.4.1C, one of 11.06% for learners who partially answered question 10.4.1.

❖ **LEVELS OF MATHEMATICAL PROFICIENCY IN QUESTION 10**

In Table 4.23, is a summary of the levels of mathematical proficiency to question 10. From the table it is evident that most learners (87.26%) were not proficient in geometry questions. A very small proportion of learners (11.16% and 1.58%) were moderate and proficient in geometry questions respectively. This is irrespective of earlier findings in the current study that ANA tested analogical reasoning in geometry, a weaker form of reasoning and proof (Amir-Mofidi et al., 2012).

Table 4.23: summary of learners' levels of mathematical proficiency to question 10

ANA 2012 Question 10 (Geometry)	**Codes for levels of mathematical proficiency**			
ANA questions	*Not Proficient*	*Moderately Proficient*	*Proficient*	*Total*
10.2 Proving that two sides are equal.	(178)SMP10.2A (912)SMP10.2B	(149)SMP10.2C	(11)SMP10.2D	1250
10.3.1 Proving that two lines are equal.	(174)SMP10.3.1A (962)SMP10.3.1B	(97)SMP10.2C	(17)SMP10.3.1D	1250
10.3.2 Proving that two triangles are similar.	(236)SMP10.3.2A (783)SMP10.3.2B	(194)SMP10.3.2C	(37)SMP10.3.2D	1250
10.4.1 Proving that two triangles are similar.	(250)SMP10.4.1A (868)SMP10.4.1B	(118)SMP10.4.1C	(14)SMP10.4.1D	1250
Totals	**4363**	**558**	**79**	**5000**
Percent	**87.26**	**11.16**	**1.58**	**100**

4.3.4 Synopsis: Levels of Mathematical Proficiency in 2014 ANA

Table 4.24 outlines mean deviations with direction for the levels of SMP to question 3, 6 and 10 of the 2014 ANA. These values were generated from the data in Table 4.23, Table 4.19 and Table 4.14 by calculating the mean deviation for each category of levels of SMP in learners' responses to ANA. Algebra and Algebraic Fractions are learners' responses to question 3, Number Patterns are learners' responses to question 6 and Geometry are learners' responses to question 10.

Table 4.24: mean deviations for levels of mathematical proficiency

	Levels of mathematical proficiency			
ANA questions	***Not Proficient***	***Moderate Proficient***	***Proficient***	***Mean***
Algebra and Algebraic Fractions	3703.67	-1798.33	-1909.33	0.00
Number Patterns	2088	-1224	-864	0.00
Geometry	2296.33	-1108.67	-1587.67	0.00

The mean deviations for levels of SMP (Figure 4.30) outlines the stronger levels in cases where graphs protrude upwards above zero, and weaker levels in cases where the graph protrude downward below zero. The strength and weaknesses were in terms of Algebra and Algebraic fractions, Number Patterns and Geometry in the 2014 ANA.

The stronger level of SMP was *not proficient* in Algebra and Algebraic Fractions, Number Patterns and Geometry. Such findings are consistent with those in a study by Dhlamini and Luneta (2016) who found that the level *not proficient* was common in learners' responses to Grade 12 mathematics final examinations. Algebra and Geometry are regarded as abstract content areas that are essential for university mathematics and natural sciences (Luneta, 2015b). Hence such findings illustrate that policymakers need to address these content areas to ensure quality in mathematics education in South Africa. Concerning the strength of the level *not proficient* as observed in ANA, an evaluative assessment must be a serious concern to the monitoring of quality of curriculum reform (Graven & Venkat, 2014).

The weaker levels of SMP were moderate proficient and proficient in terms of Algebra and Algebraic Fractions, Number Patterns and Geometry. Again these findings are consistent with those in a study by Dhlamini and Luneta (2016) that the levels moderately proficient and proficient were weaker in learners' responses to Grade 12 mathematics final examinations. They are also consistent with findings by Ally and Christensen (2013) who found absence of reasoning in elementary mathematics classrooms. The quality of mathematics education is a serious concern where the system produced less proficient learners in Algebra and Geometry which are foundations of abstract mathematics and conceptual knowledge (Luneta, 2015b).

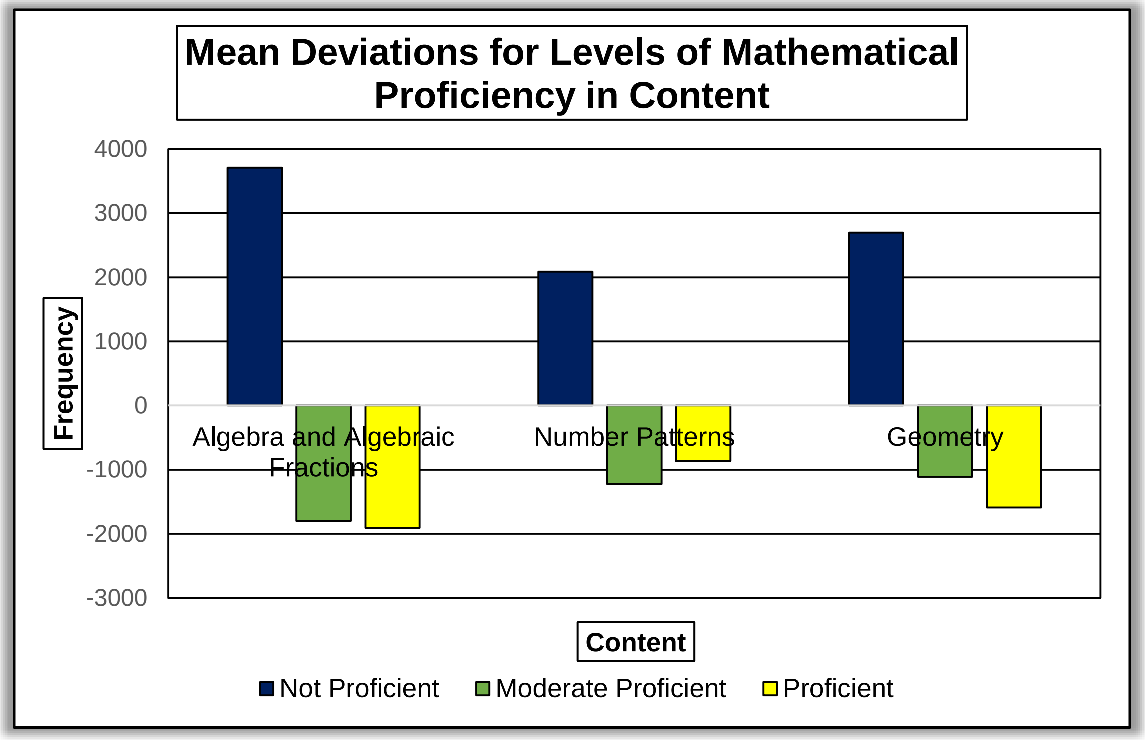


Figure 4.30: Mean deviations with direction for levels of mathematical proficiency

4.4 Results and Discussion for Alignment of ANA and TIMSS

According to Porter (2002), to calculate the Porter's alignment index, first there must be cells of matrices for content and cognitive levels for the documents that are to be aligned. In this study, the Porter (2002) alignment index was used to calculate the alignment index. A set of percentages of two tables was developed to form matrices of proportions that were used to make comparisons. In the current study, to calculate the Porter's alignment index, matrices were generated for cognitive levels for the ANA and the TIMSS.

4.4.1 Content and Cognitive levels in the 2012 ANA

To identify content and cognitive levels in the ANA 2012 mathematics Grade 9 test, the content in the 2012 ANA are classified into cognitive levels that each test item examined. There are four content areas: Numbers and Operations; Algebra and Functions; Geometry and Measurement; and Data handling and Probability. These were classified according to their cognitive demand into four cognitive levels: (1) Knowing Facts & Procedures; (2) Applying Concepts & Procedures; (3) Routine Problem Solving; and (4) Complex Problem Solving & Reasoning.

The matrix for the 2012 ANA mathematics test is shown in Table 4.25 below and is derived from the results of the question by question analysis of the 2012 ANA question paper. The totals in the matrix in Table 4.25 resulted from the number of hits of cognitive levels that the 2012 Grade 9 ANA mathematics test required. According to Porter (2002) this is a matrix of cognitive demands per content strand which is compared with another matrix and in this study it will be compared with the 2011 Grade 8 TIMSS, this is matrix (X_i).

Table 4.25: Matrix for 2012 ANA mathematics Grade 9 for topics and cognitive levels

Content	Cognitive levels				
	1	2	3	4	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	6	1	6	2	15
Algebra & functions	7	7	2	3	19
Geometry & measurement	5	2	6	8	21
Data Handling & probability	1	2	1	0	4
Sub-total	19	12	15	13	59

4.4.2 Content and Cognitive Levels in the 2013 Grade 9 ANA

To identify content and cognitive levels in the ANA 2013 mathematics Grade 9 test, the content in the 2013 ANA is classified into cognitive levels that each test item examined. There are four content areas: Numbers and Operations; Algebra and Functions; Geometry and Measurement; and Data handling and Probability. These were classified according to their cognitive demand to four cognitive levels: (1) Knowing Facts & Procedures; (2) Applying Concepts & Procedures; (3) Routine Problem Solving; and (4) Complex Problem Solving & Reasoning.

The matrix for the 2013 ANA mathematics test is shown in Table 4.26 below and is derived from the results of the question by question analysis of the 2013 ANA question paper. The totals in the matrix in Table 4.26 resulted from the number of hits of cognitive levels that the 2013 Grade 9 ANA mathematics test required. According to Porter (2002), this is a matrix of cognitive demands per content strand which is compared with another matrix and in this study it will be compared with the 2011 Grade 8 TIMSS, this is matrix (X_j).

Table 4.26: Matrix for 2013 ANA mathematics Grade 9 for topics and cognitive levels

Content	Cognitive levels				
	1	2	3	4	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	5	0	3	1	9
Algebra & functions	7	5	2	4	18
Geometry & measurement	5	7	4	6	22
Data Handling & probability	9	2	2	0	13
Sub-total	26	14	11	11	62

4.4.3 Content and Cognitive levels in the 2014 Grade 9 ANA

To identify content and cognitive levels in the ANA 2014 mathematics Grade 9 test, the content in the 2014 ANA is classified into cognitive levels that each test item examined. There are three content areas, Numbers and Operations, Algebra and Functions, and Geometry and Measurement. These were classified according to their cognitive demand to four cognitive levels: (1) Knowing Facts & Procedures; (2) Applying Concepts & Procedures; (3) Routine Problem Solving; and (4) Complex Problem Solving & Reasoning.

The matrix for the 2014 ANA mathematics test is shown in Table 4.27 below and is derived from the results of the question by question analysis of the 2014 ANA question paper. The totals in the matrix in Table 4.27 resulted from the number of hits of cognitive levels that the 2014 Grade 9 ANA mathematics test required. According to Porter (2002) this is a matrix of cognitive demands per content strand which is compared with another matrix, and in this study it will be compared with the 2011 Grade 8 TIMSS. This is called this matrix (X_p).

Table 4.27: Matrix for 2014 ANA mathematics Grade 9 for topics and cognitive levels

Content	Cognitive levels				
	1	2	3	4	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	5	2	3	2	12
Algebra & functions	9	4	7	5	25
Geometry & measurement	7	8	2	7	24
Data Handling & probability	0	0	0	0	0
Sub-total	21	14	12	14	61

4.4.4 Content and Cognitive levels in the 2011 TIMSS

To identify content and cognitive levels in the TIMSS 2011 mathematics Grade 8 test items, the content is classified in the 2011 TIMSS into cognitive levels that each test item examined. There are four content areas: Numbers and Operations; Algebra and Functions; Geometry and Measurement; and Data Handling and Probability. These were classified according to their cognitive demand to four cognitive levels: (1) Knowing Facts & Procedures; (2) Applying Concepts & Procedures; (3) Routine Problem Solving; and (4) Complex Problem Solving & Reasoning.

The matrix for the 2011 TIMSS mathematics test is shown in Table 4.28 below and is derived from the results of the question by question analysis of the 2011 TIMSS mathematics test items. The totals in the matrix in Table 4.28 resulted from the number of hits of cognitive levels that the 2011 TIMSS Grade 8 mathematics test demanded. According to Porter (2002), this is a matrix of cognitive demands per content strand which is compared with another matrix and in this study it is compared with matrices formed using the Grade 9 mathematics ANAs 2012 (X_i), 2013 (X_j) and 2014 (X_p) respectively. I shall call this matrix (Y_i).

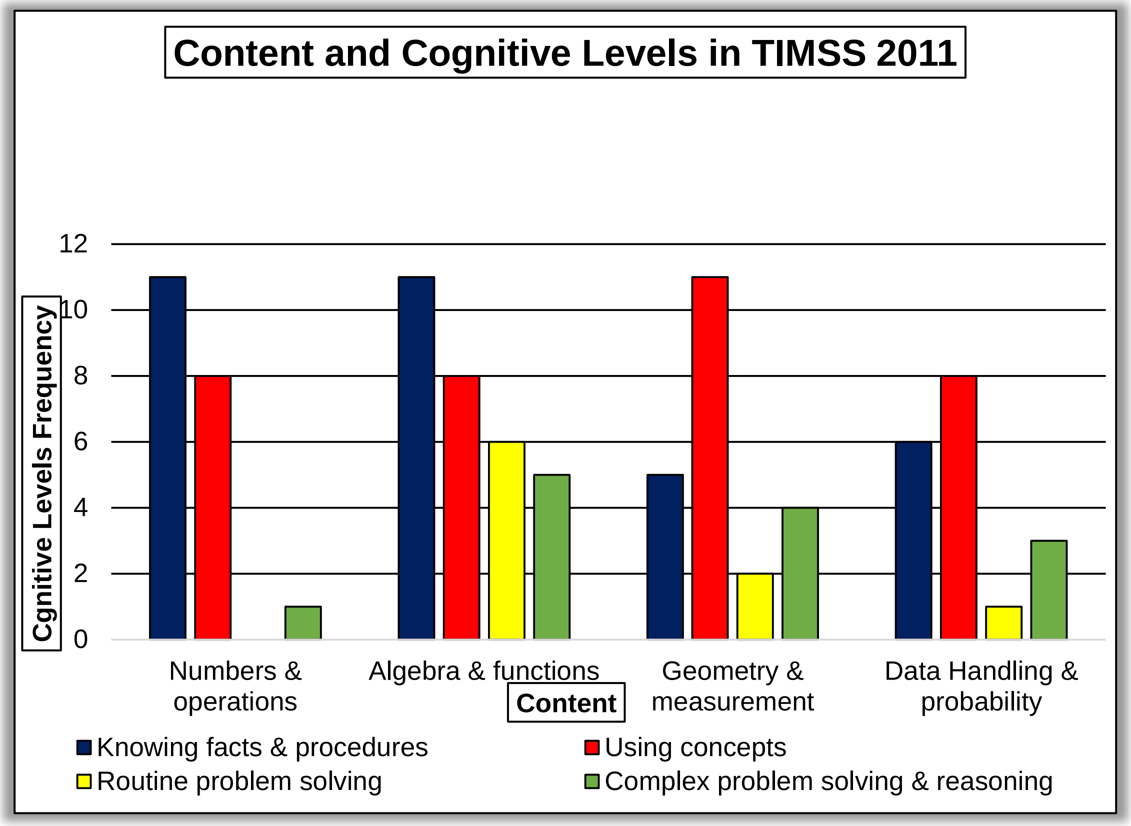
Table 4.28: Matrix for 2011 TIMSS mathematics Grade 8 for topics and cognitive levels

Content	**Cognitive levels**				
	1	2	3	**4**	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	11	8	0	1	20
Algebra & functions	11	8	6	5	30
Geometry & measurement	5	11	2	4	22
Data Handling & probability	6	8	1	3	18
Sub-total	33	35	9	13	90

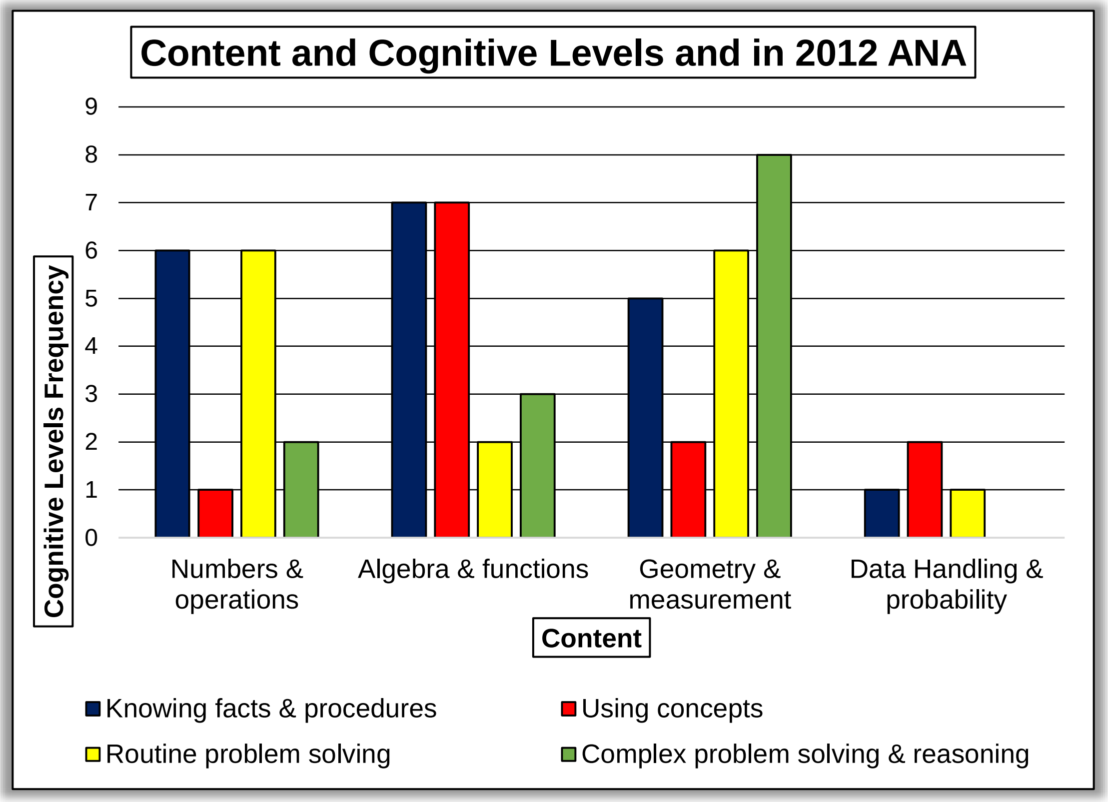
4.4.5 Comparing Content and Cognitive Levels in ANA and TIMSS

The data in Figure 4.31a-d is a synopsis of cognitive levels and content in the 2012, 2013, 2014 ANA and 2011 TIIMSS. In Numbers and Operations, Knowing Facts and procedures and routine questions and routine problems are mostly tested. This is an indication that ANA in this content area tested mostly lower order questions, a challenge gauging numeracy in mathematics education in South Africa (Graven & Venkat, 2014; Greenleess, 2011). Subsequently, the questions tested less questions on Using Concepts (conceptual understanding). As such the ANA did not focus on comprehending numbers and this allows learners to coherently solve various concepts, procedures and relations (Star & Stylianides, 2013). These findings confirms data in this study on ANA questions. This confirms the fact that in this content area there was a deficit of questions that test at higher order, a challenge to a systemic assessment such as ANA. For Algebra and Functions the same results were observed for knowing facts and complex problems. However, more questions were on using concepts (conceptual understanding) and less on routine problem solving. By contrast, the TIMSS have a different trend, as they test more on knowing facts and using concepts in most content areas. Subsequently, the 2011 TIMSS test less routine and complex problems. Consequently, the ANA in all three consecutive years are not aligned to TIMSS in terms of content and cognitive levels. This is worrisome and not regular in mathematics assessment that only Algebra and Functions must test more conceptual understanding (Schneider & Stern, 2010; Stein et al., 1996).

a)

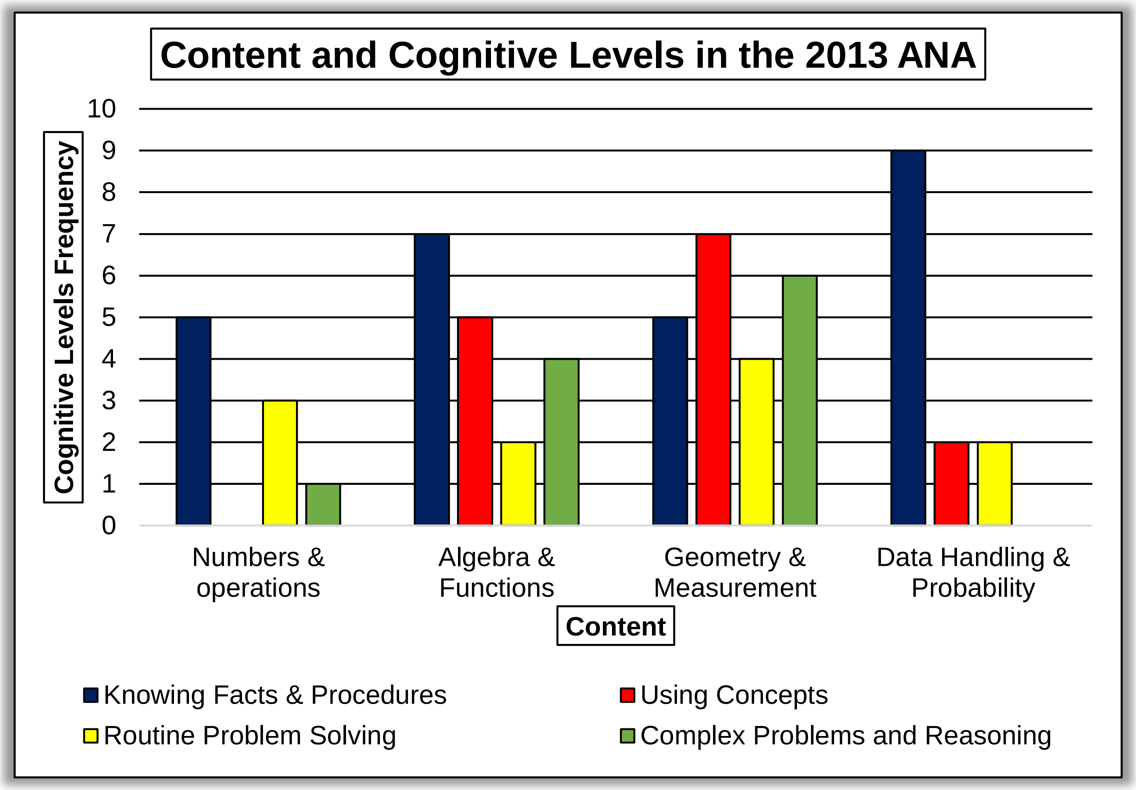


b)



Content and Cognitive Levels and in 2012 ANA
Cognitive Levels Frequency
9
8
7
6
5
4
3
2
1
0
Numbers & operations
Algebra & functions
Geometry & measurement
Data Handling & probability
Content
Knowing facts & procedures
Using concepts
Routine problem solving
Complex problem solving & reasoning

c)



Content and Cognitive Levels in the 2013 ANA
Cognitive Levels Frequency
10
9
8
7
6
5
4
3
2
1
0
Numbers & operations
Algebra & Functions
Geometry & Measurement
Data Handling & Probability
Content
Knowing Facts & Procedures
Using Concepts
Routine Problem Solving
Complex Problems and Reasoning

d)

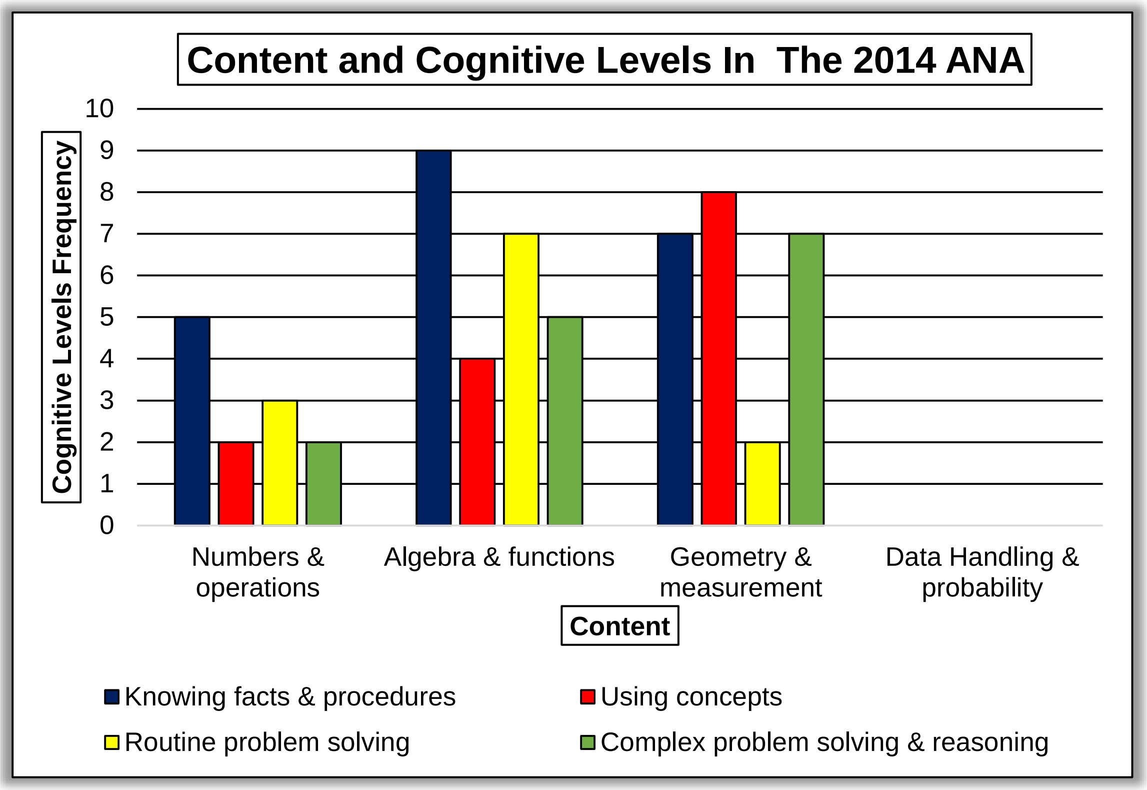


Figure 4.31 a-d: Cognitive Levels and Content in ANA and TIMSS

4.4.6 Alignment of the Grade 8 and the 2012 ANA Questions

In the matrix for the 2012 ANA mathematics, that is in Table 4.29, shows percentage ratios calculated by finding the quotient of the numbers in the cells with the total hits in all the cells as shown in Table 4.29 (X_i) below. Table 4.30 shows the percentage ratios, matrix Y_i derived from Table 4.28, and the 2011 TIMSS Grade 8 mathematics and are compared with those in cell X_i to calculate the alignment index between the 2012 ANA Grade 9 mathematics test and the 2011 Grade 8 TIMSS mathematics test. The formula for calculating the alignment index is used:

Alignment index = $1 - \frac{\sum |X-Y|}{2}$.

Table 4.29: (X_i) ANA 2012 mathematics Grade 9 ratios

Content	Cognitive levels				
	1	2	3	**4**	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	$\frac{6}{59} = 0.1$	$\frac{1}{59} = 0.02$	$\frac{6}{59} = 0.1$	$\frac{2}{59} = 0.034$	0.25
Algebra & functions	$\frac{7}{59} = 0.12$	$\frac{7}{59} = 0.12$	$\frac{2}{59} = 0.035$	$\frac{3}{59} = 0.05$	0.325
Geometry & measurement	$\frac{5}{59} = 0.085$	$\frac{2}{59} = 0.03$	$\frac{6}{59} = 0.1$	$\frac{8}{59} = 0.14$	0.355
Data Handling & probability	$\frac{1}{59} = 0.02$	$\frac{2}{59} = 0.03$	$\frac{1}{59} = 0.02$	$\frac{0}{59} = 0.0$	0.07
Sub-total	0.325	0.2	0.254	0.22	1.00

Table 4.30: (Y_i) TIMSS 2011 mathematics Grade 8 ratios

Content	Cognitive levels				
	1	2	3	**4**	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	$\frac{11}{90} = 0.12$	$\frac{8}{90} = 0.09$	$\frac{0}{90} = 0.0$	$\frac{1}{90} = 0.01$	0.22
Algebra & functions	$\frac{11}{90} = 0.12$	$\frac{8}{90} = 0.09$	$\frac{6}{90} = 0.07$	$\frac{5}{90} = 0.06$	0.34
Geometry & measurement	$\frac{5}{90} = 0.06$	$\frac{11}{90} = 0.12$	$\frac{2}{90} = 0.02$	$\frac{4}{90} = 0.04$	0.24
Data Handling & probability	$\frac{6}{90} = 0.07$	$\frac{8}{90} = 0.09$	$\frac{1}{90} = 0.01$	$\frac{3}{90} = 0.03$	0.2
Sub-total	0.37	0.39	0.1	0.14	1.00

When calculating the sum of the absolute difference between the 2012 ANA mathematics question paper and the TIMSS 2011 mathematics test items; the alignment index is 0.6566904991.

4.4.7 Alignment of the TIMSS Grade 8 and the 2013 ANA Questions

The matrix for the 2013 ANA mathematics in Table 4.31 is percentage ratios that were calculated by finding the quotient of the numbers in the cells with the total hits in all the cells as shown in Table 4.31 (X_j). The percentage ratios in Table 4.30 of the 2011 TIMSS Grade 8 mathematics (Y_i), are compared with those in cell X_j to calculate the alignment index between the 2013 ANA Grade 9 mathematics test and the 2011

Grade 8 TIMSS mathematics test. The formula for calculating the alignment index is used:

Alignment index = 1 - $\frac{\sum|X-Y|}{2}$

Table 4.31: (X_j) ANA 2013 mathematics Grade 9 ratios

Content	Cognitive levels				
	1	2	3	**4**	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	$\frac{5}{62} = 0.08$	$\frac{0}{62} = 0.0$	$\frac{3}{62} = 0.05$	$\frac{1}{62} = 0.02$	0.15
Algebra & functions	$\frac{7}{62} = 0.113$	$\frac{5}{62} = 0.08$	$\frac{2}{62} = 0.03$	$\frac{4}{62} = 0.06$	0.283
Geometry & measurement	$\frac{5}{62} = 0.08$	$\frac{7}{62} = 0.113$	$\frac{4}{62} = 0.06$	$\frac{6}{62} = 0.1$	0.353
Data Handling & probability	$\frac{9}{62} = 0.15$	$\frac{2}{62} = 0.032$	$\frac{2}{62} = 0.032$	$\frac{0}{62} = 0.0$	0.214
Sub-total	0.423	0.225	0.172	0.18	1.00

Calculate the sum of the absolute difference between the 2013 ANA mathematics question paper and the TIMSS 2011 matrix (Y_i), mathematics response items and divide it by two, and the value of the alignment index is 0.7281362007.

4.4.8 Alignment of the TIMSS Grade 8 and the 2014 ANA Questions

The matrix for the 2014 ANA mathematics in Table 4.32, percentage ratios is calculated by finding the quotient of the numbers in the cells with the total hits in all the cells as shown in Table 4.32 (X_p) below. Percentage ratios in Table 4.28 are for the 2011 TIMSS Grade 8 mathematics (Y_i) and are compared with those in cell X_p to calculate the alignment index between the 2014 ANA Grade 9 mathematics test and the 2011 Grade 8 TIMSS mathematics test. The formula for calculating the alignment index is used:

Alignment index = 1 - $\frac{\sum|X-Y|}{2}$

Table 4.32: (X_p) ANA 2014 mathematics Grade 9 ratios

Content	Cognitive Levels				
	1	2	3	**4**	
	Knowing facts & procedures	**Using concepts**	**Routine problem solving**	**Complex problem solving & reasoning**	**Sub-total**
Numbers & operations	$\frac{5}{61}=0.08$	$\frac{2}{61}=0.035$	$\frac{3}{61}=0.05$	$\frac{2}{61}=0.035$	0.2
Algebra & functions	$\frac{9}{61}=0.15$	$\frac{4}{61}=0.07$	$\frac{7}{61}=0.115$	$\frac{5}{61}=0.08$	0.415
Geometry & measurement	$\frac{7}{61}=0.11$	$\frac{8}{61}=0.13$	$\frac{2}{61}=0.035$	$\frac{7}{61}=0.11$	0.385
Data Handling & probability	$\frac{0}{61}=0.0$	$\frac{0}{61}=0.0$	$\frac{0}{61}=0.0$	$\frac{0}{61}=0.0$	0
Sub-total	0.34	0.235	0.2	0.225	1.0

When calculating the sum of the absolute difference between the 2014 ANA mathematics question paper and the TIMSS 2011 mathematics response items matrix (Y_i) and dividing it by two, this value of the alignment index is 0.6805922792.

4.4.9 The Value of the Alignment Index

The calculation of the alignment index between ANA and 2011 TIMSS revealed the following: the alignment index of the 2012 ANA Grade 9 mathematics test and the 2011 TIMSS is 0.657 (66%). The alignment index of the 2013 ANA Grade 9 mathematics test and the 2011 TIMSS is 0.728 (73%) and the alignment index of the 2014 ANA Grade 9 mathematics test and the 2011 TIMSS is 0.681 (68%) (Figure 4.31). Subsequently, the alignment index is interpreted as follows: 0 to 0.5 (no alignment to moderate), and 0.51 to 1.1 (moderate to perfect alignment) as shown by Porter (2002). Consequently, alignment between ANA and TIMSS is in the range ‘moderate to perfect). As such, there are content standards that are misaligned due to the alignment not being perfect and certain standards not tested (Polikoff et al., 2011). However, such alignment results between ANA and TIMSS (Figure 4.32) are contrary to findings by Ndlovu and Mji (2012) who found poor alignment between TIMSS and RNCS. It is suggested that ANA test developers must constantly monitor the alignment between ANA and other international assessment. In systemic assessments, the performance of a system is often compared with international standards to gauge how well it is doing (Volante & Cherubini, 2010). South African systemic assessment testing is not an exception.

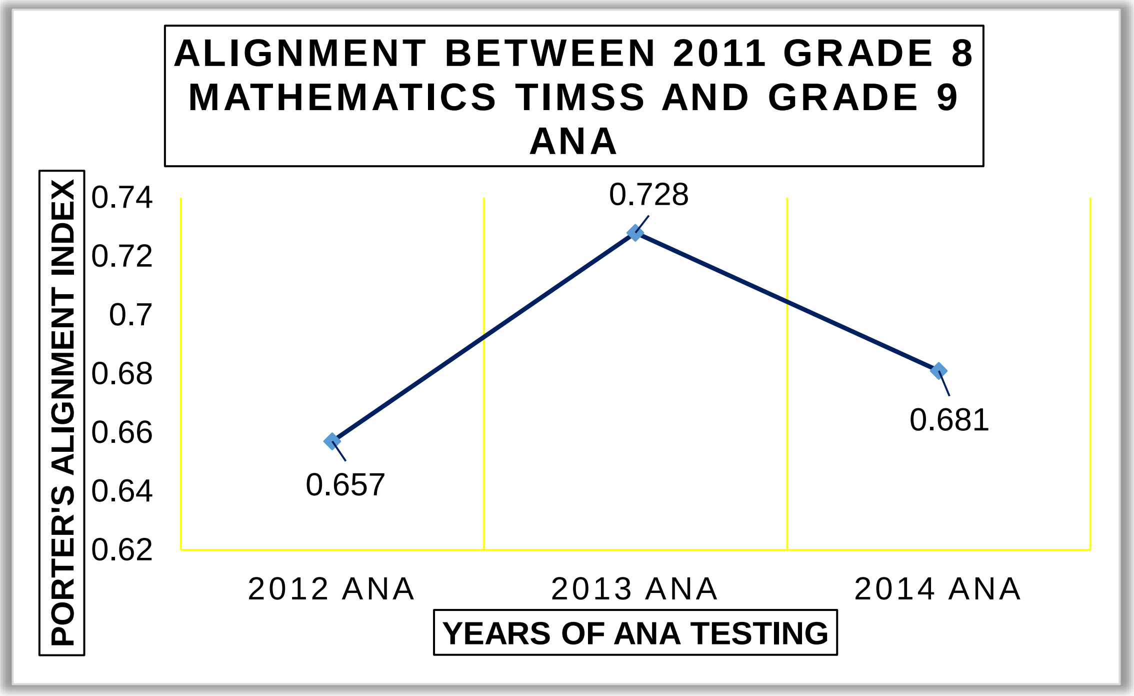


Figure 4.32: The Porters' alignment index for TIMSS and ANA

4.4.10 Discussion: Content and Cognitive Levels in ANA and TIMSS

Table 4.33 summarises the mean discrepancies for content in the 2012, 2013, 2014 ANA and TIMSS. The mean discrepancies for content were derived from the totals of content from data in Table 4.29, Table 4.31 and Table 4.32. The values show negative and positive discrepancies. This outlines the misalignment of content in the cells of ANA tests and TIMSS 2011. It is important for policymakers to align content in assessments to achieve standards of what is expected from learners (Martone & Sireci, 2009).

Table 4.33: Mean deviations for content with direction

ANA tests	*Numbers and operations*	*Algebra and functions*	*Geometry and measurement*	*Data handling and probability*	*Means*
2012 ANA	0	0.075	0.105	-0.18	0,0
2013 ANA	-0.1	0.033	0.103	-0.036	0.0
2014 ANA	-0.05	0.165	0.135	-0.25	0.0
TIMSS 2011	-0.03	0.09	-0.01	-0.05	0.0

Figure 4.33 illustrates the mean deviations for content with direction for the 2012, 2013, 2014 ANA tests and TIMSS 2011. From the graph it is evident that ANA and the 2011 TIMSS were weaker in terms of content in the instances where the bars protrude downwards below zero, and stronger in the instances where the bars protrude upwards above zero. The weakness and strength were in terms of content in the ANA and 2011 TIMSS.

Strengths and weaknesses were observed between ANA and the TIMSS in terms of content. First, ANA was stronger than the 2011 TIMSS in terms of geometry and measurement in the 2012, 2013 and 2014 tests. Second, strength of ANA was in the 2014 test in Algebra and Functions. In terms of weaknesses, ANA was weaker than the 2011 TIMSS in 2013 and 2014 in terms of Numbers and Operations and this was not the same in 2012. The 2012 paper was stronger than the 2011 TIMSS in terms of Numbers and Operations. Another weakness of ANA was in data handling and probability in the 2012 and 2014. However, in 2013 ANA was slightly stronger than 2011 TIMSS.

The strengths and weaknesses that have been observed between ANA and TIMSS confirms that there was misalignment between ANA and TIMSS in 2012, 2013 and 2014 in terms of content. Such misalignment was also observed by Ndlovu and Mji (2012) in the RNCS and TIMSS assessment standards. According to the observation made here, this seems to filter the ANA and TIMSS assessments. It is surprising because Ndlovu and Mji (2012) suggested that action must be taken to align the South African and TIMSS assessment standards. According to Porter (2002), making policymakers aware of content of instruction assists in monitoring the quality a curriculum that is being implemented. Additionally, Polikoff et al. (2011) also pointed out that content of instruction is key in judging whether the standard based curriculum reform is succeeding. The results shown in this analysis are worrisome when there is an obvious revelation that a curriculum monitoring tool such as ANA is misaligned with international assessments such as the 2011 TIMSS. Judging from these results, one can safely say that there is a need to align TIMSS and ANA in terms of content of instruction, which is key to the success of curriculum reform in South Africa (Graven & Venkat, 2014; Porter et al., 2011).

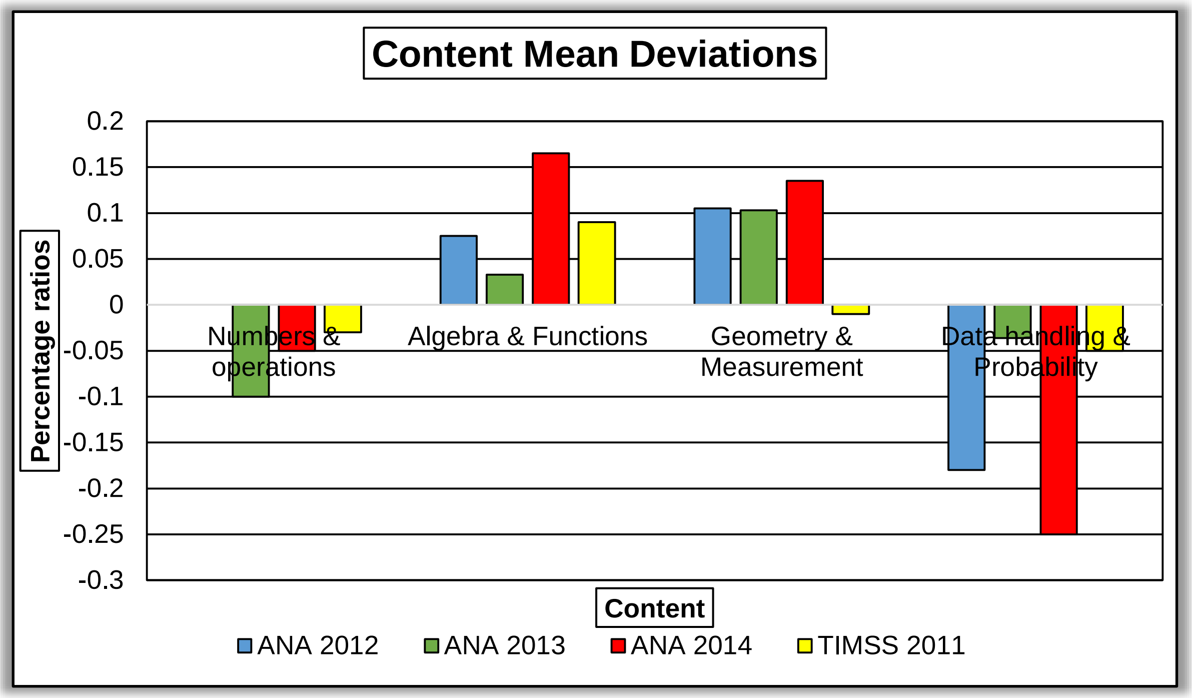


Figure 4.33: Mean discrepancies for content with direction

Table 4.34 summarises the mean deviations for cognitive in the 2012, 2013, 2014 ANA and the 2011 TIMSS. The mean discrepancies for cognitive levels were derived from the totals of cognitive levels from data in Table 4.29, Table 4.30, Table 4.31 and Table 4.32. The values show negative and positive discrepancies and this outlines the misalignment of cognitive levels in the cells of ANA tests and TIMSS 2011. Policymakers must assure that assessments are aligned and benchmarked internationally to inform monitoring of quality of an enacted curriculum positively (Ndlovu & Mji, 2012; Porter, 2002).

Table 4.34: Mean deviations for cognitive levels with direction

ANA tests	*Knowing facts*	*Using concepts*	*Routine Problem solving*	*Complex problem solving and reasoning*	*Means*
2012 ANA	0.075	-0.05	0.004	-0.03	0.0
2013 ANA	0.173	-0.025	-0.078	-0.07	0.0
2014 ANA	0.09	-0.015	-0.05	-0.025	0.0
TIMSS 2011	0.12	0.14	-0.15	-0.11	0.0

Figure 4.34 illustrates the mean deviations for cognitive levels in the 2012, 2013, 2014 ANA tests and TIMSS 2011. From the graph it is evident that ANA and the 2011 TIMSS were weaker in terms of cognitive levels in the instances where the bars protrude downwards below zero, and stronger in the instances where the bars protrude upwards above zero. The weakness and strength were in terms of mathematics cognitive levels in the ANA and 2011 TIMSS.

Strengths and weaknesses were observed between ANA and the TIMSS in terms of cognitive levels. First, ANA was stronger than the 2011 TIMSS in terms of knowing facts only in 2013 but weaker in 2012 and 2014. Secondly, ANA was stronger than TIMSS in 2012, 2013 and 2014 in terms of routine problem solving and complex problems and reasoning. Thirdly, ANA was weaker than TIMSS in terms of using concepts in 2012, 2013 and 2014.

The strengths and weaknesses that have been observed between ANA and TIMSS confirm that there was misalignment between ANA and TIMSS in 2012, 2013 and 2014 in terms of cognitive levels. In justifying the use of cognitive levels, Fulmer (2011) pointed that, for learners to master standards, there is almost complete reliance on the adequate testing of the standards. Hence, from the results of the current study, it is safe to say that the adequacy of the tested cognitive levels is irregular judging from the content and cognitive levels discrepancies. A more coherent approach is needed to create high quality ANA which when benchmarked internationally, is aligned to those standards (Polikoff et al., 2011).

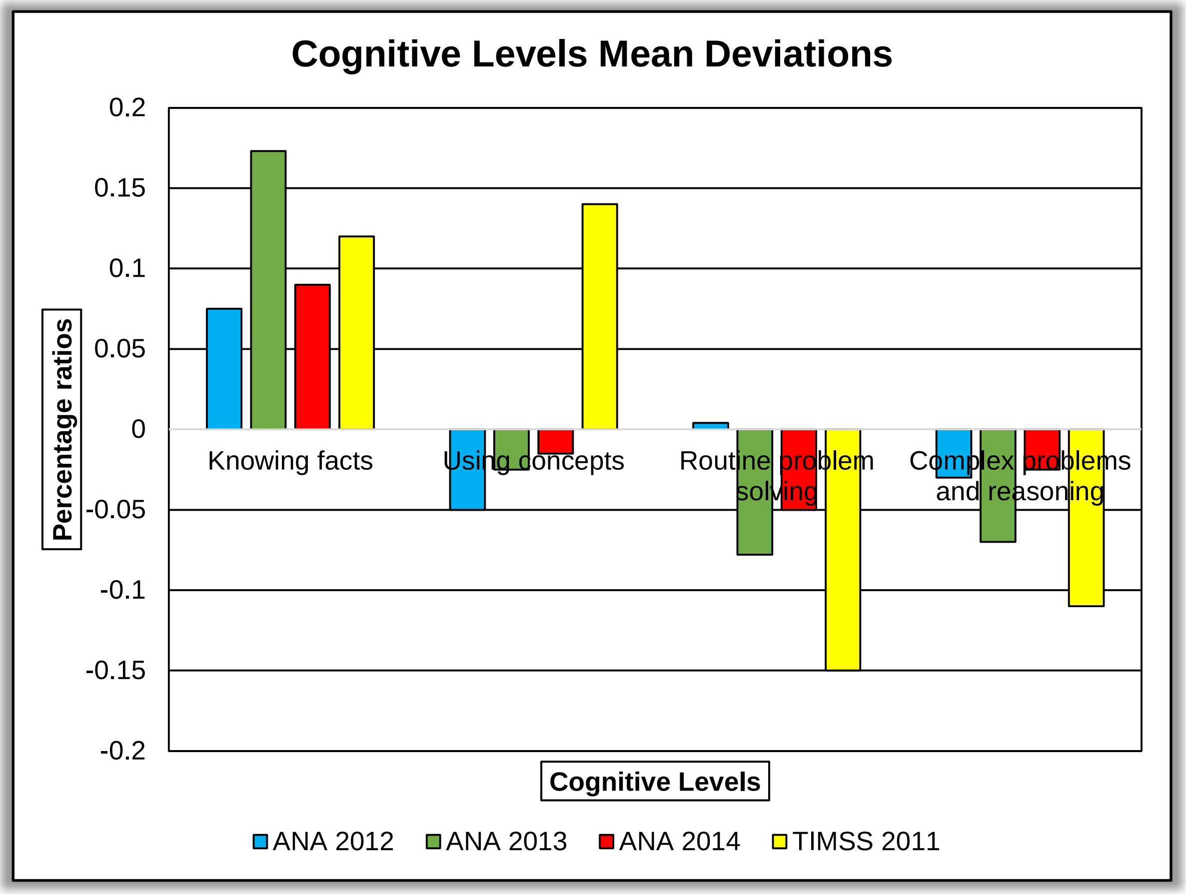


Figure 4.34: Mean deviations for cognitive levels with direction

4.4.11 Synopsis: Alignment ANA and TIMSS

To calculate alignment between ANA and TIMSS, two methods were used, the Porter's alignment index and the mean deviations. The Porters alignment index has revealed that ANA and TIMSS were in the range moderate to perfect. Such findings are not discursive in assessing how content and cognitive levels relate within the documents that are aligned (TIMSS and ANA). By contrast, the descriptive data explores the relationship between the content and the cognitive levels by explicitly showing those favoured in ANA and TIMSS. This disputes the data in the Porter's alignment index by showing misalignment. The use of the Porter's alignment index gives the impression that the alignment is good, whilst the mean deviations assist in showing misalignment. Thus, the current researcher can safely infer that ANA and TIMSS were misaligned irrespective of an alignment index that is moderate to perfect.

4.5 Conclusion

This chapter has presented and interpreted the results. Key findings have been captured that relate to the cognitive levels of mathematics tested by ANA, the levels of SMP exhibited by learners in response to ANA and the alignment of ANA and TIMSS. The next chapter presents the concluding statements and recommendations.

5. CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Introduction

This study was mainly concerned with exploring the effectiveness of Annual National Assessment in monitoring the mathematics education standard in South Africa through assessing the SMP tested by ANA and exhibited by Grade 9 learners in South Africa. This was premised on the notion that in assessments, when learners are tested, they exhibit only the mathematical knowledge, skills and dispositions that test items examine. The emphasis placed on the effectiveness of ANA was triggered by low achievement, especially in Grade 9, by South African learners in the three consecutive ANA tests in 2012, 2013 and 2014. Although the focus was not on achievement, this researcher opted to assess the effectiveness of ANA as an evaluative assessment for curricula in use. Such assessment was done using three dimensions. First, through the use of SMP, cognitive complexity of ANA was assessed. This notion responded to the first research question of the study. Secondly, the assessment of the levels of SMP exhibited by learners in their responses to ANA gave an indication of mathematical knowledge, skills and dispositions that learners exhibited. This assessment answered the second research question. Finally, the relation to content and cognitive levels in ANA tests and TIMSS, the results revealed the content messages that the assessment instruments related when benchmarked internationally. This answered the third research question. The discussion that follows in the next three sections captures the study's responses to the research questions and the key findings of the study.

5.2 Research Design and Method

The study used mixed methods in the context of the exploratory sequential design. The use of various methods assisted this study to confirm the inferences made in relation to the research problem. The research design was divided into three parts as follows: (1) Part one Phase one and Part one Phase two; (2) Part two Phase one and

Part two Phase; and (3) Part three Phase one and Part three Phase two as shown in Table 3.1. For part one phase one, content areas were explored to assess SMP that were tested by ANA. Codes were used to document the categories of SMP that emerged from the analysis. These codes were matched with mathematics cognitive levels which were informed by the NAEP Taxonomy to view the complexity of ANA. Means were used to generate mean deviations and charts that showed discrepancies in the various SMP.

Part two Phase one, assessed learners' responses exhibited in response to ANA 2014 questions. Learners' responses were captured in tables using four categories: correctly answered, partially answered, incorrectly answered, and no response. Subsequently, vignettes were also coded and used to explain proficiency levels of learners' responses to questions on Algebra, Number Patterns and Geometry. In part two phase one, the data in tables of categories was used to generate charts of radars that showed the trend in learners' responses. Additionally, the means were used to generate descriptive statistics, discrepancies in the levels of SMP.

For part three phase one, document analysis was used to capture content and cognitive levels in ANA and TIMSS to generate matrices with ratios. In part three phase two, the Porter's alignment index was calculated, together with descriptive statistics, and charts for discrepancies in content and cognitive levels to explain misalignment in cells of ANA and TIMSS.

5.3 Summary of Findings and the Research Findings

This section summarises the findings of the study by reflecting on how the study responded to the research questions. This assist the researcher to ascertain how this study addressed the research problem and subsequently achieved its purpose.

5.3.1 Research Question one

The first research question was: "*How are the cognitive levels of mathematics tested by ANA reflective of SMP?*" First, the results revealed that ANA testing was biased towards procedural fluency, while very few questions were aligned to conceptual understanding. Next, in strategic competence that was tested, notably, there was the absence of one element of strategic competence, and that was problem solving. Thirdly, adaptive reasoning that was only tested in Geometry was in the form of analogical proofs and no other forms of proof such as deductive and inductive proofs were tested. Fourthly, there was inconsistency in the content tested by ANA in the three consecutive years, such as the absence of Data Handling and Probability in the 2014 ANA. Fifth, ANA tests were biased towards low complexity as shown in the descriptive statistics. These results pointed towards the assumptions made earlier in this study, *'what you test is what you get',* which are further explored in the next research question.

5.3.2 Research Question two

The second research question was: "*What levels of mathematical proficiency do learners exhibit in response to the ANA tests?*" First, learners exhibited knowledge of lower Grades when computing procedures in Algebra. Secondly, the marking of ANA focused on correct answers and disregarded thinking processes that learners exhibited, in which some were procedurally correct. Thirdly, learners incorrectly used algorithms for other concepts to execute some computations. Fourthly, a majority of learners were not proficient to the sampled ANA questions, despite there being a majority of low complexity questions.

5.3.3 Research Question three

The third research question was: "*How do the content and cognitive levels tested by ANA compare with TIMSS?*" First, the Porter's alignment index between Grade 9 ANA mathematics in three consecutive years and Grade 8 TIMSS mathematics response items was in the range moderate to perfect. Secondly, ANA and TIMSS were misaligned, meaning that there were discrepancies in the content standards that the two assessments tested. Thirdly, in terms of content, ANA testing was biased towards Algebra and Geometry at the expense of Numbers and Operations, and Data Handling and Probability. Fourthly, in terms of content, TIMSS testing was biased towards Algebra at the expense of Numbers and Operations, Geometry, and Data Handling and Probability. Fifthly, in terms of cognitive levels, ANA was biased towards Knowing Facts at the expense of Using Concepts, Routine Problem Solving, and Complex Problem Solving and Reasoning. Finally, in terms of cognitive levels, TIMSS was biased towards Knowing Facts and Using Facts at the expense of Routine Problem Solving, and Complex Problem Solving and Reasoning.

5.4 Synthesis of Findings

This section synthesises the research questions with reference to the reviewed literature. The findings have been revealed in relation to the research questions and now I reflect on what they entail and why they appear the way they are. First the results indicate that ANA testing was biased towards procedural fluency with a very few questions aligned to conceptual understanding. The CAPS document stresses the need for conceptual understanding, as such, when ANA was configured, this aspect was not considered. Second, the results indicated that ANA testing was biased towards Algebra at the expense of Numbers and operations. This indicates that during the configuration of ANA, the test responded positively to the mathematics curriculum which has a lot of content in Algebra (35%) and less content in terms of Numbers, Operations and Relationships (15%) as a policy requirement (AMESA, 2012). Thirdly, the findings pointed that ANA seemed to favour Geometry testing and not TIMSS. These findings were reflective of the curriculum requirements in South Africa (30%) and the TIMSS assessment framework focus for mathematics (13%). Fourthly, ANA focused on the complexity levels of lower nature with less recognition of problem solving and higher order thinking skills. The use of SMP in this study was rigorous and it shows that the higher levels are minimal. This may emanate from the background knowledge and understanding of quality education by curriculum designers. This suggests that the focus area was at a policy statement level and not at implementation level. This implies that the classroom teacher has no guidance on what to act upon during the period since there was no clear actions for the implementers to follow. It is no wonder many teacher emphasise memorisation of content to show that there is improvement in learners' content understanding. This ultimately has compromised critical thinking at such an early age. To address this shortfall there should be a section in the curriculum explaining the pros and cons of critical thinking at various levels of education. Teachers can visit other countries where quality education has been identified especially at grades 1-9 for a term or two. Also, some pilot schools could use expatriates from the benchmarked countries to in-house educate teachers in the country for a specified time on how to manage quality education at the grass root level through teaching and assessment. The implantation of such drastic measures will in no doubt enhance quality education through assessment at various levels of achievements. Fifth, the results show that the majority of learners were not proficient during ANA, despite the low level type of questions used in test items in the middle complexity sections. Learners' responses to the 2014 ANA, from the sampled questions, were fragmented when analysed against the SMP. Additionally, the rigor in ANA was low as compared to the SMP and learners were challenged with such low level questions. As such, these challenges are attributed to the quality of assessment that learners received during teaching and learning. Bensilal (2017) highlighted discrepancies between Grade 9 classroom assessment and ANA testing. This was attributed to the fact that the majority teachers in South Africa lack skills of preparing quality classroom assessment. Finally, there is a misalignment between ANA and the learners' mathematics curriculum. This is evident in terms of content and cognitive levels between ANA and TMSS. The benchmark between ANA and TIMSS assist in gauging the quality of an education system. Such discrepancies in content and cognitive levels have high possibility of filtering into the mathematics classrooms. For an example, TIMSS has reasoning as a cognitive level, which is not explicit in the South African mathematics curriculum. This is an indication that ANA which is configured using the policy document is limited on reasoning. Hence, contents of the mathematics curriculum need to be revisited in terms of cognitive levels and mathematics content in order to improve quality of ANA.

In addition, the contrasting purposes of ANA and TIMSS attributed to this discrepancy. ANA focuses on the performance of a single system whilst TIMSS focuses on global education systems. The curricula requirements of these multiple systems in TIMSS differ drastically in content and cognitive levels. Hence the TIMSS assessment framework originates from a benchmark of all participating countries and this justifies its discrepancy to the curriculum assessment guidelines in South Africa.

5.5 Conclusions

In conclusion, I reflect on, firstly what prompted the researcher to embark on this study and secondly, whether the purpose of the study has been achieved. The purpose of this study was:
To explore the effectiveness of ANA in monitoring the standard of mathematics education and to assess the mathematical proficiencies tested and exhibited by Grade 9 learners in South Africa.

The study addressed its purpose in three ways first, using the SMP to determine mathematics cognitive levels tested by ANA. The results on the themes show that ANA favoured low cognitive skills. Additionally, there was total lack of other elements of SMP in strategic competence and adaptive reasoning. Three levels of SMP were tested by ANA, which were: low complexity, moderate complexity and high complexity. The results suggest that the ANA was skewed towards low complexity. This is a serious challenge as the literature points out that assessment must be neither too easy nor too difficult. A conclusion may be reached only towards that ANA posed reliability and validity concerns. This was as a result of the ANA tests being low complexity because of not paying attention to higher order thinking skills.

Secondly, the purpose of the study was addressed by assessing the levels of SMP that learners exhibited in their responses to ANA. Three levels of SMP were exhibited by learners in their responses to ANA, and these were: not proficient, moderate proficient and proficient. The results show that the levels were skewed too much towards not proficient. If I may generalise for the chosen sample could infer that the standard of mathematics in South Africa is a serious concern when considering that a bulk of learners in the chosen sample were not proficient when examined with ANA that was mainly low complexity.

Last, the purpose of the study was addressed by benchmarking the content and cognitive levels with international tests, TIMSS to calculate the Porter’s alignment index. The study found that the Porter’s alignment index was between moderate and perfect in 2012 (65.7%), 2013 (72.8%) and 2014 (68.1%). The mean and cognitive levels discrepancies showed that there was misalignment in the cells. Algebra and Numbers and Operations were favoured by ANA and TIMSS, while Geometry was favoured by ANA and not TIMSS and Data handling and Probability was not favoured by ANA and TIMSS. For cognitive levels, there was an indication from the data that ANA and TIMSS were skewed towards knowing facts, a low cognitive demand at the expense of complex problem solving and reasoning, which are high cognitive demands. For Grade 9 mathematics it can be concluded that ANA there was a serious challenge in using ANA as an instrument to monitor the standard of mathematics education in South Africa due to the misalignment that was observed when ANA was benchmarked internationally. The key findings that have been outlined suggest that the purpose of the study was achieved and the research problem was addressed only in the issues on which this study focused.

5.6 Recommendations

First, it is recommended that the examiners of the ANA tests must consider the complexity of the tests using assessment protocols such as the SAGM for mathematics. There must be consistency in consecutive tests to maintain the reliability of ANA. The test must be neither too easy nor too difficult, which implies that in terms of low, moderate and high complexity, the test must be normally skewed. Kilpatrick et al. (2001) pointed out that reasoning is the glue that holds everything together. Hence, ANA must include other forms of reasoning such as deductive reasoning and inductive reasoning, which will strengthen higher order problem solving abilities.

Secondly, it is recommended that mathematical knowledge, skills and dispositions exhibited by learners in response to ANA must be verified using theory such as the SMP which advocates coherence (Groves, 2012). This will provide policymakers with relevant information to use in redressing situations such as those where learners are not proficient or moderate proficient. Where learners are found to be proficient, policymakers must gather information that contributed to that success to share good practice with struggling schools. Once policymakers have gathered information, they must revisit the purpose of ANA. The literature identified the purpose of ANA as quality, equity and provision. There is a practice of not reporting on all these issues. Before embarking on the next ANA, policymakers must ascertain that they gather information on the following: (1) teaching and learning that is aligned to the implementation of curriculum; (2) mathematics knowledge, skills and dispositions that learners exhibit as a result of engaging with that type of ANA; (3) how the system is responding to issues of gender, socio-economic, ethnic groups and school governance; and (4) how the education system is responding to curriculum reform, system restructuring and factors of achievement. Once this information has been gathered, policymakers must design redress strategies that are not only informed by outcomes but also widely renowned mathematical practice such as SMP.

Thirdly, it is recommended that alignment of ANA must be done frequently in terms of content standards and cognitive levels to inform teaching and learning (Porter, 2002). This must begin by reviewing the cognitive levels in the SAGM to include higher order problem solving skills and reasoning such as adaptive reasoning. Frequent alignment of ANA will assist in testing relevant content of instruction and well-balanced cognitive levels. The use of mixed methodologies in calculating alignment will help the overreliance on the alignment index, which may be misleading when not supplemented with other quantitative descriptive statistics such as charts for mean deviations for content and cognitive levels and qualitative document analysis.

5.7 Experiences of Engaging With this Study

Methodologically, SMP have not been compatible for document analysis as observed in previous studies, such as seminal work by Graven and Stott (2012). Studies (Dhlamini & Luneta, 2016; Rittle-Johnson & Schneider, 2015; Star, 2004) often use two, or four SMP, which is contrary to claims by Kilpatrick et al. (2001) that the five strands are interconnected, interwoven and inseparable. An exception was a study by Groves (2012) who used all five strands. However, methodologically that study had a different focus from this study. Its focus was on classroom practice that uses SMP. Contrary, this study has used all five SMP methodologically in assessments, focusing on SMP tested by ANA and exhibited by learners in their responses. Further, the use of document analysis in this study, adapted SMP and the use of descriptive statistics, such as radar, which allowed me to elucidate explicitly SMP exhibited by learners in various schools

The model in Figure 2.3 which could be an option for policymakers on the use of systemic assessment such as ANA for monitoring the quality of mathematics education. This model was adapted from models on basic school functionality input-process-output that shows the functionality of a system from teachers and parents to school and classroom factors and how these contribute to the achievements of students (Drent et al., 2013). However, in the input-process-output model there is a lack of a key element that is relevant to systemic assessment, especially national assessment. This is *redress,* an addition to the model contributed by this study. Through applying the model, as an option for policymakers might reflect first on the input that outlines a national assessment and secondly the processes that are driven by well-known mathematical conceptions such as SMP and last outputs that are reflective of these. The redress may then focus on how the deficit of SMP can be addressed in terms of the input, quality, equity and provision. The input-process-output model for school functionality lacks a mathematical component. The introduction of the SMP could make the model suitable for the monitoring of mathematics education. Success in the use of SMP has been well-documented in the relevant literature (Groves, 2012 Suh, 2007) and also in the mathematics curriculum that is driven by SMP such as the SMC (Naroth & Luneta, 2015).

5.8 Limitations of the Study

First, ANA was delimitated to the Grade 9 when it is also administered to Grade 3 and 6. However, it is argued that ANA gathers information at regular intervals. An argument would be that it is premature to generalise about the effectiveness of ANA in monitoring mathematics education in South Africa, if one interval was assessed while leaving out two more intervals. However, the information that is presented by this study can be used to assess the performance of the system at that level since it seemed problematic for the DBE. This limitation suggests that further studies need to be done in the other intervals, Grade 3 and Grade 6 to gather comprehensive information and make conclusive inferences about the effectiveness of ANA.

5.9 Concluding Remarks

In my concluding remarks, I reflect on the assumptions that I brought to this study. The current study was informed by the transformative paradigm. I mentioned earlier that a transformative paradigm is a framework that allows researchers to identify inequalities in society and the promotion of social justice (Mertens, 2010b).

Firstly, the current study has revealed that ANA posed mostly low complexity questions which deprives learners of the opportunity to display high level problem solving abilities. The current study has revealed this inequality and now policymakers have the information to promote social justice by including high complexity test items in ANA. The literature has revealed that powerful assessments such as ANA have the ability to shape the mathematics that is taught in classrooms.

Secondly, another assumption that was brought to this study was that 'WYTIWUG' (Schoenfeld, 2007: 72). ANA has prioritised low complexity questions and irrespective of the fact that learners were not proficient. This is a sign of a system that is not doing well and directs policymakers to address the deficit of knowledge, skills and dispositions in terms of quality, equity and provision. Although most of the findings point to challenges on ANA as a tool to monitor the quality of the mathematics education standard in South Africa, the lessons learnt from testing in the three consecutive years are that the system needs a vigorous revamp in numeracy.

Lastly, research such as that of Porter (2002) stressed the importance of content standards as key to the success of educational reform. Subsequently, these are visible when calculating the alignment between curriculum and assessment. As such, frequent alignment of ANA is essential in South Africa to ensure that standards of mathematics education are achieved and maintained. This will be effective if policy documents such as the SAGM are revised (Berger et al., 2010) to include reasoning and proof which will then filter through to classroom practice such as assessments.

6. REFERENCES

Aaron, WR & Herbst, PG. 2015. Teachers' Perceptions of Students' Mathematical Work While Making Conjectures: An Examination of Teacher Discussions of an Animated Geometry Classroom Scenario. *International Journal of STEM Education*, 2(10): 1-13.

Adu, EO & Olaoye, O. 2015. Language Proficiency and Method of Instruction as a Determinant of Grade 9 Students' Academic Performance in Algebra. *International Journal of Educational Sciences*, 8(3): 547-555.

Adu, EO, Akinloye, GM & Adu, KO. 2015. School Input and Teacher Effectiveness in Some Local Government Areas of Lagos State, Nigeria. *International Journal of Educational Sciences,* 8(3): 461-472.

Alcock, L & Inglis, M. 2008. Doctoral Students' Use of Examples in Evaluating and Proving Conjectures. *Educational Studies in Mathematics,* 69: 111-129.

Ally, N & Christiansen M. 2013. Opportunities to Develop Mathematical Proficiency in Grade 6 Mathematics Classroom in Kwazulu-Natal. *Perspectives in Education*, 31(3): 106-121.

Ally, N. 2011. The Promotion of Mathematical Proficiency in Grade 6 Mathematics Classes from the UMgungundlovu District in Kwazulu-Natal (Master's Thesis), University of KwaZulu Natal, Pietermaritzburg. From http://hdl.handle.net/10413/5791 (accessed 11 August 2015).

Amir-Mofidi, S, Amiripour, P & Bijan-zadeh, M. 2012. Instruction of Mathematical Concepts through Analogical Reasoning Skills. *Indian Journal of Science and Technology,* 5(6): 2916-2922.

Amit, M & Fried, MN. 2002. High-Stakes Assessment as a Tool for Promoting Mathematical Literacy and the Democratization of Mathematics Education. *Mathematical Behavior,* 21*:* 499-514.

Angrist, JD, Pathak, PA & Walters, CR. 2013. Explaining Charter School Effectiveness. *American Economic Journal: Applied Economics*, 5(4): 1-27.

Association for Mathematics Education of South Africa (AMESA). 2012. *2012 Annual National Assessment Analysis.* Johannesburg, University of the Witwatersrand.

Ayalon, M & Even, R. 2008. Deductive Reasoning: In the Eye of the Beholder. *Educational Studies in Mathematics*, 69(3): 235-247.

Bailey, AL, Blackstock, A & Heritage, M. 2015. At The Intersection of Mathematics and Language: Examining Mathematical Strategies and Explanations by Grade and English Learner Status. *The Journal of Mathematical Behavior*, 40: 6-28.

Bantwini, BD. 2010. How Teachers Perceive the New Curriculum Reform: Lessons from a School District in The Eastern Cape Province, South Africa. *International Journal of Education,* 3: 1-25.

Baroody, A, Feil, Y & Johnson, AR. 2007. An Alternative Reconceptualization of Procedural and Conceptual Knowledge. *Journal for Research in Mathematics Education*, 38: 115-131.

Baroody, AJ. 2005. Discourse and Research on an Overlooked Aspect of Mathematical Reasoning. *The American Journal of Psychology,* 118(3): 484-489.

Bass, H. 2003. Computational Fluency, Algorithms and Mathematical Proficiency: One Mathematician's Perspective. *Teaching Children Mathematics,* 9(6): 322-327.

Bansilal, S. 2017. The Difficulty Level of a National Assessment of Grade 9 Mathematics: The Case of Five Schools. *South African Journal of Childhood Education*, 7(1): From: doi.org/10.4102/sajce.v7i1.412 (accessed 18 August 2017).

Bansilal, S. 2012. What can we learn from the KZN ANA results? *SA-eDUC Journal*, *9*(2): retrieved on 12 August 2014 from: www.nmu.ac.za/webfm_send/58394.

Berger, M, Bowie, L & Nyaumwe, L. 2010. Taxonomy Matters: Cognitive Levels and Types of Mathematical Activities in Mathematics Examinations. *Pythagoras*, 71: 30-40.

Bergqvist, T & Lithner, J. 2012. Mathematical Reasoning in Teachers' Presentations. *The Journal of Mathematical Behaviour*, 31: 252-269.

Berman, J. 2013. Utility of a Conceptual Framework within Doctoral Study: A Researcher's Reflections. *Issues in Educational Research,* 23(1): 1-18.

Bieda, KN. 2010. Enacting Proof-Related Tasks in Middle School Mathematics: Challenges and Opportunities. *Journal for Research in Mathematics Education,* 41(4): 351-382.

Blanton, M & Stylianou, DA. 2014. Understanding the Role of Transactive Reasoning in Classroom Discourse as Students Learn to Construct Proofs. *The Journal of Mathematical Behaviour*, 34: 76-98.

Bleiler-Baxter, SK & Pair JD. 2017. Engaging Students in Roles of Proof. *The Journal of Mathematical Behavior*, 47: 16-34.

Boesen, J, Lithner, J & Palm, T. 2010. The Relation between Types of Assessment Tasks and the Mathematical Reasoning Students Use. *Educational Studies in Mathematics,* 75: 89-105.

Bordage, G. 2009. Conceptual Frameworks to illuminate and magnify. *Medical Education,* 43: 312-319.

Bournot-Trites, M & Belanger, J. 2005. Ethical Dilemmas Facing Action Researchers. *The Journal of Educational Thought*, 39(2): 197-215.

Boyce, S & Norton, A. 2016. Co-Construction of Fractions Schemes and Units Coordinating Structures. *The Journal of Mathematical Behavior*, 41:10-25.

Breen, S & O'Shea, A. 2010. Mathematical Thinking and Task Design. *Irish Mathematical Society,* 66: 39-49.

Brodie, K. 2007. Dialogue in Mathematics Classrooms: Beyond Question-and-Answer Methods. *Pythagoras*, 66: 3-13.

Brodie, K. 2010. *Teaching Mathematical Reasoning in Secondary Classrooms.* 1st Edition. Dordrecht: Springer.

Callahan, LG & Garofalo, J. 1987. Metacognition and School Mathematics. *The Arithmetic Teacher*, 34(9): 22-23.

Cantlon, D. 1998. Kids + Conjecture = Mathematics Power. *Teaching Children Mathematics,* 5(2): 108-112.

Capps, LR & Pickreign, J. 1993. Language Connections in Mathematics: A Critical Part of Mathematics Instruction. *The Arithmetic Teacher,* 41(1): 8-12.

Carnoy, M, Ngware, M & Oketch, M. 2015. The Role of Classroom Resources and National Educational Context in Student Learning Gains: Comparing Botswana, Kenya, and South Africa. *Comparative Education Review*, 59(2): 199-233.

Chapman, O. 2013. Mathematical-Task Knowledge for Teaching. *Journal of Mathematics Teacher Education.* From: doi 10.1007/s10857-013-9234-7 (accessed 26 April 2014).

Chen, Q. 2014. Using TIMSS Data to Build Mathematics Achievement Model of Fourth Graders in Hong Kong. *International Journal of Science and Mathematics Education.* From: link.springer.com/article/10.1007%2 (accessed 3 August 2014).

Cheung, KC. 1988. Outcomes of Schooling: Mathematics Achievement and Attitudes towards Mathematics Learning in Hong Kong. *Educational Studies in Mathematics,* 19(2): 209-219.

Cohen, L, Manion, L & Morrison, K. 2011. *Research Methods in Education*. 7th Edition. London: Routledge.

Cragg, L & Gilimore, C. 2014. Skills Underlying Mathematics: The Role of Executive Function the Development of Mathematics Proficiency. *Trends in Neuroscience and Education.* From: http://dx.doi.org/10.1016/j.tine.2013.12.001 (accessed 10 May 2014).

Creghan, KA & Creghan, C. 2013. Assessing for Achievement. *Science and Children*, 51(3): 29-35.

Creswell, J. 2014. *Educational Research: Planning, Conducting and Evaluating Quantitative and Qualitative Research*. 5th Edition. New Jersey: Pearsons Education.

Davis, JD. 2005. Connecting Procedural and Conceptual Knowledge of Functions. *The Mathematics Teacher*, 99(1): 36-39.

DBE, 2013c. *Annual National Assessment 2013 Grade 9 Mathematics Test*, Pretoria. From: www.education.gov.za. (accessed 6 December 2013).

DBE, 2014b. *Annual National Assessment 2014 Grade 9 Mathematics Test*, Pretoria. From: www.education.gov.za (accessed 4 December 2014).

DBE. 2012a. *Report on the Annual National Assessment*, Pretoria. From: www.education.gov.za (accessed 20 August 2013).

DBE. 2012b. *Diagnostic Report. Annual National Assessment 2012*, Pretoria. From: www.education.gov.za (accessed 20 August 2013).

DBE. 2012c. *Annual National Assessment 2012 Grade 9 Mathematics Test*, Pretoria. From: www.education.gov.za (accessed 20 August 2013).

DBE. 2012d. Annual National Assessment Grade 9 Mathematics Exemplar Questions. Pretoria. From: www.education.gov.za (accessed 20 August 2013).

DBE. 2013a. *National Senior Certificate Examination: Technical report 2012*. Pretoria. From: www.education.gov.za (accessed 20 October 2014).

DBE. 2013b. *Report on the Annual National Assessment of 2013:* Grades 1 to 6 & 9, Pretoria From: www.education.gov.za (accessed 18 December 2013).

DBE. 2014. *Report on the Annual National Assessment of 2014: Grades 1 to 6 & 9.* Pretoria. From: www.education.gov.za (accessed 20 December 2014).

Department for International Development. [Sa]. *Guidance Note: A DFID practice paper.* From: http://www.gov.uk/government/uploads/.../nat-int-assess-stdnt-ach.pdf (accessed 15 September 2013).

Department of Basic Education (DBE). 2011. *Curriculum and Assessment Policy Statements Grades 7-9.* Pretoria. From: www.education.gov.za (accessed 30 July 2015).

Department of Education and Training (DoET). 2002a. *National Curriculum Statements*, Pretoria. From: www.education.gov.za (accessed 22 August 2013).

Dhlamini, ZB & Kibirige, I. 2014. Grade 9 Learners' Errors And Misconceptions In Addition Of Fractions. *Mediterranean Journal of Social Sciences*, 5(8): 227-235.

Dhlamini, ZB & Luneta, K. 2016. Exploration of the Levels of Mathematical Proficiency Displayed by Grade 12 Learners in Responses to Matric Examinations. *International Journal of Educational Sciences*, 12(2): 231-246.

Dieltiens, V & Meny-Gilbert, M. 2012. In class? Poverty, Social Exclusion and School Access in South Africa. *Journal of Education*, 55: 127-144.

Dobbie, WB & Fryer, RG. 2013. Getting Beneath the Veil of Effective Schools: Evidence from New York City. *American Economic Journal: Applied Economics*, 5(4): 28-60.

DoET. 2002b. *Revised National Curriculum Statement Grades R-9 RNCS (Schools) Mathematics,* Pretoria. From: www.education.gov.za (accessed 22 August 2013).

DoET. 2003. *National Curriculum Statement Grades 10-12 (General) Mathematics*, Pretoria: Government Printer.

DoET. 2007. *Subject Assessment Guidelines: Mathematics,* Pretoria. From: www.education.gov.za (accessed 22 August 2013).

Drent, M, Meelissen, MR & van der Kleij, FM. 2013. The Contribution of TIMSS to the Link between School and Classroom Factors and Student Achievement. *Journal of Curriculum Studies,* 45(2): 198-224.

DuBois, JM. 2002. When is Informed Consent Appropriate in Educational Research? Regulatory and Ethical Issues. *Ethics & Human Research*, 24(1): 1-8.

Dunne, T, Long, C, Graig, T & Venter, E. 2012. Meeting the Requirements of both Classroom-Based and Systemic Assessment of Mathematics Proficiency: The Potential of Rasch Measurement Theory. *Pythagoras,* 33(3): From: http://dx.doi.org/10.4102/pythagoras.v33i3.19 (accessed 15 September 2013).

Ebdon, SA, Coakley, MM & Legnard, D. 2003. Mathematical Mind Journeys: Awakening Minds to Computational Fluency. *Teaching Children Mathematics*, 9(8); 486-493.

Edwards, N. 2010. An Analysis of the Grade 12 Physical Sciences Examination and the Core Curriculum in South Africa. *South African Journal of Education,* 30: 571-590.

Emmett, J & McGee, D. 2013. Extrinsic Motivation for Large-Scale Assessments: A Case Study of a Student Achievement Program at One Urban High School. *The High School Journal*, 96(2): 116-137.

Fagnant, A & Crahay, M. 2011. Theories of Mind and Personal Epistemology: Their Interrelation and Connection with the Concept of Metacognition. *European Journal of Psychology in Education*, 26(2): 257-271.

Feldman, R. 2013. Filled Radar Charts Should not be used to Compare Social Indicators. *Social Indicators Research,* 111: 709-712.

Ferrer-Esteban, G. 2016. Trade-Off between Effectiveness and Equity? An Analysis of Social Sorting between Classrooms and between Schools. *Comparative Education Review.* From: http://www.jastor.org/stable/10.1086/684490 (accessed 8 January 2016).

Ferrini-Mundy, J & Schmidt, WH. 2005. International Comparative Studies in Mathematics Education: Opportunities for Collaboration and Challenges for Researchers. *Journal for Research in Mathematics Education,* 36(3): 164-175.

Flowers, J, Kline, K & Rubenstein, RN. 2003. Developing Teachers' Computational Fluency: Examples in Subtraction. *Teaching Children Mathematics*. From, www.nctm.org (accessed 5 August 2015).

Foster, C. 2011. Productive Ambiguity in the Learning of Mathematics. *For the Learning of Mathematics*, 31(2): 3-7.

Fraenkel, JR & Wallen, NE. 2010. *How To Design and Evaluate Research in Education*. 6th Edition. New York: MacGraw Hill.

Froneman, S, Plotz, M & Vorster, H. 2015. A Comparison of the Mathematical Knowledge and Skills of First-year Student Cohorts from a Transmission and an Outcomes-based Curriculum. *African Journal of Research in Mathematics, Science and Technology Education*, 19(1): 45-56.

Fulmer, GW. 2011. Estimating Critical Values for Strength of Alignment among Curriculum, Assessments and Instruction. *Journal of Educational and Behavioral Statistics.* From: http://jeb.sagepub.com/content/early/2011/02/02/1076998610381381397 (accessed 13 May 2014).

Gardee, A & Brodie, K. 2015. A Teacher's Engagement with Learner Errors in her Grade 9 Mathematics Classroom. *Pythagoras*. From: http://dx.doi.org/10.4102/pythagoras.v36i2.293 (accessed 11 January 2016).

Gay, LR, Mills, GE & Airasian, PW. 2014. *Educational Research: Competences for Analysis and Applications.* 10th Edition. London: Pearson Education Limited.

Gibbs, GR. 2012. Different Approaches to Coding. *Sociological Methodology*, 42: 82-84.

Gierl, MJ & Bisanz, J. 1995. Anxieties and Attitudes Related to Mathematics in Grade 3 and 6. *The Journal of Experimental Education,* 63(2): 139-158.

Goldman, S. 2006. A New Angle on Families: Connecting the Mathematics of Life with School Mathematics. *Counterpoints,* 249: 55-76.

Gonzales, R & Fuggan, CG. 2012. Exploring the Conceptual and Psychometric Properties of Classroom Assessment. *The International Journal of Educational and Psychological Assessment,* 9(2), 45-60.

Gorard, S. 2005. The Advantages of the Mean Deviation. *British Journal of Educational Studies*, 53(4): 417-430.

Gough, J. 2007. Conceptual Complexity and Apparent Contradictions in Mathematics Language. *Australian Mathematics Teacher*, 63(2): 8-16.

Granberg, C. 2016. Discovering and Addressing Errors during Mathematics Problem-Solving-A Productive Struggle? *The Journal of Mathematical Behavior*, 42: 33-48.

Graven, M & Stott, D. 2012. Conceptualising Procedural Fluency as a Spectrum of Proficiency. *Proceedings of the 18th Annual National Congress for the Association for Mathematics Education in South Africa,* 1: 146-156.

Graven, M & Venkat, H. 2014. Primary Teachers' Experiences Relating to the Administration Processes of High-Stakes Testing: The Case of Mathematics Annual National Assessments. *African Journal of Research in Mathematics, Science and Technology Education*. From: http://dx.doi.org/10.1080/10288457.2014.965406 (accessed 2 September 2016).

Graven, M & Venkatakrishnan, H. 2013. ANAs: Possibilities and Constraints for Mathematical Learning. *Learning and Teaching Mathematics,* 14: 12-16.

Greenes, C. 2014. Tasks to Advance the Learning of Mathematics. *Journal of Learning of Mathematics,* 5*(*1): 1-7.

Greenleess, J. 2011. The Fantastic Four of Mathematics Assessment Items. *Australian Primary Mathematics Classroom*, 16(2): 23-29.

Groves, S. 2012. Developing Mathematical Proficiency. *Journal of Science and Mathematics, 3*(2): 119-145.

Guberman, R & Leikin, R. 2013. Interesting and Difficult Mathematical Problems: Changing Teachers' Views by Employing Multiple-Solution Tasks. *Journal of Mathematics Teacher Education.* From: doi10.1007/s10857-012-9210-7 (accessed 26 May 2014).

Gysling, J. 2016. The Historical Development of Educational Assessment in Chile: 1810-2014. *Assessment in Education: Principles, Policy & Practice*, 23(1): 8-25.

Haapasalo, L & Kadijevich, D. 2000. Two Types of Mathematical Knowledge and Their Relation. *Journal of Mathematics Didaktik*, 21: 139-157.

Hallett, D, Nunes, T, Bryant, P & Thorpe, CM. 2012. Individual Differences in Conceptual and Procedural Fraction Understanding: The role of abilities and school experience. *Journal of Experimental Child Psychology*, 113: 469-486.

Halverson, R, Kelley, C & Shaw, J. 2014. A Call for Improved School Leadership. *The Phi Delta Kappan*, 95(6): 57-60.

Hanna, G. 2000. Proof, Explanation and Exploration: An Overview. *Educational Studies in Mathematics,* 44: 5-23.

Hannula, MS. 2002. Attitudes towards Mathematics: Emotions, Expectations and Values. *Educational Studies in Mathematics,* 49(1): 25-46.

Harel, G. 2017. The Learning and Teaching of Linear Algebra: Observations and Generalizations. *The Journal of Mathematical Behavior*, 46: 69-95.

Harrits, GS. 2011. More than Method?: A Discussion of Paradigm Differences within Mixed Methods Research. *Journal of Mixed Methods Research,* 5(2): 150-166.

Hauserman, CP & Stick, SL. 2013. The Leadership Teachers Want from Principals: Transformational. *Canadian Journal of Education*, 36(3): 184-203.

Hecht, SA & Vagi, KJ. 2012. Patterns in Children's Knowledge about Fractions. *Journal of Experimental Child Psychology,* 111: 212-229.

Hess, KK. 2006. *Exploring Cognitive Demand in Instruction and Assessment*. National Center for the Improvement of Educational Assessment, Dover NH. From: http://www.org/publications/DOK_ApplyingWebb_KH08.pdf (accessed 11 September 2016).

Hiebert, J & Lefevre, P. 1986. *Conceptual and Procedural Knowledge. The case of mathematics,* 1-27, Hillsdale, NJ: Erlbaum.

Hiebert, J. 1986. *Conceptual and Procedural Knowledge: The Case of Mathematics.* (Eds.). London: Lawrence Erlbaum Associates.

Hoffert, SB. 2009. Mathematics: The Universal Language? *The Mathematics Teacher*, 103(2): 130-139.

Hofman, RH, Hofman, WH & Gray, JM. 2015. Three Conjectures about School Effectiveness: An Exploratory Study. *Cogent education*. From: http://dx.doi.org/10.1080/2331186x.2015.1006977 (accessed 28 July 2015).

Hofstee, E. 2015. *Constructing a Good Dissertation: A Practical Guide to Finishing a Master's MBA or PhD on Schedule.* 3rd Edition. Johannesburg: EPE.

Howie, S. 2003. Language and Other Background Factors Affecting Secondary Pupils' Performance in Mathematics in South Africa. *African Journal of Research in Science Mathematics and Technology Education,* 7: 1-20.

Howie, S. 2004. A National Assessment in Mathematics within an International Comparative Assessment. *Perspectives in Education,* 22(2): 149-162.

Hsu, HY & Silver EA. 2014. Cognitive Complexity of Instructional Tasks in a Taiwanese Classroom: An Examination of Task Sources. *Journal of Research in Mathematics Education*, 45(4): 460-496.

Human Sciences Research Council (HSRC). 2013. Highlights from TIMSS 2011 the South African Perspective, Pretoria: HSRC Press.

Hungi, N, Makuwa, D, Ross, K, Saito, M, Dolata, S, van Cappelle, F, Paviot, L & Vellien, J. 2010. *SACMEQ III Project Results: Pupil achievement levels in reading and mathematics,* working document number 1.

Hwang, WY, Su, JH, Huang, YM & Dong, JJ. 2009. A Study of Multi-Representation of Geometry Problem Solving with Virtual Manipulatives and Whiteboard System. *Educational Technology & Society*, 12(3): 229–247.

Hyde, A, George, K, Mynard, S, Hull, C, Watson, S & Watson, P. 2006. Creating Multiple Representations in Algebra: All Chocolate, No Change. *Mathematics Teaching in the Middle School*, 11(6): 262-268.

International Association for the Evaluation of Educational Achievement (IEA). 2013. *TIMSS 2011 International Results in Mathematics, Executive Summary*, TIMSS & PIRLS International Study Center, and Lynch School of Education, Boston College, and Chestnut Hill.

Jantjies, M & Joy, M. 2015. Mobile Enhanced Learning in a South African Context. *Educational Technology & Society,* 18 (1): 308–320.

Johnson, B & Christensen, L. 2012. *Educational Research: Quantitative, Qualitative and Mixed Approaches.* 4th Edition. California: SAGE Publications.

Jonsson, B, Norqvist, M, Liljekvist, Y & Lithner, J. 2014. Learning mathematics through algorithmic and creative reasoning. *The Journal of Mathematical Behavior*, 36: 20-32.

Jonassen, DH. 2001. Review: How Can We Learn Best from Multiple Representations? *The American Journal of Psychology*, 114(2): 321-327.

Jones, BR, Hopper, PF, Franz, DP, Knott, L & Evitts, TA. 2008. Mathematics: A Second Language. *The Mathematics Teacher,* 102(4): 307-312.

Kagan, J, Pearson, L & Welch, L. 1966. Conceptual Impulsivity and Inductive Reasoning. *Child Development,* 37(3): 583-594.

Kalchman, M. 2011. Using the Everyday Life to Improve Student Learning: The Math in Everyday Life Homework Assignment Builds Student Confidence and Competence in Mathematics. *Middle School Journal,* 43(1): 24-31.

Kanjee, A & Moloi, Q. 2016. A Standard-based Approach for Reporting Assessment Results in South Africa. *Perspectives in Education*, 34(4): 29-51.

Kanjee, A & Moloi, Q. 2014. South African Teachers' Use of National Assessment Data. *South African Journal of Childhood Education*, 4(2): 90-113.

Karakoc, G & Alacaci, C. 2015. Real World Connections in High School Mathematics Curriculum and Teaching. *Turkish Journal of Computer and Mathematics Education*, 6(1): 31-46.

Kaulinge, P. 2012. Analysing the Relationship between Conceptual and Procedural Knowledge in a Departmentally Provided Numeracy Work Book. *Proceedings of the 18th Annual National Congress of the Mathematics Education in South Africa,* 1: 203-211.

Kazemi, E & Stipek, D. 2001. Promoting Conceptual Thinking in Four Upper-Elementary Mathematics Classrooms. *The Elementary School Journal*, 102(1): 59-80.

Kellaghan, T. Greaney, V & Murray, ST. 2009*. Using the Results of a National Assessment of Educational Achievement,* Washington, DC: World Bank.

Kepner, HS & Huinker, D 2012. Assessing Students' Mathematical Proficiencies on the Common Core. *Journal of Mathematics Education at Teachers College,* 3: 26-32.

Khairani, AZ & Nordin, MS. 2011. The Development and Construct Validation of the Mathematical Proficiency Test for 14-year-old Students. *Asia Pacific Journal of Educators and Education*, 26(1): 33-50.

Khashan, KH. 2014. Conceptual and Procedural Knowledge of Rational Numbers for Riyadh Elementary School Teachers. *Journal of Education and Human Development*, 3(4): 181-197.

Kilpatrick, J, Swafford, J & Findell, B. 2001. (Eds.). *Adding it up: Helping Children Learn Mathematics*, Washington D.C: National Academy Press, 115-155.

Kim, YR, Park, MS, Moore, TJ & Varma, S. 2013. Multiple Levels of Metacognition and their Elicitation through Complex-Solving Tasks. *The Journal of Mathematical Behavior*, 32: 377-396.

Kimberlin, CL & Winterstein, AG. 2008. Validity and Reliability of Measurement Instruments used in Research. *American Journal for Health-System Pharmacy,* 65: 2276-2284

Kiwanuka, HN, Van Damme, Noortgate, WVD, Anumendem, DN & Namusisi, S. 2015. Factors Affecting Mathematics Achievement of First-Year Secondary Students in Central Uganda. *South African Journal of Education*, 35(3): 1-16.

Klauer, KJ & Phye, GD. 2008. Inductive Reasoning: A Training Approach. *Review of Educational Research,* 78(1): 85-123.

Klieger, A. 2016. Principals and Teachers: Different Perceptions of Large-scale Assessment. *International Journal of Educational Research*, 75: 134-145.

Knuth, EJ. 2002. Proof as a Tool for Learning Mathematics. *The Mathematics Teacher,* 95(7): 486-490.

Koh, KH, Tan, C & Ng, PT. 2012. Creating Thinking Schools through Authentic Assessment: The Case of Singapore. *Educational Assessment, Evaluation and Accountability,* 24: 135-149.

Koichu, B & Leron, U. 2015. Proving as Problem Solving: The Role of Cognitive Decoupling. *The Journal of Mathematics Behavior*, 40: 233-244.

Komatsu, K, Jones, K, Ikeda, T & Narazaki, A. 2017. Proof Validation and Modification in Secondary School Geometry. Journal of Mathematical Behavior, 47: 1-15.

Kontorovich, I, Koichu, B, Leikin, R & Berman, A. 2012. An Exploratory Framework for Handling the Complexity of Mathematical Problem Posing in Small Groups. *The Journal of Mathematical Behavior*, 31: 149-161.

Koretz, D. 2009. How Do American Students Measure Up? Making Sense of International Comparisons. *The Future of Children,* 19(1): 37-51.

Kotze, GS & Strauss, JP. 2006. Contextual Factors of the Mathematics Learning Environment of Grade 6 Learners in South Africa. *Pythagoras,* 63: 38-45.

Kramarski, B, Mevarech, ZR & Arami, M. 2002. The Effects of Metacognitive Instruction on Solving Mathematical Authentic Tasks. *Educational Studies in Mathematics*, 49(2): 225-250.

Krefting, L. 1991. Rigor in Qualitative Research: The Assessment of Trustworthiness. *The American Journal of Occupational Therapy*, 45(3): 214-222.

Lachmy, R & Koichu, B. 2014. The Interplay of Empirical and Deductive Reasoning in Proving "If" And "Only If" Statements in a Dynamic Geometry Environment. *The Journal of Mathematical Behavior,* 36: 150-165.

Land, TJ. 2017. Teacher Attention to Number Choice in Problem Posing. *The Journal of Mathematical Behavior*, 45: 35-46.

Lathrop, RL. 1961. A Quick-but Accurate-Approximation to the Standard Deviation of a Distribution. *The Journal of Experimental Education*, 29(3): 319-321.

Leckie, G & Baird, J. 2011. Rater Effects on Essay Scoring: A Multilevel Analysis of Severity Draft, Central Tendency, and Rater Experience. *Journal of Educational Measurement in Education,* 48(4): 399-418.

Lee, CY & Chen, MJ. 2015. Effects of Worked Examples Using Manipulatives on Fifth Graders' Learning Performance and Attitude toward Mathematics. *Journal of Educational Technology & Society,* 18(1): 264-275.

Lee, J. 2012. Prospective Elementary Teachers' Perceptions of Real-Life Connections Reflected in Posing and Evaluating Story Problems. *Journal for Mathematics Teacher Education,* 15: 429-452.

Lee, K & Sriraman, B. 2011. Conjecturing Via Reconceived Classical Analogy. *Educational Studies in Mathematics,* 76(2): 123-140.

Lee, K. 2016. Students' Proof Schemes for Mathematical Proving and Disproving of Propositions. *The Journal of Mathematical Behavior*, 41: 26-44.

Leung, F. 2005. Can Achievement be attributed to Classroom Practices? Learning from the TIMMS Video Study. *Proceedings of the 11th Congress of the Association for Mathematics Education in South Africa,* 1: 14-25.

Leung, F. 2014. What can and should we learn from International Studies of Mathematics Achievement? *Mathematics Education Research Journal*. From: doi 10.1007/s13394-013-0109-0 (accessed 13 March 2014).

Levenson, E. 2013. Tasks that May Occasion Mathematical Creativity: Teachers' Choices. *Journal of Mathematics Teacher Education,* 16: 269-291.

Long, C & Wendt, H. 2017. A Comparative Investigation of South Africa's High Performing Learners on Selected TIMSS Items Comprising Multiplicative Concepts. *African Journal of Research in Mathematics, Science and Technology Education*, 21(2): 109-124.

Luneta, K & Dhlamini, Z. 2012. Mathematical Proficiencies Displayed By Grade 12 Learners in their Response to Final Examination Questions in the Context of South African Schools in the Gauteng Province. *Proceedings of the 18th Annual National Congress for the Association for Mathematics Education in South Africa,* 1: 217-230.

Luneta, K. 2015a. Issues in Communicating Mathematically in Rural Classrooms in South Africa. *Journal of Communication,* 6(1): 1-9.

Luneta, K. 2015b. Understanding Students' Misconceptions: An Analysis of Final Grade 12 Examination Questions in Geometry. *Pythagoras,* 36(1): From: http://dx.doi.org/10.4102/pythagoras.v36i1.261 (accessed 7 August 2015).

Luyt, R. 2012. A Framework for Mixing Methods in Quantitative Measurement Development, Validation and Revision: A Case Study. *Journal of Mixed Methods Research* 6(4): 294-316.

Maharaj, A, Brijlall, D & Narain OK. 2015. Improving Proficiency in Mathematics through Website-based Tasks: A Case of Basic Algebra. *International Journal of Educational Sciences*, 8(2): 369-386.

Maoto, S, Masha, K & Maphutha, K. 2016. Where is the bigger picture in the teaching and learning of mathematics? *Pythagoras*, 37(1): From: http://dx.doi.org/10.4102/pythagoras. v37i1.338 (accessed 12 January 2017).

Markovits, H & Doyon, C. 2011. Using Analogy to Improve Abstract Reasoning in Adolescents: Not as Easy as it Looks. *European Journal of Psychology of Education,* 26(3): 355-372.

Martin, WG & Kasmer, L. 2010. Reasoning & Sense Making. *Teaching Children Mathematics,* 15(5): 284-291.

Martone, A & Sireci, SG. 2009. Evaluating Alignment between Curriculum, Assessment and Instruction. *Review of Educational Research*, *79*: 1322-1361.

McCormick, R. 1997. Conceptual and Procedural Knowledge. *International Journal of Technology and Design Education*, 7: 141-159.

McLeod, DB & Briggs, JT. 1980. Interactions of Field Independence and General Reasoning with Inductive Instruction in Mathematics. *Journal for Research in Mathematics Education,* 11(2): 94-103.

McMillan, JH & Schumacher, S. 2014. *Research in Education: Evidence-Based Inquiry.* 7th edition. London: Pearson Education Limited.

Mellone, M, Verschaffel, L & Van Dooren, W. 2017. The Effect of Rewording and Dyadic Interaction on Realistic Reasoning in Solving Word Problems. The *Journal of Mathematical Behavior*, 46: 1-12.

Mertens, DM. 2007. Transformative Paradigm: Mixed Methods and Social Justice. *Journal of Mixed Methods Research,* 1(3): 212-225.

Mertens, DM. 2010a. Philosophy in Mixed Methods Teaching: The Transformative Paradigm as Illustration. *International Journal of Multiple Research Approaches,* 4(1): 9-18.

Mertens, DM. 2010b. Transformative Mixed Methods Research. Qualitative Inquiry. From: doi: 10.1177/1077800410364612 (accessed 12 May 2015).

Mhlolo, MK, Venkat, H & Schafer, M. 2012. The Nature and Quality of the Mathematical Connections Teachers Make. *Pythagoras*, 33(1): From: http://dx.doi.org/10.4102/pythagoras.v33i1.22 (accessed 20 August 2013).

Modzuka, CM. 2017. *An Investigation of the 2012 Annual National Assessment Grade 6 Mathematics Instrument*. Unpublished Master's Dissertation, University of Pretoria.

Moraa, BS & Nyaga, MP. 2015. Relationship between Paternal Involvement in Pupils' Education and Academic Achievement among Primary School Pupils in Nairobi County, Kenya. *International Journal of Education and Social Science*, 2(10): 19-29.

Morgan, C. 1996. "The Language of Mathematics" Towards a Critical Analysis of Mathematics Tasks. *For the Learning of Mathematics,* 16(3): 2-10.

Morris, AN. 2002. Mathematical Reasoning: Adults' Ability to Make the Inductive-Deductive Distinction. *Cognition and Instruction,* 20(1): 79-118.

Morrow, SL. 2005. Qualitative and Trustworthiness in Qualitative Research in Counselling Psychology. *Journal of Counselling Psychology*, 52(2): 250-260.

Mueller, M, Yankelewitz, D & Maher, C. 2011. Sense Making as Motivation in Doing Mathematics: Results from Two Studies. *The Mathematics Educator,* 20(2): 33-43.

Mwakapenda, W. 2008. Understanding Connections in the School Mathematics Curriculum. *South African Journal of Education,* 28, 1: 189-202.

Naicker, I. 2015. School Principals Enacting the Values of Ubuntu in School Leadership: The Voices of Teachers. *Studies on Tribes and Tribals*, 13(1): 1-9.

Naroth, C & Luneta, K. 2015. Implementing the Singapore Mathematics Curriculum in South Africa: Experiences of Foundation Phase Teachers. *African Journal of Research in Mathematics, Science and Technology Education*. From: http://dx.doi.org/10.1080/10288457.2015.1089675 (accessed 11 January 2016).

National Council of Teachers of Mathematics. 1989. *Principals and Standards for School Mathematics*, Reston, VA: NCTM.

National Council of Teachers of Mathematics. 2000. *Principals and Standards for School Mathematics*, Reston, VA: NCTM.

Nazeem, E. 2010. An Analysis of the Alignment of the Grade 12 Physical Sciences Examination and the Core Curriculum in South Africa. *South African Journal of Education*, 30: 571-590.

Ndlovu, M & Mji, A. 2012. Alignment between South African Mathematics Assessment Standards and the TIMSS Assessment Frameworks. *Pythagoras,* 33(3): From: http://dx.doi.org/10.4102/pythagoras.v33is.182 (accessed 6 November 2013).

Neal, D. 2010. Aiming for Efficiency Rather Than Proficiency. *Journal of Economic Perspectives,* 24(3): 119-132.

Neuenhaus, N, Artelt, C, Lingel, K & Schneider, W. 2011. Fifth Graders Metacognitive Knowledge: General or Domain-Specific? *European Journal of Psychology of Education*, 26(2): 163-178.

Niemi, D. 1996. Assessing Conceptual Understanding in Mathematics: Representations, Problem Solutions, Justifications, and Explanations. *The Journal of Educational Research,* 89(6): 351-363.

Nurse, JRC, Agrafiotis, I, Creese, S, Goldsmith, M & Lamberts, K. [Sa]. Communicating Trustworthiness Using Radar Graphs: A Detailed Look. From: www.innovateuk.org (accessed 11 August 2015).

O'Halloran, KL. 2015. The Language of Learning Mathematics: A Multimodal Perspective. *The Journal of Mathematical Behavior*, 40: 63-74.

Otten, O, Bleiler-Baxter, SK & Engledowl, C. 2017. Authority and Whole-Class Proving in High School Geometry: The Case of Ms. Finley. *The Journal of Mathematical Behavior*, 46: 112-127.

Palha, S, Dekker, R, Gravemeijer, K & van Hout-Wolters, B. 2013. Developing Shift Problems to Foster Geometrical Proof and Understanding. *The Journal of Mathematical Behavior*, 32: 142-159.

Pedemonte, B & Balacheff, N. 2016. Establishing Links between Conceptions, Argumentation and Proof through the Ck¢-Enriched Toulmin Model. *The Journal of Mathematical Behavior*, 41:104-122.

Perry, JA & Atkins, SL. 2002. It's Not Just Notation: Valuing Children's Representations. *Teaching Children Mathematics*, 9(4):196-201.

Phillips, KJR. 2010. What Does "Highly Qualified" Mean for Student Achievement? Evaluating the Relationships between Teacher Quality Indicators and At-Risk Students' Mathematics and Reading. *The Elementary School Journal*, 110(4): 464-493.

Polikoff, MS, Porter, CA & Smithson, J. 2011. How well Aligned Are State Assessments of Student Achievement with State Content Standards. *American Educational Research Journal.* From: http://aerja.aera.net (accessed 12 October 2013).

Polya, G. 1945*. How to Solve it*. New Jersey: Princeton Science Library.

Porter, CA, MacMaken, J, Hwang, J & Yang, R. 2011. Common Core Standards: The U.S Intended Curriculum. *Educational Researcher,* 40(3): 103-116.

Porter, CA, Smithson, J, Blank, R & Zeidner, T. 2007. Alignment as a Teacher Variable. *Applied Measurement in Education*, 20: 27-51.

Porter, CA. 2002. Measuring the Content of Instruction. Uses in Research and Practice. *Educational Researcher,* 31(7): 3-14.

Pournara, C, Mpofu, S & Sanders, Y. 2015. The Grade 9 Maths ANA - What Can We See After Three Years? *Learning and Teaching of Mathematics*, 18: 34-41.

Poyraz, C, Gulten, C & Bozkurt, S. 2013. Analysis of the Relationship between Students' Success in Mathematics and Overall Success. *International Journal on New Trends in Education and Their Implications,* 4(1): 28-38.

Prochazkova, LT. 2013. Mathematics for Language, Language for Mathematics. *European Journal of Science and Mathematics Education,* 1(1): 23-28.

Quinn, AL, Evitts, TA & Heinz, K. 2009. Count on Number Theory to Inspire Proof. *The Mathematics Teacher,* 103(4): 298-304.

Ramful, A. 2014. Reversible Reasoning in Fractional Situations: Theorems-in-Action and Constraints. *The Journal of Mathematical Behavior*, 33: 119-130.

Reddy, V. 2006. *Mathematics and Science Achievement at South African Schools in TIMSS 2003*. Pretoria: Human Sciences Research Council Press.

Reid, DA. 2002. Conjectures and Refutations in Grade 5 Mathematics. *Journal for Research in Mathematics Education,* 33(1): 5-29.

Richardson, K, Carter, T & Berenson, S. 2010. Connected Tasks: The Building Blocks of Reasoning and Proof. *Australian Primary Mathematics Classroom,* 15(4): 17-23.

Richland, LE, Stigler, JW & Holyoak, KJ. 2012. Teaching the Conceptual Structure of Mathematics. *Educational Psychologist*, 47(3):189-203.

Riordain, MI & O'Donoghue, J. 2009. The Relationship between Performance on Mathematical Word Problems and Language Proficiency for Students Learning Through the Medium of Irish. *Educational Studies in Mathematics*, 71: 43-64.

Rittle-Johnson, B & Alibali, W. 1999. Conceptual and Procedural Knowledge of Mathematics: Does One Lead to the Other? *Journal of Educational Psychology*, 91(1): 175-189.

Rittle-Johnson, B & Schneider, M. 2015. *Developing Conceptual and Procedural Knowledge of Mathematics*. In R. Cohen Kadosh & A. Dowker (Eds.). Oxford Handbook of Numerical Cognition (pp. 1102-1118). Oxford UK: Oxford University Press. From: doi:10.1093/oxfordhb/9780199642342.013.014 (accessed 6 December 2016).

Roth, W-M, Ercikan, K, Simon, M & Fola, R. 2015. The Assessment of Mathematical Literacy of Linguistic Minority Students: Results of a Multi-Method Investigation. *The Journal of Mathematical Behavior*, 40: 88-105.

Russell, L. 2008. Connecting Mathematics. *Mathematics in School*, 37(4): 40-41.

Russell, SJ. 2000. Developing Computational Fluency with Whole Numbers in the Elementary Grades. *The New England Math Journal,* 32: 40-54.

Saary, MJ. 2008. Radar Plots: A Useful Way for Presenting Multivariate Health Care Data. *Journal of Clinical Epidemiology*, 60: 311-317.

Samkoff, A & Weber, K. 2015. Lessons Learned From an Instructional Intervention on Proof Comprehension. *The Journal of Mathematical Behaviour*, 39: 28-50.

Sanni, R. 2009. Mathematical Proficiency and Practices in Grade 7. *Learning and Teaching Mathematics,* 7: 25-28.

Savic, M. 2015. The Incubation Effect: How Mathematicians Recover from Proving Impasses. *The Journal of Mathematical Behavior*, 39: 67-78.

Schmidt, WH & Houang, RT. 2012. Curricular Coherence and the Common Core State Standards for Mathematics. *Educational Researcher,* 41(8): 294-308.

Schmidt, WH & McKnight, CC. 1998. What can We Really Learn from TIMSS? *American Association for the Advancement of Science*. From: http://www.jastor.org/stable/2896998 (accessed 6 October 2014).

Schneider, M & Stern, E. 2010. The Developmental Relations between Conceptual and Procedural Knowledge: A Multi-method Approach. *Developmental Psychology*, 46(1): 178-192.

Schneider, M, Rittle-Johnson, B & Star, JR. 2011. Relations among Conceptual Knowledge, Procedural Knowledge, and Procedural Flexibility in Two Samples Differing in Prior Knowledge. *Developmental Psychology,* 47(6): 1525-1538.

Schoenfeld, A. 1985. *Mathematical Problem Solving*, Orlando, FL Academic Press.

Schoenfeld, A. 2007. What is Mathematical Proficiency and how can it be assessed? *Assessing Mathematical Proficiency,* 53: 59-73.

Sherin, B & Fuson, K. 2005. Multiplication Strategies and Appropriation of Computational Resources. *Journal for Research in Mathematics Education,* 36(4): 347-395.

Sibanda, L. 2017. Grade 4 Learners' Linguistic Difficulties in Solving Mathematical Assessments. *African Journal of Research in Mathematics Science and Technology Education*, 21(1): 86-96.

Siegler, RS & Stern, E. 1998. Conscious and Unconscious Strategy Discoveries: A Microgenetic Analysis. *Journal of Experimental Psychology*, 127: 273-397.

Siegler, RS, Thompson, CA & Schneider, M. 2011. An Integrated Theory of Whole Number and Fractions Development. *Cognitive Psychology*. From: doi:10.1016/j.cogpsych.2011.03.001 (accessed 20 August 2013).

Sigley, R & Wilkinson, LC. 2015. Ariel's Cycles of Problem Solving: An Adolescent Acquires the Mathematics Register, *The Journal of Mathematics Behavior*, 40: 75-87.

Silver, EA, Mamona-Downs, J, Leung, S & Kenney, PA. 1996. Posing Mathematical Problems: An Exploratory Study. *Journal for Research in Mathematics Education,* 27(3): 293-309.

Simon, MA. 1996. Beyond Inductive and Deductive Reasoning: The Search for a Sense of Knowing. *Educational Studies in Mathematics*, 30(2): 197-209.

Singer, FM & Voica, C. 2013. A Problem-Solving Conceptual Framework and its Implications in Designing Problem-Posing Tasks. *Educational Studies in Mathematics*, 83: 9-26.

Skemp, RR. 1976. Relational Understanding and Instrumental Understanding. *Mathematics Teaching*, 77: 20-26.

Sofroniou, N & Kellaghan, T. 2004. The Utility of Third International Mathematics and Science Study Scales in Predicting Students' Examination Performance. *Journal of Educational Measurement,* 41(4): 331-329.

Sosibo, L. 2015. The Notion of School 'Functionality' in a Teaching-Practice Placement Policy. *International Journal of Educational Sciences*, 8(1): 43-52.

Soto-Johnson, H & Troup, J. 2014. Reasoning on Complex Plane via Inscriptions and Gesture. *The Journal of Mathematical Behavior*, 36: 109-125.

South African Teachers' Democratic Union (SADTU) [Sa]. Annual National Assessment (ANA): A SADTU Perspective, Johannesburg.

Spaull, N. 2010. *A Preliminary Analysis of SACMEQ III South Africa.* Stellenbosch Economic Working Papers: 11/11: University of Stellenbosch.

Spaull, N. 2016. Disentangling the Language Effect in South African Schools: Measuring the Impact of 'Language of Assessment' in Grade 3 Literacy and Numeracy. South African Journal of Childhood Education, 6(1): From: doi.org/10.4102/sajce.v6i.475 (accessed 15 March 2017).

Springer, K. 2010. *Educational Research: A Contextual Approach.* 1st Edition Danvers: John Wiley & Sons Inc.

Stacey, K & Vincent, J. 2009. Modes of Reasoning in Explanations in Australian Eighth Grade Mathematics Textbooks. *Educational Studies in Mathematics*, 72(3): 271-288.

Stalvey, HE & Vidakovic, D. 2015. Students' Reasoning about Relationships between Variables in a Real-World Problem. *The Journal of Mathematical Behavior*, 40: 192-210.

Stanic, GM, Silver, EA & Smith, MS. 1990. Teaching Mathematics and Thinking. *The Arithmetic Teacher*, 37(8): 34-37.

Star, JR & Stylianides, GJ. 2013. Procedural and Conceptual Knowledge: Exploring the Gap between Knowledge Type and Quality. *Canadian Journal of Science, Mathematics, and Technology Education*, 13(2): 169-181.

Star, JR. 2004. The Development of Flexible Procedural Knowledge in Equations Solving. Paper Presented at the Annual Meeting of the American Research Association, San Diego. From: https://msu.edu/~jonstar/papers/AERA04.pdf (accessed 23 June 2017).

Star, JR. 2005. Conceptualizing Procedural Knowledge. *Journal for Research in Mathematics Education,* 36: 404-411.

Stein, M, Grover, B & Henningsen, M. 1996. Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms. *American Educational Research Journal*, 33(2): 455-488.

Stephen J, Pape, SJ & Tchoshanov, MA. 2001. The Role of Representation(s) in Developing Mathematical Understanding. *Theory into Practice*, 40(2):118-127.

Stoyanova, E. 1998. Problem Posing in Mathematics Classrooms. In A. McIntosh, & N. Ellerton (Eds.). *Research in Mathematics Education: A Contemporary Perspective* (pp. 164-185). Edith Cowan University: MASTEC.

Stylianides, GJ, Stylianides, AJ & Shilling-Traina, LN. 2013. Prospective Teacher's Challenges in Teaching Reasoning-and-Proving. *International Journal of Science and Mathematics Education,* 11: 1463-1490.

Stylianides, GJ. 2008. An Analytic Framework of Reasoning-and-Proving. *For the Learning of Mathematics*, 28(1): 9-16.

Suh, JM. 2007. Typing it all Together: Classroom Practices That Promote Mathematical Proficiency for All Students. *Teaching Children Mathematics*. From: www.nctm.org (accessed 10 August 2013).

Sullivan, P, Borcek, C, Walker, N & Rennie, M. 2016. Exploring a Structure of Mathematics Lessons That Initiate Learning By Activating Cognition on Challenging Tasks. *The Journal of Mathematical Behavior*, 41: 159-170.

Sumpter, L & Hedefalk, M. 2015. Preschool Children's Collective Mathematical Reasoning during Free Outdoor Play. *The Journal of Mathematical Behavior*, 39: 1-10.

Suter, LE. 2000. Is Student Achievement Immutable? Evidence from International Studies on Schooling and Student Achievement. *Review of Educational Research,* 70(4): 529-545.

Suurtamm, C. 2012. Assessment Can Support Reasoning: Teachers Can Use Data From a Research Project to Enhance their Classroom Assessment Practices. *The Mathematics Teacher*, 106(1): 28-33.

Tchoshanov, MA. 2011. Relationship between Teacher Knowledge of Concepts and Connections, Teaching Practice, and Student Achievement in Middle Grades Mathematics. *Educational Studies in Mathematics*, 76(2): 141-164.

Tevfik, I & Ahmet, I. 2003. Conceptual and Procedural Learning in Mathematics. *Journal of the Korea Society of Mathematics Education Series D: Research in Mathematics Education,* 7(2): 91-99.

Thanheiser, E. 2014. Developing Prospective Teachers' Conceptions with Well-Designed Tasks: Explaining Successes and Analyzing Conceptual Difficulties. *Journal of Mathematics Teacher Education,* 45(3): From: doi 10.1007/s10857-014-9272-9 (accessed 26 May 2014).

Thomas, JP. 2000. Influences on Mathematics Learning and Attitudes among African American High School Students. *The Journal of Negro Education,* 69(3): 165-183.

Tillema, E & Hackenberg, A. 2011. Developing Systems of Notation as a Trace of Reasoning. *For the Learning of Mathematics,* 31(3): 29-35.

Tomic, W & Klauer, KJ 1996. On the Effects of Training Inductive Reasoning: How Far Does It Transfer and How Long Do the Effects Persist? *European Journal of Psychology of Education,* 11(3): 283-299.

Torrance, H. 2012. Triangulation, Respondent Validation and Democratic Participation in Mixed Methods Research. *Journal of Mixed Methods Research,* 6(2): 111-123.

Tripathi, PN. 2008. Mathematical Understanding through Multiple Representations. *Mathematics Teaching in the Middle School*, 13(8): 438-445.

Tsamir, P, Tirosh, D, Tabach, M & Levenson, E. 2010. Multiple Solution Methods and Multiple Outcomes-Is it a Task for Kindergarten Children? *Educational Studies in Mathematics*, 73: 217-231.

Veloo, A, Krishnasamy, HN & Abdullah, WSW. 2015. Types of Student Errors in Mathematical Symbols, Graphs and Problem-Solving. *Asian Social Science*, 11(15): 324-334.

Venkat, H & Adler, J. 2012. Coherence and Connections in Teachers' Mathematical Discourses in Instruction. *Pythagoras*. From: http://dx.doi.org/10.4102/pythagoras.v33i3.188 (accessed 22 June 2014).

Volante, L & Cherubini, L. 2010. Understanding the Connections between Large-Scale Assessment and School Improvement Planning. *Canadian Journal of Educational Administration and Policy Issue,* 115: 1-26.

Wang, Z, Osterlind, SJ & Bergin, DA. 2012. Building Mathematics Achievement Models in Four Countries Using TIMSS 2003. *International Journal of Science and Mathematics Education,* 10: 1215-1242.

Webb, NL. 2007. Issues related to Judging the Alignment of Curriculum Standards and Assessments. *Applied Measurement in Education*, 20: 7-25.

Wheeler, D. 2001. Mathmatisation as a Pedagogical Tool. *For the Learning of Mathematics,* 21(2): 50-53.

Whitacre, I, Azuz, B, Lamb, LLC, Bishop, JP, Schappelle, BP & Philipp, RA. 2017. Integer Comparisons across the Grades: Students' Justifications and Ways of Reasoning. *The Journal of Mathematical Behavior, 45*: 47-62.

Wium, AM & Louw, B. 2012. Continued Professional Development Of Teachers to Facilitate Language Used in Numeracy and Mathematics. *South African Journal of Communication Disorders*, 59: 8-15.

Wong, WK, Yin, SK, Yang, HH & Cheng, YH. 2011. Using Computer-Assisted Multiple Representations in Learning Geometry Proofs. *Educational Technology & Society*, 14 (3): 43–54.

Yilmaz-Tuzun, O & Topcu, MS. 2010. Investigating the Relationship among Elementary School Students' Epistemological Beliefs, Metacognition, and Constructivist Science Learning Environment. *Journal of Science Teacher Education*, 21(2): 255-273.

Yopp, DA. 2010. Inductive Reasoning to Proof. *Mathematics Teaching in the Middle School,* 15(5): 286-291.

Yopp, DA. 2015. Prospective Elementary Teachers’ Claiming In Responses to False Generalizations. *The Journal of Mathematical Behavior*, 39: 79-99.

Zandieh, M, Roh, KH & Knapp, J. 2014. Conceptual blending: Student Reasoning When Proving “Conditional Implies Conditional” Statements. *The Journal of Mathematical Behavior*, 33: 209-229.

Zazkis, D, Weber, K & Mejia-Ramos, JP. 2015. Two Proving Strategies of Highly Successful Mathematics Majors. *The Journal of Mathematics Behavior*, 39: 11-27.

Zohrabi, M. 2013. Mixed Method Research: Instruments, Validity, Reliability and Reporting Findings*. Theory and Practice in Language Studies*, 3(2): 254-262.

7. APPENDICES

Appendix A: Approval from the University of Limpopo

UNIVERSITY OF LIMPOPO

University of Limpopo
Faculty of Humanities
Executive Dean
Private Bag X1106, Sovenga, 0727, South Africa
Tel: (015) 268 4895, Fax: (015) 268 3425, Email:richard.madadzhe@ul.ac.za

DATE: 20 September 2016

NAME OF STUDENT:	**DHLAMINI, ZB**
STUDENT NUMBER:	**[201324949]**
DEPARTMENT:	**PhD – Mathematics Education**
SCHOOL:	**Education**

Dear Student

FACULTY APPROVAL OF PROPOSAL (PROPOSAL NO. FHDC2016/2330)

I have pleasure in informing you that your PhD proposal served at the Faculty Higher Degrees Meeting on 24 August 2016 and your title was approved as follows:

TITLE: THE EFFECTIVENESS OF ANNUAL NATIONAL ASSESSMENT IN MONITORING MATHEMATICS EDUCATION STANDARD IN SOUTH AFRICA

Note the following:

Ethical Clearance	**Tick One**
Requires no ethical clearance Proceed with the study	
Requires ethical clearance (Human) (TREC) (apply online) Proceed with the study only after receipt of ethical clearance certificate	√
Requires ethical clearance (Animal) (AREC) Proceed with the study only after receipt of ethical clearance certificate	

Yours faithfully

Prof RN Madadzhe
Executive Dean: Faculty of Humanities

Director:	**Dr RS Maoto**
Supervisor:	**Dr K Masha**
Co-supervisor:	**Prof I Kibirige**

Finding solutions for Africa

Appendix B: Ethical Clearance from the University of Limpopo

University of Limpopo
Department of Research Administration and Development
Private Bag X1106, Sovenga, 0727, South Africa
Tel: (015) 268 2212, Fax: (015) 268 2306, Email:noko.monene@ul.ac.za

TURFLOOP RESEARCH ETHICS COMMITTEE CLEARANCE CERTIFICATE

MEETING:	**25 January 2017**
PROJECT NUMBER:	**TREC/04/2017: PG**

PROJECT:

Title:	The effectiveness of annual national assessment in monitoring Mathematics Education standard in South Africa
Researchers:	Mr ZB Dhlamini
Supervisor:	Dr K Masha
Co-Supervisor:	Prof I Kibirige
School:	Education
Degree:	PhD in Mathematics Education

PROF TAB MASHEGO
CHAIRPERSON: TURFLOOP RESEARCH ETHICS COMMITTEE

The Turfloop Research Ethics Committee (TREC) is registered with the National Health Research Ethics Council, Registration Number: **REC-0310111-031**

Note:
i) **Should any departure be contemplated from the research procedure as approved, the researcher(s) must re-submit the protocol to the committee.**
ii) **The budget for the research will be considered separately from the protocol.**
PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES.

Finding solutions for Africa

Appendix C: Letter Seeking Consent from the Department of Education: Limpopo Province

University of Limpopo
Centre for Academic Excellence
Private Bag X1106, Sovenga, 0727, South Africa
Tel: (015) 268 3439/3109, Fax: (015) 268 3109/3208, Email: Kwena.Masha@ul.ac.za

Date: 03 November 2014

The Head of Department
Limpopo Department of Education
Corner 113 Biccard & 24 Excelsior Street,
Polokwane

Subject: Application for the use of Grade 9 Annual National Assessment scripts for a PhD Research Project

This letter serves to humbly request access to grade 9 Annual National Assessment (ANA) scripts in some schools in the Capricorn District of the Limpopo Province. The access to the actual question papers is been sort with the National Department of Basic Education

The request is in line with my current Doctor of Philosophy (Mathematics Education) with the University of Limpopo under the supervision of Dr Masha. The study is entitled: **Using Mathematics Annual National Assessment Tests to Monitor South Africa's Education System.**

The purpose of the study is to investigate the nature of mathematics that is promoted through these tests and the challenges and experiences of the learners as they respond to those. Other national or international mechanisms or tests such as TIMMS are also forming part of the investigation. Permission to access and use the tests is being solicited through the National Office.

The University of Limpopo subscribes to high research ethics which will not at any stage of this research bring the Department of Basic Education into disrepute. I can also confirm that the copyrights of the learners' responses will not be breached in the process. Learners' identities and those of their schools will not be reflected in all the research reports.

Kind regards

Zwelithini Bongani Dhlamini
Student number: 201324949

Dr. Kwena Masha (Supervisor)

Finding solutions for Africa

UNIVERSITY OF LIMPOPO

TURFLOOP CAMPUS

Faculty of Humanities

School of Education

Department of Mathematics, Science and Technology Education

Email: Zwelithini.Dhlamini@ul.ac.za

28 November 2016

TO : The Head of Department

Limpopo Department of Education

Corner 113 Biccard & 24 Excelsior Street

Polokwane

RE: CHANGE OF TITLE IN PHD RESEARCH PROJECT

This letter serves to humbly request change of title for the previously approved permission to conduct research. The letter of previous approval is attached.

After engaging with the piloting of the study, the title has now been finalised as follows:

The Effectiveness of Annual National Assessment in Monitoring Mathematics Education Standard in South Africa.

Please amend the approval to conform to the newly approved title.

Regards,

Mr. Z.B. Dhlamini

……………………

Student No: 201324949

Finding solutions for Africa

Appendix D: Letter of Approval: Department of Education: Limpopo Province

LIMPOPO
PROVINCIAL GOVERNMENT
REPUBLIC OF SOUTH AFRICA

DEPARTMENT OF
EDUCATION

Enquiries: MC Makola PhD, Tel No: 015 290 9448 .E-mail: MakolaMC@edu.limpopo.gov.za

UNIVERSITY OF LIMPOPO

PRIVATE BAG X1106

SOVENGA

0727

DHLAMINI ZB

RE: Request for permission to Conduct Research

1. The above bears reference.
2. The Department wishes to inform you that your request to conduct research has been approved. Topic of the research proposal: **"USING MATHEMATICS ANNUAL NATONAL ASSESSMENT TEST TO MONITOR SOUTH AFRICA`S EDUCATION SYSTEM".**
3. The following conditions should be considered:

   3.1 The research should not have any financial implications for Limpopo Department of Education.

   3.2 Arrangements should be made with the Circuit Office and the schools concerned.

   3.3 The conduct of research should not anyhow disrupt the academic programs at the schools.

   3.4 The research should not be conducted during the time of Examinations especially the fourth term.

   3.5 During the study, applicable research ethics should be adhered to, in particular the principle of voluntary participation (the people involved should be respected).

Cnr. 113 Biccard & 24 Excelsior Street, POLOKWANE, 0700, Private Bag X9489, POLOKWANE, 0700
Tel: 015 290 7600, Fax: 015 297 6920/4220/4494

The heartland of southern Africa - development is about people!

3.6 Upon completion of research study, the researcher shall share the final product of the research with the Department.

4. Furthermore, you are expected to produce this letter at Schools/ Offices where you intend conducting your research as an evidence that you are permitted to conduct the research.
5. The department appreciates the contribution that you wish to make and wishes you success in your investigation.

Best wishes.

2015/02/02.

Mashaba KM **Date**

Acting Head of Department.

LIMPOPO
PROVINCIAL GOVERNMENT
REPUBLIC OF SOUTH AFRICA

DEPARTMENT OF
EDUCATION

Ref: 2/2/2 Enq: MC Makola PhD Tel No: 015 290 9448 E-mail:MakolaMC@edu.limpopo.gov.za

Dhlamini ZB
UNIVERSITY OF LIMPOPO
Private bag X1106
Sovenga
0727

RE: REQUEST FOR PERMISSION TO CONDUCT RESEARCH

1. The above bears reference.
2. The Department wishes to inform you that your request to conduct research has been approved. Topic of the research proposal: **"THE EFFECTIVENESS OF ANNUAL NATIONAL ASSESSMENT IN MONITORING MATHEMATICS EDUCATION STANDARD IN SOUTH AFRICA".**
3. The following conditions should be considered:

3.1 The research should not have any financial implications for Limpopo Department of Education.

3.2 Arrangements should be made with the Circuit Office and the schools concerned.

3.3 The conduct of research should not anyhow disrupt the academic programs at the schools.

3.4 The research should not be conducted during the time of Examinations especially the fourth term.

3.5 During the study, applicable research ethics should be adhered to; in particular the principle of voluntary participation (the people involved should be respected).

3.6 Upon completion of research study, the researcher shall share the final product of the research with the Department.

REQUEST FOR PERMISSION TO CONDUCT RESEARCH: DHLAMINI ZB

CONFIDENTIAL

4 Furthermore, you are expected to produce this letter at Schools/ Offices where you intend conducting your research as an evidence that you are permitted to conduct the research.

5 The department appreciates the contribution that you wish to make and wishes you success in your investigation.

Best wishes.

Ms NB Mutheiwana
Head of Department

Date 06/12/16

REQUEST FOR PERMISSION TO CONDUCT RESEARCH: DHLAMINI ZB

CONFIDENTIAL

Appendix E: Access to Use the 2011 TIMSS Grade 8 test items

Dear Zwelithini,

Thank you for your interest in TIMSS.
You are free to use the grade 8 mathematics items for your research. We ask that you properly acknowledge the source, including the year and name of the assessment you are using. You can find our released items policy at http://timssandpirls.bc.edu/timss2011/international-released-items.html below is the text from the website that explains the policy.
TIMSS and PIRLS are copyrighted and are registered trademarks of IEA. Released items from TIMSS and PIRLS assessments are for non-commercial, educational, and research purposes only. Translated versions of items remain the intellectual property of IEA. Although the items are in the public domain, please print an acknowledgement of the source, including the year and name of the assessment you are using. If you publish any part of the released items from TIMSS 2011, please use the following acknowledgement:

SOURCE: TIMSS 2011 Assessment. Copyright © 2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College, Chestnut Hill, MA and International Association for the Evaluation of Educational Achievement (IEA), IEA Secretariat, Amsterdam, the Netherlands.

Best of luck with your research!

Kerry

Appendix F: Certificate of Editing

BERNICE BRADE EDITING Member of the Professional Editors' Guild

FREELANCE WRITER, PROOF READER AND EDITOR
WEB RESEARCHER AND RESEARCH STRATEGIST
ENGLISH SPECIALIST
ESTABLISHED 1987

Tel. and Fax +27 11 465 4038
Cell 072 287 9859
Email edit@iafrica.com

P O Box 940
LONEHILL 2062
South Africa

3 September 2017

To whom it may concern: Certificate of Editing

This letter serves to confirm that **in August 2017** I did the proofreading and the language editing for the thesis of

ZWELITHINI BONGANI DHLAMINI

Student Number 201324949

Titled: ***THE EFFECTIVENESS OF ANNUAL NATIONAL ASSESSMENT IN MONITORING MATHEMATICS EDUCATION STANDARDS IN SOUTH AFRICA***

This document is being submitted in fulfilment of the requirements for the degree

PhD in MATHEMATICS EDUCATION

In the Faculty of Humanities in the SCHOOL OF EDUCATION

At the UNIVERSITY OF LIMPOPO

I have proofread and edited the entire thesis, including the introductory pages, the list of references and the appendices. This editing principally involves proofreading, language, style and grammar editing; and also checking the text for clarity of meaning, sequence of thought and expression and tenses. I have also noted any inconsistencies in thought, style or logic, and any ambiguities or repetitions of words and phrases, and have corrected those errors which creep into all writing. I have written the corrections on the hard copy and have returned the document to the author, who is responsible for inserting these. Please note that this confirmation refers only to editing of work done up to the date of this letter and does not include any changes which the author or the supervisor may make later.

3 September 2017

Bernice McNeil ***BA Hons NTSD***

If editors respect the academic purpose of thesis writing and the priority of the supervisor, we can help students (and ourselves). As one member told us: "We are a valuable resource for students as long as we edit these papers in an ethical way—a way in which … the work that students submit is indeed their own, only more polished." Guidelines for Editing Theses - The Editors' Association of Canada/l'Association canadienne des réviseurs

Material for editing or proofreading should ideally be submitted in hard copy. In electronic copy, it is too easy for the student to accept editorial suggestions without thinking about their implications Queensland University of Technology Higher Degree Research Guidelines

Proprietor: Bernice McNeil BA Hons, NSTD Member of the Classical Association of South Africa
Member of the English Academy of Southern Africa