What does an "A" symbol in Physical Sciences represent about a learner's skills and knowledge in the subject? A study of the cognitive demand of the National Senior Certificate Physical Sciences examination question papers.

Nagesh Munsamy

Submitted in partial fulfilment of the academic
requirements for the degree of
Master of Education in the
School of Education
University of KwaZulu-Natal

October 2014

ABSTRACT

In this study, I report on the cognitive demand of the South African National Senior Certificate (NSC) final public physical sciences examination question papers from 2008 to 2013. A cognitive demand taxonomy was developed and strengthened by me together with the assistance of four subject experts. The raters' analysis for the November 2012 physics and chemistry papers was found to be reliable. I have analysed six physics and six chemistry examination papers. On average, my study analysis found that the NSC Physical Sciences papers from 2008 to 2013 tested 94% lower order thinking skills of "retrieval", "basic comprehension" and "routine execution" while 6% comprised of "application" type skills. No higher order skills on "analysis and reasoning" were tested in the NSC papers. The physics and chemistry papers were both over-represented with the combined "routine execution" and "application" type questions, with the chemistry papers also over-represented with the "recall" type questions. Both physics and chemistry papers were under-represented in "evaluation" and "synthesis" type questions when compared to the recommendations made in the National Curriculum Statement. Studies reviewed have indicated that teaching higher order thinking fused into everyday learning may lead to better performances and economic productivity. The findings of this study may be useful to physical science teachers, curriculum specialists, academics, assessment experts, quality assurance specialists, education officials, employers and other stakeholders in the field of science and technology. The study makes recommendations on the use of taxonomies to analyse question papers to ensure tests and examinations are keeping with recommended weightings regarding cognitive demands. The study addresses the problem of the quality of a symbol awarded to a learner in terms of cognitive skills and knowledge tested in the physical sciences NSC examination.

DECLARATION

I, Nagesh Munsamy (student number 212 558 711) declare that:

(i) The research reported in this thesis, except where otherwise indicated is my original work;

(ii) This thesis has not been submitted for any degree or examination at any other university;

(iii) This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons;

(iv) This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

a) their words have been re-written but the general information attributed to them has been referenced;

b) where their exact words have been used, their writing has been placed inside quotation marks, and referenced.

(v) The work described in this thesis was carried out in the School of Education, University of KwaZulu-Natal, from September 2013 to October 2014 under the supervision of Professor Paul Hobden (Supervisor).

(vi) The Ethical clearance No. HSS/1189/013M [Appendix A] was granted prior to undertaking the fieldwork.

Signed: ___________________________ Date:_________________________

As the candidate's Supervisor I, Paul Hobden, agree to the submission of this thesis.

Signed:_____________________________ Date:__________________________

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

My sincere gratitude to:

my supervisor, Professor Paul Hobden for his dedication, time, guidance and expertise;

the subject experts who sacrificed their time to assist me with the development of the cognitive demand taxonomy;

my family, friends and colleagues who assisted, supported and encouraged me through this academic journey.

DEDICATION

This dissertation is dedicated to my late mother Daley M. Pillay who, despite having limited formal schooling opportunities, greatly valued knowledge, reasoning and academic excellence.

Chapter 1

INTRODUCTION

"Education is not the learning of facts, but the training of the mind to think"

Albert Einstein

Assessment is an integral part of formal learning and is used to gauge the competencies of the candidates in a particular learning area using certain criteria. In South Africa, the results of high stakes examinations like the National Senior Certificate (NSC) exit high school examination inform parents, learners, teachers and other stakeholders of the performance of the learners. The Department of Basic Education (DBE) uses the NSC (also referred to as matric) results to monitor its own practices and plan improvement strategies. The performance of learners in the NSC examination is often used as a measure to grant them entry to higher learning institutions or the world of work. With so much emphasis placed on this examination, it is important that the NSC examination is viewed as credible by education authorities as well as academics, researchers, policy makers and the general public. This chapter acquaints the reader with the study by first discussing the educational context of the study. This is followed by the rationale, focus and development of the cognitive demand taxonomy. The research design is then discussed and this is followed by an overview of the remaining chapters of the study.

1.1 EDUCATIONAL CONTEXT OF THE STUDY

In attempting to redress the inequality in education, the Department of Education introduced reforms and revisions over a very short period of time resulting in teachers having little time to adapt to the changes. With these revisions came changes to the high school exit examinations. These changes are described below and form part of the background within which this study took place.

The first democratic government that came into power in 1994, made administrative and curricula changes in an attempt to transform the 19 unequal, apartheid education departments. This transformation was aimed at improving the opportunities, standards and quality of education of all learners (Orkin, 2006). The Department of Basic Education (2011a) describes the changes that are summarised below. The phasing in of curriculum 2005 with its outcomes-based approach in 1997 experienced implementation problems. The implementation prompted a review which was followed by the Revised National Curriculum Statement at the General Education and Training phase in 2000 and the National Curriculum Statement (NCS) phased in at grade 10 in 2006. The more recent changes to the NCS called the Curriculum and Assessment Policy Statements (CAPS) was phased in at grades R to 3 and grade 10 in 2012, grades 4 to 6 and grade 11 in 2013 and grades 7 to 9 and grade 12 in 2014. Education stakeholders had to adapt to the changing curriculum and assessment practices across all grades and particularly at senior certificate level.

The senior certificate examination standards are the subject of much debate (wa Kivilu, 2006). A brief description of the senior certificate examination councils follows, and provides a background to this study. Prior to 1994, nineteen education departments administered the public senior certificate examination where candidates could take their subjects at higher grade (HG) or standard grade (SG) levels. This resulted in different standards across the different departments but an equivalent senior certificate. The debate about the senior certificate examination standards appears ongoing irrespective of which organisation controls this high stakes examination (Lolwana, 2006). The control and assuring of standards of the senior certificate examination may be described by three phases namely the Joint Matriculation Board (JMB) from 1918 to 1992, the South African Certification Council, SAFCERT (1992 to 2001) and Umalusi (2002 to date). After the new government took over in 1994, changes to a single national examination system ensuring a common certificate for having passed a common examination at grade 12 level has been gradual. The nine provincial education departments controlled the setting of papers and administration of the examinations in 1996, still offering subjects at HG and SG levels and 14 years after the new government was formed, the first common public exit examination called the National Senior Certificate (NSC) examination was written by grade 12 learners in 2008. The new government grappled with the integrity and administration of the examination (wa Kivilu, 2006) and debates about the standards of this high stakes examination continue.

Foxcroft (2006) states that the criteria used for admission to higher education institutions are intensely debated. This brings into focus the quality of a National Senior Certificate pass. The requirements to obtain the NSC are summarised below. The NSC is a three year South African secondary school exit qualification at level 4 on the National Qualifications Framework covering the curriculum in grades 10, 11 and 12 (KwaZulu-Natal Department of Education, 2014). The Department of Basic Education (2013a) notes the following. Learners are required to offer two South African languages, Mathematical Literacy or Mathematics, Life Orientation and choose three other designated subjects. The NSC encompasses 12 years of formal schooling and is awarded when a candidate sits and passes the NSC examination together with the school-based continuous assessment based on the NCS. In order for a candidate to receive a NSC, s/he must achieve 40% in three subjects, one of which is an official language at Home Language level and achieve 30% in three other subjects. For a candidate to receive a NSC for admission to pursue a bachelor’s degree, Higher Education South Africa (HESA) requirements state that s/he must receive the NSC with a minimum of 30% in the language of learning and teaching of the university together with four designated subjects above 50% (Department of Education, 2003).

Naidoo (2006) indicates that the senior certificate examination serves a critical role for different stakeholders. He goes further to mention that for candidates it is a means to satisfy “career aspirations”; for parents, it reflects the success or failure of their “children’s efforts”; for teachers, they are a measure of their “effectiveness as professionals” and for society, they indicate “competency, achievement and worth of a nation” (p. 10). With such low requirements to obtain the NSC, academics, employers and the general public are questioning the standards of the NSC examination. Despite the criticism, the Department of Basic Education (2013a) stands firm on the issue that the NSC requirements are equal to, if not higher than, the old senior certificate (matric) requirements. This study analyses NSC physical sciences examination papers from 2008 to 2013 in terms of cognitive demand with a view to informing the standards debate with evidence of the cognitive demand of the examination questions.

1.2 RATIONALE

In my opinion, the problem with the NSC examination is that it does not afford learners adequate opportunity to engage in higher order thinking. This view is supported by Crowe (2012), Edwards (2010), Mhlolo, (2011) and Muller (2005) who indicate that the examinations tend to emphasise lower order thinking skills. More employers want their workers to think for themselves, solve problems and come up with new ideas to improve on current technology and systems. Governments across the globe want their citizens to engage in critical thinking and problem solving to improve the lives of all citizens. Teachers, academics and the general public are questioning the standards of high stakes examinations like the NSC. Mhlolo (2011) reveals that learners who are able to think for themselves are more economically productive when exposed to cognitively rich assessments. The composition of the NSC examination questions should thus be examined to see if it is cognitively challenging to address the concerns of the general public, teachers, academics, employers and government leaders.

As a physical sciences teacher, with 20 years of teaching experience, I have observed that over time, an increasing number of learners struggle with problem solving, critical thinking, making decisions and scientific reasoning. Even some learners who obtain “A” symbols in the subject don’t seem to have developed the skills to answer questions that require higher order thinking. These learners appear to have mastered the routine type of questions from the recent past years of examination papers and struggle to think beyond the routine when questions are designed to test application, analysis and reasoning. The Department of Basic Education (2003), in formulating the NCS, acknowledges that “ higher order thinking and problem solving skills are required to meet the demands of the labour market and for active citizenship within communities with increasingly complex technological, environmental and societal problems” (p. 13). However, a number of authors such as Edwards (2010), Mhlolo (2011), Muller (2005), Umalusi (2009), wa Kivilu (2006) bring into question, the standards of this high stakes examination and the value of “A” symbols.

The Department of Basic Education maintains that the matric standards are improving, citing national overall pass rates having improved from 47% in 1997 to 78% in 2013. Muller (2005) claims that despite this upward progression in pass rates, actual standards have dropped because the level of cognitive demand was low. These improvements in the pass rates merely indicate that candidates have met the minimum requirements for the NSC and passed subjects like physical sciences at 30% and do not give an indication of the cognitive demand of the examination questions. This may create a false sense of improvement in standards. A study like this research undertaking will present evidence of the levels of the cognitive demand of the grade 12 NSC examination questions in physical sciences.

From the literature surveyed, I found that in-depth studies of the National Senior Certificate (NSC) content standards in physical sciences were limited. Among the studies on cognitive demand, Umalusi (2004) looked at examination question papers, memoranda, marking process and the findings revealed that there was evidence pointing to declining levels in conceptual understanding. An Umalusi (2009) report compared the cognitive demands of 2008 NSC with that of past senior certificate examination papers and found that the emphasis on the area of problem solving had declined in the NSC. Muller (2005) claimed that the standards of the senior certificate examination were of low cognitive demand. Edwards' (2010) study investigated an alignment of the physical sciences curriculum and the 2008 and 2009 examination question papers. Crowe's (2012) study on Senior Certificate (SC) Biology standards analysed question papers and answer scripts in an attempt to understand standards in terms of cognitive demand from an evidence-based methodological perspective. International studies such as Liu, Zhang, Liang, Fulmer, Kim and Yuan (2008) compare standards in physics among the New York State, Singapore and Chinese examinations by analysing cognitive levels. This research study will add to the literature on the cognitive demand of the South African NSC physical sciences examination question papers.

The findings of this study may be relevant to physical science teachers, curriculum specialists, tertiary academics, assessment experts, examiners, moderators, quality assurance specialists, education policy makers and employers in the field of science and technology. The study may make recommendations on the use of taxonomies to analyse question papers and ensure tests and examinations are keeping with recommended weightings regarding cognitive demands. In this study, all physical sciences examination papers will be analysed in terms cognitive skills and knowledge tested. This will determine the cognitive demand of the papers and indicate the value of a symbol awarded for the

NSC physical sciences examination. Stakeholders will be able to make judgements by examining the findings of this study and similar studies to address their concerns with regard to the type of skills and knowledge that learners acquire in the NSC examinations.

1.3 FOCUS OF THE STUDY

The focus of this study is the cognitive demand of the grade 12, public, NSC physical sciences final examination question papers. My main objectives in this study were to develop a taxonomy to measure the cognitive demand of physical sciences examination questions, determine the cognitive demand of the NSC physical sciences examination question papers by a process of deductive analysis using the taxonomy, and draw comparisons between the cognitive demand of the NSC physical sciences examination question papers and the recommendations made by the DBE in the NCS. In order to do this I framed this study with the following critical research questions:

i) What are the cognitive demands of the NSC physical sciences examination question papers?

ii) How do the cognitive demands of the NSC physical sciences examination question papers compare with the weighting of cognitive levels recommended in the national curriculum statement?

To answer my research questions, I first developed the cognitive demand taxonomy, analysed the examination questions according to this taxonomy and thereafter made comparisons between the percentage weighting of cognitive levels recommended and the findings of the study.

1.4 DEVELOPMENT OF THE COGNITIVE DEMAND TAXONOMY

I initially examined various cognitive demand taxonomies to find a relevant taxonomy to analyse the Physical Sciences NSC examination questions. The following is a list of published taxonomies which I surveyed: Bloom's taxonomy (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956) [Appendix B]; the revised Bloom's taxonomy (Anderson et al., 2001) [Appendix C]; the TIMMS taxonomy (Reddy, 2006); Marzano's new taxonomy of educational objectives (Marzano & Kendall, 2007); the Department of

Education (2008a) physical sciences assessment taxonomy [Appendix D]; Umalusi (2009) cognitive demand instrument; the Department of Basic Education (2011a) physical sciences assessment taxonomy [Appendix E]; the performance expectations taxonomy (Crowe, 2012) [Appendix F]; and the taxonomy of introductory physics problems (Teodorescu, Bennhold, Feldman, & Medsker, 2013). The revised Bloom's taxonomy appears to be favoured by researchers as an instrument to analyse the cognitive demand of examination questions. I initially proceeded to use the revised Bloom's taxonomy customised for physical sciences to conduct my analysis. As I proceeded, I encountered problems with the classification of questions into cognitive categories. I then re-evaluated the instrument and began developing a taxonomy which I found to be more relevant and specific to analyse the cognitive demand of NSC physical sciences examination questions. This process is fully described in my study.

1.5 RESEARCH DESIGN

Researchers operate in different paradigms. Maree (2007) states that post-positivists acknowledge that there is a reality independent of our thinking that science can study and that all observation and measurement is fallible and all theory is reversible. As a post-positivist researcher, I entered the study with the belief that there is an independent reality that we can understand partially and with more evidence our understanding of this reality improves. Knowledge that develops as a result of a post-positivistic worldview is based on observation and measurement (Creswell, 2003). I employed a quantitative design approach requiring the deductive analysis of examination question papers.

In order to answer my research questions, a taxonomy for the assessment of science skills and knowledge (TASSK) [Appendix G] was developed as an instrument relevant to measure cognitive demand of physical sciences assessment items. To measure the cognitive demand, the examination questions had to be classified into the different cognitive domains of the taxonomy. The taxonomy was first piloted with four subject experts and the researcher. Discussions of the study took place with all these experts at the outset and during the initial stages of the analysis. These took the form of face to face, telephonic and electronic mail discussions. Thereafter, I together with the four subject experts analysed the DBE November 2012 paper one and paper two. The examination papers used in this study were from a collection made by the researcher. Modifications to the taxonomy were made until there was agreement among raters that the taxonomy was reliable. The inter-rater reliability calculated was found to be acceptable. In this study, the group of physical sciences subject experts were purposively chosen because of their extensive experience in the subject. The subject experts helped to refine the taxonomy created by the researcher and also to verify that the taxonomy is reliable for the cognitive demand analysis of examination question papers.

The researcher then conducted the analysis of all the NSC examination question papers from 2008 to 2013 using the newly developed taxonomy. Each examination question paper was divided into the smallest unit that could be scored using the taxonomy and is termed an assessment item in this study. Each assessment item was measured against the newly developed taxonomy for the assessment of science skills and knowledge (TASSK) and categorised as cognitive level one to five representing the cognitive domains of “retrieval”, “basic comprehension”, “routine execution”, “application” and “analysis and reasoning”. Similar approaches applying other taxonomies were used in the studies by Crowe (2012), Edwards (2010), Liu et al. (2008), Näsström (2009) and Umalusi (2009) for measuring the cognitive demand of examination question papers.

1.6 STRUCTURE OF THE STUDY

This dissertation focuses on the cognitive demand of NSC physical sciences examination question papers and is divided into five chapters. The current chapter has provided a background and rationale to orientate the reader to the study. Chapter two reviews the literature on cognitive demand under the following headings: the South African senior certificate examination, cognitive demand and alignment studies. It discusses the role of assessment, historical development of assessment, large-scale assessments, types of assessment items, high stakes tests and examination standards. It also discusses cognitive demand taxonomies, higher order thinking skills and the alignment between the written, taught and tested curriculum. Chapter three describes the research design, validity and reliability, data collection strategy, ethical issues of the study and the development of the cognitive demand taxonomy. Chapter four presents the findings from the deductive analysis of all final NSC physical sciences examination question papers from 2008 to 2013. It also discusses inter-rater reliability, the DBE and TASSK analyses and compares the cognitive demand of examination papers with recommended weightings in the NCS. Chapter five provides a summary of findings, limitations, implications, recommendations and concludes the study report.

Chapter 2

LITERATURE REVIEW

Assessments, and in particular examinations that differentiate what skills and knowledge an individual retains from a course of study, have become an integral part of education in modern times. Many school teachers administer examinations without the full understanding of their cognitive composition, the value of the symbols obtained and how the results of the examination affect the candidate. This chapter will critically discuss the cognitive demand of assessment tasks, analyse higher order thinking skills that the current workplace is looking for, and examine whether the NSC examination questions adequately address these concerns. I first discuss the South African senior certificate examination system with a focus on examination standards. This is followed by investigating cognitive demand, discussing various cognitive demand taxonomies and critically analysing higher order thinking skills. I then discuss alignment studies on the intended and tested curriculum.

2.1 THE SOUTH AFRICAN SENIOR CERTIFICATE EXAMINATION

In this section, I discuss the role of assessment with the focus on summative assessment. I then describe the historical development of assessment to position the South African examination system in context. This is followed by a discussion of large-scale assessments and the types of assessment items in the NSC. I then explore high stakes examinations. Thereafter, I investigate whether the standards of the NSC examination are on a decline. I also show the effect of grade inflation on the actual acquisition of skills and knowledge in a certain field of study.

2.1.1 The role of assessment.

If we look into the role of assessment we find that it serves two main functions, namely, acting as gatekeepers to another stage of study or entry into the world of work and providing a quantitative measure of what is known by the learner. Assessment can be used in different contexts and can have a variety of meanings (Doran, Lawrenz, & Helgeson, 1994). Madaus and Kellaghan (1996) define assessment as “an activity designed to show what a person knows or can do” (p. 120). Assessment includes various methods for determining whether students are achieving the intended learning outcomes of instruction (Gronlund &Waugh, 2009). According to Doran et al. (1994), the purposes of assessment in science are “improvement of science instruction and programmes; conveying expectations to students, parents, teachers and administrators; monitoring the status of individuals, classes, districts, states, and the nation; and accountability” (p. 395). Assessments may be classified as either formative or summative. Black (2004) notes that the formative and summative labels, describe two ends of the assessment spectrum rather than two isolated and different functions. Black and Harrison (2010) describe formative assessment tasks as activities that provide information to be used as feedback to modify teaching and learning. McMillan (2011) states that “summative assessment is conducted after instruction primarily as a way to document what students, know, understand and can do” (p. 6). Summative assessments are thus all those assessments, given at the end of units, terms, or year of study. The NSC examination can thus be classified as a summative assessment which provides a measure of a learner’s knowledge and skills that affect his/her entry into the labour market or an institution of higher education.

2.1.2 Historical development of assessments.

Throughout history candidates have undertaken various types of tests and examinations be it oral presentations, practical assessment tasks, assignments, formal tests and year-end examinations to demonstrate their skills and knowledge. The criteria to determine success or failure in these examinations have been the subject of contention. This paragraph provides a background to the NSC examination and outlines the historical development of assessments. Madaus and Kellaghan (1996) describe below how all these assessment methods started. They narrate that initially, medieval guilds measured competence by requiring the apprentice to supply a product as proof of competence. This was assessed by master craftsmen through well-defined criteria. In ancient Greece, rules for oral examinations were offered by Aristotle, and apprentice scholars demonstrated competence of a subject through questions posed by the masters. The “viva voce” (oral) examinations were designed to reveal a student’s ability to remember acceptable knowledge and present it eloquently. Catechism, with its roots in the Christian faith, enjoyed considerable popularity with its oral question and answer testing based on a set curriculum and instruction. The earliest record of written testing is traced to China with the Mandarin civil service tests having written sections. As paper became more readily available in the West, written examinations were introduced in Europe in the 14th century. When mathematics was introduced into the curriculum, the oral methods of testing became unsuitable and written tests and examinations dominated formal assessment. As explorers left Europe in search of new lands, some of their systems took root in their newly established colonies, including the Cape colony on southern tip of Africa. It can be seen that over time various communities have attempted to assess the competence and skills of learners in a similar manner as modern educational systems do.

In order to place the current issues associated with the NSC in context, an account of the history of examinations for university admission in South Africa as described by Trumpelmann (1991) is provided. He states that the first centralised formal examination board founded in 1850 in South Africa was called the “Cape Public Service Board”. The purpose of the board was mainly to train public servants. In 1858, parliament recommended the formation of “Board of Public Examiners” whose function was to improve the general educational standards of the public service and the population. The first formal examination was conducted in 1858. Examinations conducted by the board yielded poor results with only five out of thirteen candidates passing in 1862. The control of examinations and prescribing standards for issuing certificates fell under the University of the Cape of Good Hope in 1873. This university eventually conducted a range of primary, secondary and tertiary examinations with the matriculation (matric) examination holding a key position. Towards the end of the 1800s the matriculation examination had become the only recognised route to higher education in the Cape Colony and later the whole of South Africa. In 1895, the number of candidates that sat for the examination had risen to 418. Matriculation regulations changed from time to time with the university demanding competence in English, Latin, Arithmetic, Algebra and Geometry in 1873; to six compulsory subjects and optional ancillary subjects in 1896; to subjects being selected from specific groups in 1909.

By 1913, there were indications that the matriculation examination was taking strain as the “desire to advertise the examination became apparent” (Trumpelmann, 1991, p. 3). He goes on further to quote *De Zuid-Afrikaan* who in 1960 stated: “To use examination results as indicators of an individual candidate's character and ability was just as effective as to use a ruler to measure the temperature of a patient" (pp. 3-4). From this quotation it appears that questioning the validity of the matric (senior certificate) examination, is not a new phenomenon and has been part of the dialogue since the 1900s.

According to Lolwana (2006), the formalisation, control and assuring of standards of the senior certificate examination, is summarised by three distinct phases. The first phase was under the control of the Joint Matriculation Board (JMB) from 1918 to 1992. The JMB controlled syllabi, curricula and moderation of question papers. A statistical moderation process to ensure standardisation was established following a report in 1963 and is still practised to date. The JMB had its fair share of criticisms regarding the question of standards during its tenure. The South African Certification Council, SAFCERT (1992 to 2001) marked the second phase of examination quality control and saw the centralisation of the certification process. Unlike the JMB, accounts of SAFCERT are not well documented and primarily anecdotal. SAFCERT's era was in a politically contested period and public examinations were plagued with high failure rates. Minister Bhengu commissioned an investigation into the senior certificate examination in 1998, followed by a study benchmarking the senior certificate examination with the Scottish Higher Examination commissioned by Minister Asmal in 1998 (Lolwana, 2006). These studies were a result of the pressure questioning the standards of the examination. Umalusi (2002 to date) continued to build on both SAFCERT and the JMB's approach in controlling the standard of the senior certificate examination (Lolwana, 2006). Umalusi is mandated to monitor qualifications, moderate examinations, scrutinise intended and examined curricula and make judgements on educational standards at primary and secondary levels (Umalusi, 2007). Lolwana (2006) argues that speculations of falling standards continued and were fuelled when the pass rate of the senior certificate examination rapidly climbed from 47% in 1997 to 73% in 2003. This prompted an Umalusi study on the examination in 2004. Umalusi (2004) looked at examination question papers, memoranda, marking process and statistical moderation process. The findings revealed that there was evidence pointing to declining levels in conceptual understanding at standard grade level and more learners opted for the simpler standard grade instead of higher grade. This in turn led to an increase in the overall pass rate. It appears from the discussions above that no matter who controls the quality assurance of this high stakes examination, the debate on standards continues.

After the new democratic government was constituted in 1994, the ex-departments continued with their administration in 1994 and 1995. The desire of the new government was to standardise the examination system to measure learners' competency in relation to collectively agreed criteria. The journey to a single national examination system ensuring a common certificate for having passed a common examination at grade 12 level has been gradual. Prior to 1994, nineteen education departments administered the public senior certificate examination where candidates could take their subjects at higher grade (HG) or standard grade (SG) levels. This resulted in different standards across the different departments but an equivalent senior certificate. In 1996, the nine provincial education departments took control of the setting of papers and the administration of the examinations still offering subjects at HG and SG levels. Continuous Assessment (CASS), introduced in schools in 1996, is a school-based assessment (SBA) score that is added to the examination score to obtain a final score for learner promotion purposes. At present, the physical sciences examination weighting is 75% of the final assessment score while the SBA constitutes 25%. In 2000, the national department of education started setting examination papers nationally in five key subjects, later increasing this to 13 subjects (Department of Basic Education, 2014a). Finally the old Senior Certificate examination based on the National Assembly Training and Education Department (NATED) 550 curriculum was phased out with the last full time cohort completing in 2007. This made way for the first common public exit examination called the National Senior Certificate (NSC) examination that grade 12 learners wrote in 2008. The NSC examinations are currently set nationally and administered by the nine provinces. The examination system has gone through many changes since 1994 in attempts to equalise standards for all and improve its quality. Throughout these changes, the question of the NSC examination standards continues to be debated.

2.1.3 Large-scale assessments in South Africa.

Examining the results of large-scale assessments like the provincial common testing programme, the Annual National Assessments (ANA) and the TIMSS show how learners perform at other levels prior to the NSC level and at international levels respectively. The ANA is administered by the DBE and the TIMSS set at international level, is administered by the Human Sciences Research Council (HSRC) in South Africa. Britton and Schneider (2007) state that the purpose of large-scale assessments includes reporting, accountability and international comparison. Lane (2004) argues that “large-scale assessments are primary tools for communicating what teachers should be teaching and what learners should be learning” (p. 12). She adds that they should be designed in a “cognitively rich” manner, assessing various levels of worthy cognitive demand.

At a provincial level there is a common testing programme for grades 10, 11 and 12. The information that follows is gathered from my experience as a physical sciences teacher. The large-scale common testing programme for grade 10 and 11 occurs at the end of each term in March, June and September in KwaZulu-Natal province. They cover the scope as reflected in the guidelines given by the DBE. Grade 12 learners write common provincial preparatory examinations in September and common tests in March and June each year. I have noticed that some questions in these preparatory examinations are a repeat of questions that appeared in past papers. These papers are marked by the class teacher and the scores obtained in these common tests form part of the school based continuous assessment comprising 25% of the final promotion mark of the grade 12 learner.

The ANA is a large-scale formative assessment carried out with the aim of identifying and improving literacy and numeracy in grades one to nine. The information that follows is obtained from the Human Sciences Research Council, HSRC (2013). In 2012, over 7 million learners wrote the ANA. While small improvements occurred in the junior grades, the ANA report for 2011 (Department of Basic Education, 2011b) and the ANA 2012 report (Human Sciences Research Council, 2013) indicate that the performance average of Grades 3, 4, 5, 6 and 9 learners in Mathematics and Languages is below 50 per cent. The ANA 2012 report indicates that only 2% of the learners that wrote Mathematics passed at 50% and above. This percentage pass for grade 9 may have serious implications for both mathematics and physical sciences as a poor grounding in problem solving will impact on learners’ performance at NSC level.

South Africa has featured the lowest in the rankings relative to participating countries in TIMSS. This appears to indicate that South African learners are not able to apply their skills and knowledge to these international tests. The HSRC administered TIMSS in South Africa in 1995, and subsequently in 1999 to Grade 8 learners. In 2002 it was administered to Grade 8 and 9 learners, and in 2011 to Grade 9 learners. For TIMSS

2011 in South Africa, the HSRC conducted the study in 285 schools among 11969 learners. The trend analysis from 1995 to 2011 showed that the national average score remained static over the years with South Africa ranking last (Human Sciences Research Council, 2012). The aim of the TIMSS study is to assess the national curricula, school and social environment, and the achievement in mathematics and science in participating countries across the world. The TIMSS tests were designed to measure mathematics and science achievement so as to help inform governments, policy makers and educators about the proficiency of their students at key points in the educational process (Human Sciences Research Council, 1998). While educational reform points towards more cognitively challenging assessments, Looney (2009) argues that most large-scale assessments do not test students' ability to draw upon more complex skills like problem solving and collaboration but are limited to formats that are easy to score and cost-effective to implement. Howie (2001) states that the international median benchmark in the TIMSS 1999 study indicates that learners can recognise and communicate basic scientific knowledge. She reports that only 6% of South African pupils reached this benchmark as compared to 83% of Chinese and 80% of Singapore and Japanese learners. The TIMSS 2002 study showed that less than 40% of all South African learners scored correctly on all cognitive levels (Reddy, 2006). The HSRC points out that "A striking feature of the mathematics and sciences scores is that the best performing South African learners approached the average performance of the top performing countries of Singapore, Chinese Taipei, the Republic of Korea, Japan, Finland, Slovenia and the Russian Federation" (Human Sciences Research Council, 2012, p. 1). Both the Chinese and Singapore examinations drive instruction to higher order reasoning (Liu et al., 2008). This shift towards focusing on higher order thinking skills is validated by China and Singapore and above-mentioned countries occupying top positions in international assessments like TIMSS.

Large-scale assessments like those discussed above can signal important goals for teachers, learners and other stakeholders to ensure that future high stakes tests and other assessments are cognitively rich in their design to prepare learners for entrance to universities or the world of work. However, from my literature survey it appears that little is published to address the area of the cognitive demand of tests and examinations.

2.1.4 Types of assessment items in the NSC examination.

Kuhn (2011) notes that "our knowledge of the role of tests and examinations as an agent for change is limited due to the lack of evidence with regard to subject specific task formats" (p. 190). With regard to the NSC examination, grade 12 physical sciences teachers have been provided with examination guidelines (Department of Education, 2009), which outline the format of the examination paper, weighting of sections, content to be tested and weighting of the levels of cognitive demand. Teachers are also provided with subject assessment guidelines (Department of Education, 2008a) and the NCS documents (Department of Education, 2003) which provide details of the curriculum and assessments standards. In this section, I will discuss the types of assessment items in the NSC physical sciences examination and comment in terms of cognitive demand.

Within the literature surveyed, many different types of assessment are reported. McMillan (2011) states that the major distinguishing characteristic of most assessments is whether the assessment items use selected-response or constructed-response formats. He adds that in the selected-response items, learners are given two or more possible responses while in constructed-response items, the learners are required to produce their answer. The NSC physical sciences examination consists of two sections. Section A contains one-word items and multiple choice items which are selected-response while section B of the question paper consists of constructed-response items.

One-word items mainly test the recall of definitions, laws, units and other factual knowledge. Learners mainly apply the "remember" skill to answer these questions. Memorisation may be through rote learning and when answers are correct for these items, teachers who mark the papers are unable to detect whether any constructive learning has taken place.

McMillan (2011) states that multiple-choice items have a stem, in the form of a question or statement, and three or more options as possible answers. One of the options is correct while two or more may be distracters. The physical sciences examination consists of four options in multiple-choice items from 1996 onwards. Prior to 1996, there were five options given in the examination. From my teaching experience, I have found that most learners who have not well prepared, generally do not leave multiple choice items blank as it requires less effort to provide a response. I also note that these items can also test the cognitive demand across the spectrum from recall to higher order thinking skills. Harlen (2010) argues that distracters may be eliminated by obvious incorrect statements, to a choice of two options and thus a 50% chance of guessing. She adds that correct answers are often chosen for the wrong reasons. Kubiszyn and Borich (2010) note that well-constructed multiple choice items are the most time-consuming kind of objective test items to formulate and warn against writing such items at the knowledge level only. Gronlund and Waugh (2009) state that the strengths of multiple-choice items are that they are easy to score, objective and reliable and that the incorrect options chosen by learners can serve as information for diagnostic analysis. They also add that multiple choice items are ineffective for measuring the ability to organize and express ideas. Black (1998) notes that despite their strengths, multiple-choice items cannot measure the ability of a learner to solve a problem, or carry out an investigation; they do not resemble the science taught in the classroom and don't emphasize conceptual and procedural understanding. He argues that "learners may often misread the demand of the item, seem incompetent because of a single slip in a complex process, fail to show what they know because they judge it as irrelevant, and can be marked down because the marker cannot understand the quality of the thinking behind responses" (p. 815). Despite the limitations, multiple choice items continue to feature in most examination systems.

Section B of the physical sciences examination consists of longer questions that are constructed-response items. Items of low cognitive demand test the "retrieval", "basic comprehension" and "routine execution" skills and items involving "application", "analysis and reasoning" test higher order thinking skills. These are not separated per question and each question tests a range of cognitive skills. When assessment items require learners to solve non-routine complex problems/tasks, learners will be operating on a higher cognitive level using skills and knowledge learned to apply, analyse and reason what is demanded of the question. From my experience as a teacher, I have found that learners who are exposed to a variety of question types that assess a range of cognitive skills tend to be comfortable and perform well in high stakes examinations like the NSC.

2.1.5 High stakes examinations.

Stakes in assessment circles refer to the consequences of the results of an assessment. An informal class test that does not contribute to a learner's promotion into the next grade may be classified as a low stakes assessment as it does not have serious consequences attached to it. An exit high school examination, however, may attach serious life-altering consequences like influencing career trajectory, university placement or finding a desired level of employment and may be classified as high stakes assessment. Johnson and Johnson (2009) explain high stakes testing as one in which the outcomes may have serious life-altering decisions including denial of certification, repetition of a grade, labelling of students, withholding of funds or closing down of a school.

The NSC is an example of a high stakes examination. The results of this examination have far reaching implications to the destiny of a learner. The certificate obtained on passing this examination is the minimum requirement for most entry level jobs. The scores obtained may determine the placement at a higher education institution and further determine the study discipline one chooses to embark on. The NSC results that are publicised in national and provincial print and electronic media stir much discussion in the communities catapulting learners, teachers, schools and districts into the limelight. Furthermore, the matric (grade 12) teachers and their schools, officials and districts are judged by the performance of learners in the NSC examinations.

While some critics have called for more oral presentation, demonstrations, and portfolio submissions to lend more authenticity to assessment, it is currently dominated by high stakes exit standardized tests / examinations at secondary school level in most industrialized countries (Gronlund & Waugh, 2009). Wilson (2007) adds that there is a net increase internationally in all testing, including high stakes testing. Harlen (2010) argues that "when the assessment stakes are high for students and teachers, what is tested inevitably has a strong determining effect on what is taught" (p. 137). My anecdotal observations in well-performing schools with appropriately qualified science teachers lead me to believe that this trend is prevalent. Johnson and Johnson (2009) state that high stakes test opponents argue that test scores are more likely related to socioeconomic status than to school test preparation punishing disadvantaged children who must compete with middle class and wealthy students in well-funded schools on the same high stakes tests. It would appear that to make judgements with different socio-economic groups regarding high stakes test scores like the NSC is not a “fair test” as many influencing factors are not kept constant.

The high stakes NSC examinations in South Africa appears to enjoy a fair amount of public support as an exit level examination. In contrast there appears to be much opposition to high stakes testing reported in the USA (Johnson & Johnson, 2009), England (Allen, 2012) and Australia (Caldwell, 2010). Allen (2012) argues that “it is common to oppose ‘high-stakes’ testing in schools in favour of a ‘low-stakes’ alternative as high-stakes testing is dominated by demands for accountability, accreditation and scientific rigour while low-stakes testing, by contrast, has a reputation for being softer and less judgemental” (p. 1). Proponents of high stakes testing believe that high standards and high stakes tests are essential to motivate students, teachers, and administrators to work ever harder to boost achievement (Johnson & Johnson, 2009). Alexander (2010) however, makes an interesting comparison between England and Finland, which has high levels of teacher and school autonomy, no national tests, enjoys top rankings internationally and has high performing school system. Although there is opposition to high stakes testing by some scholars in various countries, these tests appear to be increasingly dominant in global examination systems.

2.1.6 Are examination standards on a decline?

This section provides an overview of the debates about examination standards. Swain (2010) notes that “making value judgements on the standards of achievement has become a national obsession, particularly at the times of the reporting of public examinations” (p. 225). In South Africa this obsession begins each year with the official release of the NSC examination results by the Minister of Basic Education. The debate is intense but gradually loses momentum during the course of the year, spiking once again when the next set of results is announced. Content standards clarify what a learner should know, understand and apply in a particular learning area (McMillan, 2011). A content standard includes a description of the nature of knowledge and skills that learners should acquire. These standards are found in the National Curriculum Statements (NCS) and more recently in the Curriculum and Assessment Policy Statement (CAPS). wa Kivilu (2006) questions whether the public and key stakeholders in education have a common understanding of what standards actually mean. He states that educators may translate the broad national goals into classroom objectives differently and set different standards for their learners. Braun and Kanjee (2006) maintain that assessment policy and practices are critical to any successful educational improvement strategy. While assessment data may be essential to teaching and is needed to monitor, evaluate, and improve the education system, the effectiveness and rate of improvement in our public education system is questionable. Resnick, Rothman, Slattery, and Vranek (2011) point out that “standards reflecting high expectations for student performance should be the focus of instruction and tests should be a fair measure of whether students have attained the standards” (p. 4). Thus while standards may be in the policy documents, the DBE must constantly monitor and evaluate the standards of the assessment tasks set.

The Department of Basic Education is confident that its examination papers are at an appropriate standard. The information that follows is taken from the NSC 2012 Report by the Department of Basic Education (2013a). All question papers for the NSC examinations, were set by a national panel of examiners consisting of between three to five examiners per panel and one internal moderator. External moderators from Umalusi verified, evaluated and approved all question papers. The Department of Basic Education (2013a) states that according to the internal and external moderators’ reports, the papers set were of appropriate standard and covered the prescribed content and that the cognitive levels of these question papers were appropriate and accommodated problem solving and critical thinking skills demanded by the curriculum.

There are, however, questions from the general public and some academics regarding the value of the standards of the NSC examination. Black (2004) points out that schools have a responsibility to the public who fund them and the public interest should focus on whether agreed standards have been reached. The analysis of the results by the Department of Basic Education (2014a) reveals that there was an improvement in the percentage of learners who passed physical sciences from 55% in 2008 to 67% in 2013. According to Liu et al. (2008) a score of 65 points (out of 100) is the pass requirement for the New York State regents physics examination while a pass of 60% in the Jiangsu Province (China) exit physics examination qualifies a learner to write the national college entrance examination. This is in stark contrast to South Africa’s requirements where the pass mark for physical science is 30%. While the DBE statistics show an improvement in grade 12 pass rates, many scholars question the standards of the examination and the value of an "A" symbol.

Is the symbol that the learner receives a realistic indication of his/her performance in the assessment tasks? Jansen (2012) states that university lecturers, through experience, say that students have become weaker "even as the matriculation results on the outside get stronger" (p. 1). He believes that the reason for setting university admissions examinations is due to the NSC marks not reflecting the true knowledge and skills of learners. He cites inflation of marks as the factor giving learners more distinctions into "the ridiculous 90s". My experiences as a teacher show that many "A" learners struggle to answer questions that are not of a routine order. The same sentiment seems to be echoed by the NSC 2013 National Diagnostic Report which states "Many candidates, including 'A' candidates, could not express themselves clearly in questions that required an explanation. Higher order questions were therefore poorly answered" (Department of Basic Education, 2014b, p. 174).

The question arises: Are scores inflated to mask the dismal performance of mainly former disadvantaged schools? Black (2004) states that while data on achievement should consider learner backgrounds in making judgement on assessment results, poor outcomes should not be hidden by some form of manipulation. The Department of Basic Education (2014a) asserts that standardisation of marks as applied to the NSC results is an international practice in large-scale assessments to "mitigate fluctuations in learner performance that are as a result of factors within the examination processes rather than the knowledge, aptitude or abilities of learners" (p. 39).

The past five years of performances are used as a norm by Umalusi and marks are adjusted upwards (grade inflation) or downwards. Nel and Kistner (2009) cite grade inflation in traditional feeder schools of the University of Cape Town. They further report that there was a 31% increase in the "aggregate A and B intervals" for first year students from 1997 to 2003. Govender and Moodley (2012) note that grade inflation is not unique to this country and is usually associated with falling standards in assessment. They indicate that in 2009, there was a dramatic decrease in both the pass rate and the average module mark for a first year physics module (PHYS151) at the University of KwaZulu-Natal. They state that academic physics staff remained fairly stable and the teaching and examining of the module remained reasonably consistent and concluded that the "new national senior certificate curriculum produced students less prepared for university study than the previous senior certificate examination" (Govender & Moodley, 2012, p. 4). From the above arguments it appears that upward adjustments to raw scores in the NSC examination may contribute towards the inflation of a symbol awarded to a learner.

2.2 COGNITIVE DEMAND

From the above discussion, it would appear that examination scores and pass rates on their own do not reveal much about the skills and knowledge that are achieved in the assessment. Mhlolo's PhD study (2011) supports the view that knowing the cognitive demand of the content tested is important because it is the only measure of student achievement. In the literature that was surveyed, I could not find a suitable, broad definition for "cognitive demand". I therefore borrow from the ideas of Bloom et al. (1956), Anderson et al. (2001), and Henningsen and Stein (1997), to suggest the following definition for "cognitive demand" to be used in this study:
Cognitive demand of assessment items refers to the type of thinking processes that is required of the learner to successfully complete the items. These thinking processes may involve using lower order skills of retrieving of knowledge, comprehending familiar concepts and executing routine procedures and algorithms or may involve using higher order thinking skills of applying, analysing and reasoning the knowledge and skills gained to new situations.

2.2.1 The cognitive demand of assessment tasks.

There should be a balance between these two major thinking categories in examinations. Muller (2005) argues that "low cognitive demand is a threat to the health of the nation" (p. 5). He is of the view that it should be addressed at every stage of the educational cycle from curricular statements, to exemplar manuals, textbooks, learning material, examination papers and in marking standards. He supported his claim that standards have dropped despite the upward trend in pass rates post 1994 with empirical evidence from Umalusi research forums which analysed the cognitive demands of papers set in 1992, 1999, and 2003. The Department of Education (2003) and the Department of Basic Education (2008a) physical sciences NCS documents outline the assessment standards and mentions cognitive levels but does not address the cognitive demand to the extent that Muller (2005) details it. Lane (2004) reports that “the advances in the study of cognition and measurement in the 1980s and 1990s prompted individuals to think differently about how students process and reason with information” (p. 1) and how assessments can capture meaningful aspects of students’ thinking. It seems that a balance of productive and reproductive thinking skills is an important consideration in the setting of examinations.

It appears that examiners and teachers may lower the cognitive demand by teaching to the test. The majority of teachers, from my experience, appear to teach for examinations and coach learners to solve previous examination type questions. Herman and Baker (2009), however, argue that teachers tend to model the pedagogical approach reflected on high-visibility tests. Looney (2009) warns that teachers who observe patterns and prepare learners to focus only on those aspects of the content are engaging in the undesirable practice of coaching. I am aware of some learners who can afford private lessons (coaching / extra tuition) attend these classes after school hours. In these classes, learners are further drilled with skills on how to approach past questions to such an extent that for these learners, the bulk of the examination is somewhat routine provided that the testing style remains consistent. If, however, the questions are phrased differently (but testing the same skills and knowledge) these learners feel that the examination was unfair or difficult. Learners who receive extra tuition may attain high scores, but will be operating mainly at lower order cognitive levels of “retrieval”, “basic comprehension” and the “routine execution” of procedures and algorithms. Braun and Kanjee (2006) affirm that the practice of extra lessons may raise test scores but they also distort inferences about the comparative effectiveness of different schools or different regions. From the above discussions, it appears that cognitive demand is an important but neglected area when designing meaningful assessment tasks.

The Department of Basic Education (2011a) states that physical sciences develop the following skills in learners: “classifying, communicating, measuring, designing an investigation, drawing and evaluating conclusions, formulating models, hypothesising, identifying and controlling variables, inferring, observing and comparing, interpreting, predicting, problem solving and reflective skills" (p. 8). Many of these higher order skills do not appear to feature in the NSC examinations. The curriculum and assessment policy document purports to focus on aspects of higher order thinking skills like inquiry skills, but has replaced practical investigations with verification type experiments since 2010 for the formal assessment tasks. Harlen (2010) states that there is massive evidence of such `recipe-following' science that gives no opportunity for thinking skills to be used or developed in learners. Furthermore the NSC examination questions testing practical work are of low cognitive demand, mainly testing the comprehension of familiar concepts and routine procedures. Wheeler (2000) argues that both the basic content and inquiry must be fused into the learners' experiences. Upon analysis of CAPS (Department of Basic Education, 2011a) and the USA framework for inquiry practices by the National Research Council (2012), it appears that the CAPS is limited with respect to inquiry when compared to the practices set up by the National Research Council (2012) in the areas of; asking questions, planning and carrying out investigations, analysing and interpreting data, constructing explanations and engaging in argument for evidence. From my experience with other teacher colleagues, the NCS syllabus appears to be driven, primary because of time constraints, to cover what is tested in the NSC high stakes examination. The studies surveyed have indicated that using a suitable taxonomy is a good approach to interpret the cognitive demand of assessment tasks.

2.2.2 Cognitive demand taxonomies.

This section examines the cognitive demand taxonomies surveyed in the literature and focuses on the Taxonomy of educational objectives (Bloom et al., 1956), the revised Bloom's taxonomy (Anderson et al., 2001) and the taxonomies used by the DBE. The approach to measuring cognitive demand is also discussed in this section.

a) Discussion of taxonomies surveyed

This section will examine the commonly used taxonomies in the literature surveyed to classify the cognitive demand of the examination questions. Researchers have developed taxonomies in an attempt to understand the cognitive demand of examinations. According to Hess (2006), many have attempted to classify the cognitive demand in different learning and assessment contexts from Bloom's taxonomy developed in 1956 (and subsequently revised in 2001), to the national assessment of educational progress in 1990, to Porter's survey of enacted curriculum, to Webb's depth of knowledge levels in 1997. I have surveyed Bloom's taxonomy (Bloom et al., 1956), the revised Bloom's taxonomy (Anderson et al., 2001), the TIMMS mathematics and science taxonomies (Reddy, 2006), the new taxonomy of educational objectives (Marzano & Kendall, 2007), the Department of Education (2008a) physical sciences assessment taxonomy, Umalusi (2009) cognitive demand taxonomy, the Department of Basic Education (2011a) physical sciences assessment taxonomy, the performance expectations taxonomy (Crowe, 2012) and the taxonomy of introductory physics problems (Teodorescu et al., 2013). From the literature surveyed, the revised Bloom's taxonomy appears to be favoured by researchers as an instrument to analyse the cognitive demand of examination questions. The revised Bloom's taxonomy has been used as an interpretive framework by Edwards (2010), Umalusi (2009), Näsström (2009) and Liu et al. (2008). The Department of Education (2008a) and Department of Basic Education (2011a) also recommend the use of a physical sciences taxonomy that is an adaptation of Bloom et al. (1956) and the revised Bloom's taxonomy (Anderson et al., 2001) to ensure that assessment tasks comply with the recommended weightings of cognitive levels.

b) Bloom's taxonomy of educational objectives

Since Bloom's taxonomy and the revised Bloom's taxonomy dominate the literature as an instrument to measure cognitive demand, they warrant a closer inspection. The idea of developing a taxonomy to measure educational objectives was the work of a group of thirty four college and university examiners formed in 1948 at an informal meeting at the American Psychological Association Convention in Boston, USA (Bloom et al., 1956). What followed was a "Taxonomy of educational objectives, Handbook I: Cognitive Domain" published in 1956 by B. S. Bloom (editor), M. D. Engelhart, E. J. Furst, W. H. Hill and D. R. Krathwohl. Since its first publication, Bloom et al. (1956) has been translated into more than twenty languages (Anderson et al., 2001).

Bloom's taxonomy was broadly divided into two categories. The first category was called "knowledge" and involved recalling of knowledge that is stored. The second broad category was the "intellectual abilities and skills" that emphasised the mental processes of organising and reorganising the subject matter. This involved the cognitive skills of comprehension, "application", analysis, synthesis and evaluation. Bloom et al. (1956) does not label any of the two categories as lower order thinking or higher order thinking. However, the description of the broad categories point towards the recall of knowledge as lower order thinking and intellectual abilities and skills as higher order thinking ranging from comprehension to evaluation. Bloom's taxonomy categorises assessment items into six hierarchical categories from least complex "knowledge" to most complex "evaluation".

The group that developed the taxonomy felt that it could facilitate the promotion of exchange of test materials and stimulate research on examinations in education (Bloom et al., 1956). The taxonomy was the first to categorise educational objectives to measure cognitive demand in an effort to quantify content standards. It brought to the surface the notion of meaningful understanding of knowledge tested in assessment tasks. Crowe (2012) argues that the taxonomy assumed a hierarchical order from "knowledge" to "synthesis" without empirical evidence. Krietzer and Madaus (1994) point out that "synthesis" involves "evaluation" and is more complex than "evaluation". I am of the view that the users of the taxonomy would have to be trained in order to use it effectively and adapt it to be applied meaningfully to different subject areas. The taxonomy gained credibility as its development was the result of the efforts of a group of more than twenty college and university examiners. The value that Bloom's taxonomy brought to education and assessment can be seen in its use across various countries, its translation into more than twenty languages (Anderson et al., 2001), and through its more than half a century of existence. It continues to exist, in part in the revised taxonomy and forms the basis of newly developed taxonomies to measure cognitive demand.

Some researchers modified and renamed Bloom's and the revised Bloom's taxonomy to classify cognitive demand. The cognitive domain of the TIMSS 2003 science taxonomy (Reddy, 2006), which appears to be a modification of Bloom's taxonomy, consists of three categories: factual knowledge; conceptual understanding; and reasoning and analysis. The TIMSS mathematics taxonomy (Reddy, 2006) has the following categories: knowing facts and procedures; using concepts; solving routine problems and reasoning. Unlike the Bloom's taxonomy that groups the "application" category to include familiar situations and new situations, the TIMSS mathematics taxonomy separates routine (familiar) from new situations where higher order reasoning is demanded. Umalusi (2009) used a conceptual tool based on Bloom's taxonomy to analyse the cognitive demand of the physical science examination papers.

Crowe (2012) reviewed 13 different classificatory systems in order to classify cognitive demand of biology examination questions. Unable to get unambiguous groups of performance expectations from existing taxonomies, she eventually devised the performance expectations taxonomy as an instrument to measure the cognitive demand of biology examination questions. She made a purposeful selection of experts who had extensive expertise in grade 12 level biology, guided them through the analysis of the question paper selected and had post analysis discussions with the expert raters. Inter-rater agreements were established and she proceeded to analyse all the other question papers selected for analysis. The performance expectations taxonomy appeared to give her an indication of the cognitive demand of the senior certificate biology examinations.

Amer (2006) notes that behaviourist theories extensively influenced school curriculum and instruction at the time of publication of Bloom's (1956) taxonomy. He goes on to argue that several theories and approaches to learning (e.g. Constructivism, Metacognition and Self-regulated learning) have since influenced educational practices and this, together with criticisms pointing out weaknesses, prompted a review of the original taxonomy.

c) The revised Bloom's taxonomy

More than four decades after the publication of the original taxonomy, a group of cognitive psychologists, curriculum and instructional researchers, and testing and assessment specialists aimed to address the weaknesses in Bloom et al., (1956) taxonomy by revising it (Anderson et al., 2001). The revised Bloom's taxonomy built on the strengths of the original taxonomy. Amer (2006) states that in contrast to the original taxonomy, the most notable change in the revised taxonomy is the move from one dimension to two dimensions: the knowledge dimension and the cognitive dimension. The knowledge dimension consists of factual, conceptual, procedural and meta-cognitive knowledge while the cognitive process dimension consists of the categories: remember, understand, apply, analyse, evaluate and create. Cognitive complexity ranges from low in "remember" to high in "create" (Näsström, 2009). The revised Bloom's taxonomy acknowledges that “knowledge” is part of all levels of the cognitive processes and vice-versa (Anderson et al., 2001).

With reference to the cognitive dimension, Amer (2006) highlights the following changes in the revised Bloom’s taxonomy. Although the six categories were retained, three of the categories were renamed as follows: knowledge to remember, comprehension to understand and synthesis to create. In addition, the order of synthesis (create)/evaluation was interchanged as Anderson et al. (2001) believed that induction which involved creating was more complex than deduction which involved breaking a whole into its constituent parts. The categories no longer formed a cumulative hierarchy (that is, each classification includes the skills and abilities which are lower than it) but followed an order of increasing cognitive complexity. Furthermore, the categories were allowed to overlap, unlike the original taxonomy. The revised taxonomy provides a framework within which teachers can model the way they teach and examine. Commenting on the revised Bloom’s taxonomy, McMillan (2011) argues that while a revised taxonomy was needed, the revised version is more complicated for teachers to work with and may not be practical. While new taxonomies continue to be developed, the value and significance that Bloom et al. (1956) contributed to education and assessment cannot be underestimated.

d) Taxonomies used by the Department of Education

The Department of Education (2008a, p. 18) has recommended the use of Bloom’s taxonomy in the NCS subject assessment guidelines for physical sciences for the setting of examinations and control tests. Despite the weaknesses highlighted by researchers and a revised Bloom’s taxonomy published seven years earlier, the Department of Education (2008a) has made use of the 1956 version of Bloom’s with superficial modifications. The taxonomy used in the NSC assessment policy is called the physical sciences assessment taxonomy [Appendix D] and describes the six categories and levels of the cognitive domain as “recall” (level 1), “comprehension” (level 2), “applying” and “analysing” (level 3), and “evaluating” and “creating” (level 4). It also gives an explanation of each category, the skills demonstrated and action verbs. The Department of Education (2008a) does not discuss examples to guide educators on how to classify questions. The noticeable changes from Bloom’s taxonomy made in the categories follow below:

The “knowledge” category is termed as “recall”. It nevertheless still involves the cognitive process of remembering factual information. The “application” category description in Bloom’s taxonomy mentions the use of “abstractions in particular and concrete situations” (Bloom et al., 1956, p. 205), while the Department of Education (2008a, p. 18) describes this category as “the ability to use (or apply) knowledge and skills in new situations”. From my understanding, this would imply that the DBE refers to a situation has not been encountered before and therefore routine problems, similar to those in class exercises standard text books and past papers cannot fit in this category. However, from the DBE analysis of the November 2011 physical science national examination paper one, it appears that the DBE has included the application of knowledge to familiar situations in this category (Department of Basic Education, 2011e). Furthermore, “evaluation” ranks higher than “synthesis” in this version despite the reverse accepted by assessment authorities like Krietzer and Madaus (1994) and Anderson et al. (2001). The Department of Education (2008a) NSC assessment policy is applicable from 2008 to 2013.

The Department of Basic Education (2011a) has made changes to the NCS in the curriculum and assessment policy statements. This policy is applicable for grade 12 from 2014 onwards. The taxonomy used in this policy document is the revised Bloom’s taxonomy but called the physical sciences assessment taxonomy [Appendix E]. It describes the six categories and levels of the cognitive domain as remembering (level 1), understanding (level 2), applying and analysing (level 3), and “evaluating” and “creating” (level 4). It also gives an explanation of each category, the skills demonstrated and action verbs. The cognitive weightings in the Department of Basic Education (2011a) still cling on to terminology used in the original Bloom’s taxonomy. The weighting has changed with “comprehension” increasing by 5% and “application and analysis” decreasing by 5% in paper one while the levels for paper two remain the same as the Department of Education (2008a) publication.

e) Summary of taxonomies

Table 2.1 below provides a summary of the taxonomies surveyed in this study. It describes the cognitive domain categories of the various taxonomies and shows how they compare with the other taxonomies. It also separates the categories into higher order and lower order thinking skills and provides cognitive levels that are characteristic of some taxonomies.

Table 2.1
Summary of taxonomies

1) BLOOM'S TAXONOMY (Bloom et al., 1956)

Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation
Lower order thinking skills-------------------------›Higher order thinking skills					

2) REVISED BLOOM'S TAXONOMY (Anderson et al., 2001)

Remember	Understand	Apply	Analyse	Evaluate	Create
Lower order thinking skills-------------------------›Higher order thinking skills					

3) TIMSS MATHEMATICS TAXONOMY (Reddy, 2006)

Knowing facts and procedures	Using concepts	Solving routine problems	Reasoning
Lower order thinking skills-----------------------------------›Higher order thinking skills			

4) TIMSS SCIENCE TAXONOMY (Reddy, 2006)

Factual knowledge	Conceptual understanding	Reasoning and analysis
Lower order thinking skills---------------------›Higher order thinking skills		

5) THE NEW TAXONOMY OF EDUCATIONAL OBJECTIVES (Marzano & Kendall, 2007)

Retrieval	Comprehension	Analysis	Knowledge utilization	Metacognitive system	Self-system
Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
Lower order thinking skills--------------›Higher order thinking skills					

6) PHYSICAL SCIENCES ASSESSMENT TAXONOMY (Department of Education, 2008a)

Recall	Comprehension	Application	Analysis	Synthesis	Evaluation
Level 1	Level 2	Level 3		Level 4	
Lower order thinking skills-----------------›Higher order thinking skills					

7) UMALUSI PHYSICAL SCIENCES TAXONOMY (Umalusi, 2009)

Remember factual Knowledge	Understand conceptual knowledge	Problem solving
Lower order thinking skills-------------------------›Higher order thinking skills		

8) PHYSICAL SCIENCES ASSESSMENT TAXONOMY (Department of Basic Education, 2011a)

Remembering	Understanding	Applying	Analysing	Evaluating	Creating
Level 1	Level 2	Level 3		Level 4	
Lower order thinking skills-------------------------›Higher order thinking skills					

9) PERFORMANCE EXPECTATION TAXONOMY (Crowe, 2012)

Memorize	Perform routine procedures	Explain	Analyse	Apply
Lower order thinking skills			Higher order thinking skills	

10) TAXONOMY OF INTRODUCTORY PHYSICS PROBLEMS (Teodorescu et al., 2013)

Retrieval	Comprehension	Analysis	Knowledge utilisation
Level 1	Level 2	Level 3	Level 4
Lower order thinking skills------------›Higher order thinking skills			

f) Approach towards measuring the cognitive demand

Schraw and Robinson (2011) maintain that in practice cognitive demand is difficult to measure consistently. I examined and noted some of the existing taxonomies and associated arguments by scholars in measuring cognitive demand. Bloom et al. (1956) mentions that in classification, that which needs to be classified must be clearly stated, the description of categories must be clear and there must be common understanding amongst users about the classification of cognitive demand according to the taxonomy. The studies surveyed showed that experienced raters (subject experts) are required to ensure consistency in the classification of assessment items.

Edwards (2010) analysed the alignment between the curriculum content and the South African grade 12 physical sciences 2008 exemplar paper, 2008 and 2009 final examination question papers. He used the revised Bloom's taxonomy to classify the curriculum content and examination items in terms of cognitive levels. Two experienced coders with a combined experience of more than 30 years were used in Edwards' (2010) study and inter-rater reliability co-efficients of 0.88 and 0.92 were obtained for the final physics and chemistry papers respectively. Differences in the classification by the raters were resolved through discussion.

Liu et al. (2008) compared the alignment among the physics papers of three education systems: Jiansu (China), New York State (USA), and Singapore. They used the revised Bloom's taxonomy to classify assessment items in terms of cognitive levels. Two raters were used to code the examination questions and an inter-rater reliability co-efficient was calculated in each case. The final results were resolved through discussion. Initially a 70% inter-rater agreement was established with the coding of the Singapore examination in terms of Bloom's cognitive levels. One of the raters had about three years teaching experience while the other had more than fifteen years of teaching experience. The two researchers then recoded the examination questions and the inter-rater co-efficient was found to be 0.83. The final classification was done with a third researcher who has a PhD and is familiar with Singapore school curriculum.

Umalusi (2009) indicate the difficulties experienced in classification using a modification of the revised Bloom's taxonomy to measure cognitive demand. Their study does not reveal any information regarding training the users of the taxonomy before its application. The researchers acknowledged problems with the analysis that made achieving an acceptable inter-rater agreement difficult. In their study, discrepancies that arose were resolved through discussion. Näsström (2009) found that the panel of assessment experts were more consistent in their interpretation of standards than the panel composed of teachers. Crowe (2012) argues that if judgements about cognitive demand are to be made there needs to be "consistency between raters and reliability within a particular rater" (p. 123).

The above discussions point to the categories of the measuring instrument to be clearly defined, unambiguous and well understood by its users. Raters need to be well trained, experts in their field and familiar with the assessment items that need to be classified in terms of cognitive demand. Furthermore, an acceptable inter-rater reliability must be calculated to ensure that the classification is reliable.

2.2.3 Higher order thinking skills.

The degree to which learners are cognitively challenged in the type of teaching and assessment that they experience at school prepares them to cope with challenges at tertiary institutions and the world of work. According to Pithers and Soden (2000), most employers prefer to have workers who are able to think for themselves and governments are putting pressure on education sectors to prepare self-thinking individuals. wa Kivilu (2006) states that "Higher Education institutions and employers in South Africa complain of the low level of skills of learners graduating from grade 12" (p. 34). Jansen (2012) concurs with this view and goes further to state that employers in business and industry regard today's graduates as weak in basic reasoning and problem solving with a huge divide between the graduates' certificate and competence. Mhlolo (2011) supports the view that higher order thinking skills and knowledge are more desirable for economic productivity.

According to Lewis and Smith (1993),

> "higher order thinking occurs when a person takes new information and information stored in the memory and interrelates and/or rearranges and extends this information to achieve a purpose or find possible answers in perplexing situations. These will include: deciding what

> to believe; deciding what to do; creating a new idea, a new object, or an artistic expression; making a prediction; and solving a non-routine problem" (p. 136).

This study will use the term higher order thinking proposed by Lewis and Smith (1993) as an all-encompassing term to include critical thinking, problem solving, higher order reasoning, decision making and creative thinking. Critical thinking experts describe the ideal critical thinker as an individual who has "a probing inquisitiveness, a keenness of mind, a zealous dedication to reason, and a hunger or eagerness for reliable information" (Facione, 2013, p. 10). The experts go on to say that s/he is not only characterised by her/his cognitive skills but also by how they approach life and living in general. Facione (1990) states that

> the critical thinking movement have argued that effective and meaningful education requires that curricular, pedagogical and assessment strategies at all levels of education be coordinated so as to foster in students those cognitive skills and habits of inquiry associated with critical thinking (p. 1).

Crowe's PhD study (2012) reports that higher order cognitive skills featured to a small extent in the senior certificate biology examinations from 1994 to 2007 and that none of the examinations approached the recommended target of higher order thinking skills. She argues that while core biology syllabus in that study period required that specified percentages be devoted to the testing of higher intellectual abilities it gave no indication of exactly what was meant by higher intellectual skills. Crowe (2012) questions whether learners were given sufficient opportunity to demonstrate their higher order thinking abilities. Umalusi's Maintaining Standards Report show that the 2008 NSC physical sciences examination comprised of 12% factual knowledge, 37% conceptual knowledge and 50% problem solving in comparison to the 2005 to 2007 matric higher grade papers which showed 12%, 30%, and 58% for the respective cognitive categories (Umalusi, 2009). On face value, this indicates that the cognitive demand on problem solving area has declined by 8% in the 2008 NSC paper.

The understanding of the meaning of problem solving is often contested. Anderson et al. (2001), Bloom et al. (1956), the Department of Basic Education (2011a), the Department of Education (2008a), TIMSS science taxonomy (Reddy, 2006) and Umalusi (2009) do not separate routine and non-routine problems in their categorisation of cognitive demand. Lewis and Smith (1993) suggest that there is general agreement that lower and higher order thinking skills can be distinguished but the learner's history may mean that a given situation that requires higher order thinking by one learner may need only lower order thinking by another. Thus for learners not exposed to certain types of routine problems, a greater cognitive demand is required to answer these "unfamiliar" questions than those learners who have been through similar problems. Hobden (2002) maintains that a task cannot be given the status of a problem task if it does not initiate some form of critical thinking. He goes on to suggest the following extended definition to capture the understanding of problem solving:

> Problem-solving is a multifaceted cognitive activity in which we engage when we are confronted with a task in which routine action or normal thinking does not allow one to go from the given existing situation to the desired goal situation, but rather there is recourse to some form of critical thinking. Such critical thinking has the task of devising some action, which may overcome the perceived barrier between the existing and goal situations. (p. 5)

Thus an item is considered as a "routine execution" and not "problem solving" from the examiner's perspective if it is familiar and similar to those set in standard learner text books, past examination papers and clearly defined in the NCS syllabus and examination guidelines. The prepared learner uses the information given to solve the familiar problem with the routine execution of an algorithm or procedure. I do concur with the views of Lewis and Smith (1993) that to learners not familiar with certain approaches / techniques to solve problems, the use of higher order reasoning to figure out a solution may prevail, while to other learners this may be somewhat of a routine exercise as similar items could have been worked out previously. Looney (2009) cites studies which have shown that students in programmes featuring teaching of higher-order skills perform as well, or outperform peers in programmes focused on test preparation, or transmission modes of teaching. Vinjevold (2005) points out that "the cognitive demand of exit-level assessment tasks must be raised and ways must be found of assessing a wide range of skills" (p. 2). The discussions tend to point towards assessment items being seen as routine by some people and higher order by others not exposed to that type of testing. High stakes examinations like the NSC thus cannot be taken at face value as the context is an important consideration when determining if an assessment item is higher order or not. The need to provide adequate teaching and assessment opportunities for learners to engage in higher order thinking is nonetheless important to enable them to become economically productive employees and self-thinking individuals.

2.3 ALIGNMENT STUDIES

Alignment studies feature prominently in the literature on curriculum and assessment reform. Alignment is the match or agreement between two categories (Squires, 2009). Webb (1997) defines alignment as "the degree to which expectations and assessments are in agreement" (p. 4). Resnick et al. (2011) claims that the success of standards based education system depends on strong standards, and assessments that measure what the standards expect.

Liu et al. (2008) notes that when a teacher interprets and implements the intended content standard in terms of the content and cognitive emphasis in the same way as the standardized tests, students are more likely to do well in the tests. The results of Liu et al. (2008) indicate that different education systems have different emphasis on topics as well as cognitive skills. They used the revised Bloom's taxonomy and showed that the New York State 2006 physics examination does not go beyond the cognitive skill of "apply" while there is a shift in Chinese and Singapore examination to de-emphasise "remember" skill and emphasise the "analyse" cognitive skill. The Singapore examination goes further to include the "evaluate" and "create" cognitive skills in the 2006 physics examination.

Edwards (2010) stresses the importance of alignment studies especially in a developing country with educational disparities like South Africa. He argues that if there is no alignment to what is set out in the curriculum and what is tested; schools may ignore the assessment standards. He emphasises a quality relationship between curriculum and assessment. Edwards (2010) claims that his alignment study is the first in the South African context and recommends examining the cognitive complexity in physical sciences as a further area of research. Edwards (2010) used the revised Bloom's taxonomy and found that the cognitive levels of the core knowledge and concepts had over-emphasised "remember" (38.9%), "understand" (29.9%) and "apply" (26.4%) in the physics curriculum with 4.8% remaining for the higher order cognitive categories. A similar picture for the grade 12 chemistry curriculum shows the first three categories totalling

96.3% with the remaining 3.7% of the curriculum catering for the higher order categories of “analyse”, “evaluate” and “create”. These findings indicate that the curriculum emphasises low cognitive demand. The examination paper analysis by Edwards (2010) found that the “apply” category was over-represented for the 2008 and 2009 final examination papers.

Mhlolo’s PhD study (2011) on the alignment of the written, tested and taught curriculum in NSC mathematics summarised the most important points of alignment studies as follows:

(a) Alignment is more powerful in predicting student achievement than race, gender or socioeconomic status.

(b) Curriculum alignment was more important for low aptitude students than for high aptitude students with low aptitude students making greater gains when alignment was present.

(c) This alignment should be anchored in cognitively demanding performance standards.

(d) When anchored in cognitively demanding standards curriculum alignment had the potential to reduce the achievement gap between the previously advantaged and previously disadvantaged learners.

(e) If we wanted students to develop the capacity to think, reason and problem solve then classroom practices need to be designed to give students opportunities to learn cognitively complex content.

(f) It was possible to use curriculum alignment as an emancipatory approach to address inequalities in the educational needs of previously disadvantaged learners (p. 38).

Mhlolo’s (2011) findings indicate that while higher order cognitive skills and processes are consistently emphasised in the curriculum documents, the 2008 examination papers emphasised lower order cognitive skills.

Resnick et al. (2011) point out that alignment is a particularly critical aspect of test validity. Leighton (2013) emphasises that the methods to validate test score inferences have become a topic of concern for researchers and practitioners as interest with using standardised achievement tests to assess students’ mastery of knowledge and skills is growing. Authors like Mhlolo (2011), Edwards (2010), Squires (2009) and Liu et al.

(2008) have emphasised the importance of alignment between the written, taught and tested curriculum and suggest that the alignment be anchored in cognitively demanding assessment standards. However, my literature search has revealed that very little is done on the area of actual cognitive demand analysis especially with high stakes examinations like the NSC.

2.4 CHAPTER SUMMARY

The world of work and higher education institutions are increasingly demanding individuals who have higher order thinking skills but schools appear to be failing to satisfy this demand adequately. While the DBE percentage pass rate of physical sciences has improved from the first NSC examination in 2008 to the November 2013 examination, the improved percentages does not reveal whether the levels of cognitive demand that were set out in the NCS are achieved. This is vital in maintaining standards and has implications on the entrance to universities, other centres of higher learning and the world of work. This review has pointed out that scholars are of the view that grade inflation does not give learners a true indication as to the skills and knowledge that they possess. The review has also indicated that while the pass rates may have risen, actual standards may have fallen. The usefulness of cognitive demand taxonomies featured across many studies surveyed was examined in this review. Large-scale studies like TIMMS give us an idea of how we compare with other countries in mathematics and science and may stimulate educational reform towards more cognitively demanding assessments which in turn will drive teaching from transmission techniques to emphasising higher order thinking skills. Studies reviewed have indicated that teaching higher order thinking fused into everyday learning may lead to better performances and economic productivity. From the literature surveyed, the study of cognitive demand in physical sciences examinations appears under-researched. Edwards (2010) recommended that more needs to be done to understand specifically the cognitive demand levels and the alignment of the written, taught and tested curriculum. This study aims to determine the cognitive demand of the NSC physical sciences examination and examine how it compares with the recommended weightings given in the NCS. The next chapter describes the research design and the development of the cognitive demand taxonomy used in this study.

Chapter 3

RESEARCH DESIGN

In this chapter the research strategy in the study is first discussed while the details of research methods employed in this study are described in the data collection and analysis. The ethical concerns of the study are then discussed followed by a detailed description of the development of the cognitive demand taxonomy.

3.1 RESEARCH STRATEGY

Researchers operate in different paradigms or research orientations. A post-positivist orientation to research was the approach used in this study. Guba and Lincoln (1994) view a paradigm as a set of basic beliefs that represents a worldview that defines for the individual the nature of the world and the individual's place in it. Maree (2007) states that post-positivists acknowledge, that there is a reality independent of our thinking that science can study and that all observation and measurement is fallible and all theory is reversible. Mertens (2005) emphasise that in this paradigm, reality is knowable within a specific level of probability. Cohen, Manion, and Morrison (2011) maintain that while post-positivists argue for an objective reality, they challenge the positivist approach and adopt a pluralist view of changing knowledge of the world. As a researcher, I will approach this study with the belief that there is an independent reality that we can understand to a degree and with more evidence our understanding improves; however, we can at best have a partial understanding of this independent reality.

Creswell (2003) maintains that the knowledge that develops through a post-positivistic lens is based on observation and measurement. As a post positivist, I am of the view that the most appropriate way to know the cognitive demand of the examination papers is to measure it. This required a quantitative analysis of examination question papers using a suitable taxonomy as my instrument to classify questions into different cognitive domains. Mertens (2005) maintains that although qualitative methods may be used within the post-positivist paradigm, quantitative methods tend to be predominant.

This study uses a quantitative, non-experimental design and categorises assessment items using a cognitive demand taxonomy as its instrument.

The purpose of this section is to describe the research strategy employed in this study. Maree (2007) states that quantitative studies are either descriptive, which show trends or experimental establishing probable causality. Cohen et al. (2011) outlines three types of reasoning that people use to comprehend the world. These are listed as inductive, deductive and a combined inductive-deductive approach. Inductive reasoning assumes that the study of a number of individual cases would lead to a hypothesis and eventually lead to a generalisation. The deductive approach assumes the formal steps of logic that can be deduced from a valid premise while the inductive-deductive approach combines both these approaches. The predominant reasoning in this study is deductive as I have analysed examination question papers using a cognitive demand taxonomy. The result of this analysis is a detailed analytical description of the cognitive demand of the papers presented in chapter four of this dissertation.

The analysis of grade 12 final physical sciences examination question papers was conducted for the period from 2008 to 2013 using a taxonomy validated by four subject experts. The analysis of the examination question papers involved the categorisation of the content standards in terms of cognitive demand. Each examination question paper was divided into the smallest unit that could be scored and is termed an assessment item in this study. Each assessment item was measured against the taxonomy for the assessment of science skills and knowledge (TASSK) and categorised as cognitive level one to five representing the cognitive domains of “retrieval”, “basic comprehension”, “routine execution”, “application” and “analysis and reasoning”. Similar approaches were used in the studies by Crowe (2012), Edwards (2010), Liu et al. (2008), Näsström (2009) and Umalusi (2009) for measuring the cognitive demand of examination papers which I have discussed in the previous chapter.

3.2 VALIDITY AND RELIABILITY

For research to be effective, an essential element to consider is validity. Validity in quantitative research is regarded as essentially a “demonstration that a particular instrument in fact measures what it purports to measure" (Cohen et al., 2011, p. 179). Maree (2007) describes content validity as the extent to which the instrument covers the content area to be measured. This study sets out to measure cognitive demand and covers the cognitive areas of "retrieval", "basic comprehension", "routine execution", "application" and "analysis and reasoning". The analysis instrument in this study was validated by the experts before the examination papers could be analysed by the researcher. In this study, validity was ensured by the employment of a quantitative deductive analysis design, utilizing a suitable cognitive demand taxonomy, by the appropriate use for data analysis and by using an effective sampling strategy (Cohen et al., 2011).

Maree (2007) describes reliability as the extent to which a measuring instrument is repeatable and consistent. In this study, this can be gauged by the analysis of the five raters (four experts and the researcher). The inter-rater reliability among the five raters in the pilot analysis was calculated and found to be 0.75. The discussion and calculation on inter-rater reliability is explained in the next chapter. Cohen et al. (2011) suggest that reliability can be improved by training of raters to ensure consistency. Therefore discussions with the experts and refinements to the taxonomy followed this trial run. Thereafter, I together with the four experts analysed the November 2012 physical science examination question papers (Department of Basic Education, 2012a; 2012b). The inter-rater reliability among the five raters was 0.87 for paper one and 0.85 for paper two. Further discussions took place to finalise the categorisation. Once the taxonomy was found to be reliable by the experts and having taken into consideration all the discussions with the experts, I proceeded to analyse all the NSC examination papers from 2008 to 2013 using the validated taxonomy to ensure consistency in the classification.

3.3 DATA COLLECTION STRATEGY

The aim of this study was to find answers to the research questions:

1. What are the cognitive demands of the NSC physical sciences examination question papers?

2. How do the cognitive demands of the NSC examination papers compare with the weighting of cognitive levels recommended in the National Curriculum Statement?

This section details the sources of data and the methods employed to collect the data in order to answer the research questions. As discussed in chapter one, the senior certificate examination prior to 2008 was written under various departments and was based on the Nated 550 syllabus. From 2014 onwards, the NSC examinations will be based on the CAPS (Department of Basic Education, 2011a). This study focuses on an era of NSC examinations from 2008 to 2013 based on the NCS (Department of Education, 2003).

Cohen et al. (2011) argue that the quality of research is determined not only by the methodology and instrumentation, but also by the sampling strategy. Maree (2007) claims that it is usually impossible to include the entire population in one’s study because of time and cost. To this end, this study did not include all senior certificate examinations ever written in physical sciences; however, the population of final NSC examination question papers in this era (2008 to 2013) of examinations was analysed.

The data was collected to analyse the physical sciences examination question papers in terms of cognitive demand. The source of data in this study has been extracted from the Department of Education (2008b, 2008c) and the Department of Basic Education (2009a 2009b, 2010a, 2010b 2011c, 2011d, 2012a, 2012b, 2013c, 2013d) NSC physical sciences final examination question papers. There were six physics (physical sciences paper one) question papers and six chemistry question papers (physical sciences paper two) in the English language that were used in this study. These 12 question papers were obtained from copies of question papers collected by the researcher and from the Department of Basic Education website.

The team of four subject experts together with the researcher categorised the assessment items from November 2012 paper one and paper two independently to develop and verify the reliability of the TASSK cognitive demand typology. The details of the analysis are described in section 4.1 of the next chapter. Once the taxonomy was considered reliable, all the examination question papers from 2008 to 2013 were analysed by the researcher using the TASSK cognitive demand typology. Before starting with the analysis the researcher read through the entire question to get a sound understanding of the question before proceeding with the analysis. Some questions required more time than others and the researcher had to discuss certain questions with some of the subject experts. Some questions required the researcher to go back to the analysed papers to ensure that analysis was consistent for that type of question. The researcher analysed all the physics examination papers and then all chemistry examination papers. Each paper took an average of ninety minutes to analyse. When all the examination questions were analysed a final check was done by the researcher and where there were queries it was discussed and resolved with the science education professor who was part of the subject expert group. The analysis of the question papers was used to answer my research question one. The findings were then used to draw comparisons to answer research question two. I chose to compare the examination analysis with the DBE recommended levels to see whether what was tested in the NSC examination was aligned to what was recommended in the NCS.

3.4 ETHICAL ISSUES

Researchers need to have ethical considerations in their pursuit to find answers to their research problems. Cohen et al. (2011) states that, researchers face an ethical dilemma trying to strike a balance between the pursuit of the truth and the rights and values of their subjects. Approval to conduct this study was granted by the Human and Social Sciences Ethics Committee [Appendix A]. The expert group were not subjects in the study. Their purpose was to refine and validate the study taxonomy by providing their expert view on the work of the researcher. They offered their assistance to analyse the examination papers in their personal capacity as expert science educators voluntarily and not as members of any organisation or department of education employees. This was done outside their working hours (private time) and at their private residence. Their participation in this study had no impact on their official duties. No permission was required from authorities and gate-keepers as no data was collected at any educational institution and no minor children were used in this study. All data (National Senior Certificate examination question papers) used in the study is in the public domain. All experts signed an informed consent form [Appendix H] undertaking that:

- their participation in this study was voluntary
- their responses would be treated in a confidential manner and no limits apply
- their anonymity would be ensured
- a decision not to participate in this study would not result in any negative or undesirable consequences to them.

- a decision to participate allowed them freedom to withdraw from the research at any stage for any reason and without any negative or undesirable consequences to them.
- the findings of the study would be made available to them on request.

3.5 THE DEVELOPMENT OF THE STUDY TAXONOMY

This section shows the development of the cognitive demand taxonomy used in the study. In order to answer my research questions, I first needed to use an appropriate taxonomy to measure the cognitive demand of the NSC physical sciences examination questions. The literature surveyed in chapter two indicates that the revised Bloom's taxonomy (Anderson et al., 2001) served as a useful instrument to analyse cognitive demand. The DBE also recommends the use of versions of Bloom's taxonomy (Department of Basic Education, 2011a; Department of Education, 2008a). I initially decided to use the Department of Basic Education (2011a) physical sciences assessment taxonomy [Appendix E] which is a version of the revised Bloom's taxonomy [Appendix C] modified by the DBE to determine the cognitive demand of the physical sciences question papers.

3.5.1 Measuring the cognitive demand of assessment items.

As I proceeded with my initial attempts at analysis, I began to encounter problems with classification of some of the assessment items. Because of my teacher training in Bloom's taxonomy and my assessment experience using this taxonomy, I applied the taxonomy without questioning it in the past. For the first time after my initial attempts at analysis in this study, I began to question the appropriateness of this version of the revised Bloom's taxonomy to classify physical sciences examination questions. I have provided my comments on the initial attempts at analysis using this version of the revised Bloom's taxonomy modified by Department of Basic Education (2011a) to classify physical sciences examination questions.

The "remember" category (cognitive level one) was fairly easy to classify the recall of items. Some complex factual questions initially posed problems but were eventually resolved. Apart from “remember”, every cognitive process in my view requires various degrees of understanding and meaningful learning. This taxonomy appears to use the “understand” category to classify questions that require a basic form of understanding and a little more thinking than recall of knowledge. Therefore the use of the term “understand” appears to be too broad a category and it does not assist in classifying questions. Bloom’s revised taxonomy of “apply” combines “application” to include both familiar and new situations. Application to new situations is more cognitively demanding than “application” to familiar situations, yet Bloom’s revised taxonomy places both these different processes in the same cognitive domain. To illustrate this, consider the example of the momentum question (Department of Basic Education, 2012a) also discussed in section **3.5.6**c.

> *A car of mass m travelling at a velocity of 20 m.s^{-1} east on a straight level road and a truck of mass 2m travelling at 20 m.s^{-1} west on the same road. The vehicles collide head-on and stick together during the collision. Ignore the effects of friction.*
> *Calculate the velocity of the truck-car system immediately after the collision.*

If the revised Bloom’s taxonomy or the Department of Basic Education (2011a) version is used, this question would fall in the “application” category. This is a familiar question that is clearly described in the syllabus (NCS), appears in this form in standard text books and is routinely asked in past examinations. The prepared learner is not applying skills or making any new connections and therefore not engaging in productive thinking. If this question was tested in this form for the very first time requiring the prepared learner to make new connections by adding the two masses and solving for the unknown final velocity, the revised Bloom’s classification will still consider it “application”. Thus the revised Bloom’s taxonomy and the DBE taxonomy (Department of Basic Education, 2011a) do not group these two distinct thinking processes differently and consider it one cognitive category. The combining of the “applying” and “analysing” (cognitive level three) categories further makes this a very broad category used by the DBE. It encompasses routine questions, application to new situations and higher order reasoning. I interpret these as three distinct cognitive processes uncomfortably boxed into one cognitive level by the DBE.

The Department of Basic Education (2011a) also combines Bloom’s revised taxonomy categories of “evaluate” and “create” (cognitive level four). Evaluate is defined as “making judgements based on criteria and standards” (Anderson et al., p. 83) while create “involves putting elements together to form a coherent or functional whole”

(Anderson et al., p. 85). The description of these categories does not appear to match any physical sciences NSC examination questions. From my experience many teachers and examiners place some questions in these categories if they consider them complicated or difficult but not actually having taken into account what “evaluate” and “create” actually describe.

The physical sciences assessment taxonomy (Department of Basic Education, 2011a), Bloom et al. (1956) and Bloom’s revised taxonomy (Anderson et al., 2001) do not distinguish between lower order and higher order thinking in their taxonomies and appear to have limitations with regard to the classification of physical sciences examination questions. The problems discussed above made it difficult to measure the cognitive demand by applying these taxonomies. Noting all the difficulty experienced, as I categorised the examination questions, I began to modify the taxonomy for the analysis of the NSC physical sciences assessment items in terms of cognitive demand to accommodate the issues raised. In attempting to create a taxonomy that can help me and other science teachers analyse the cognitive demand of assessment items, I had to ensure that it is practical and easy to use by science teachers. Such a taxonomy must also separate lower order and higher order thinking and separate each cognitive category to reflect the type of thinking demand distinctively. I created a draft taxonomy initially with cognitive domain categories that became definitive as it developed. The development of the TASSK cognitive demand typology, described in this section, was a process with many modifications as it progressed.

3.5.2 Subject experts.

To assist me with the development of the taxonomy, I required a team of physical sciences experts. Häussler and Hoffmann as cited by Crow (2012) “recommended that experts be individuals who are reflective of their professional field, are open to the opinions of other people, and are actively involved in promoting teaching and learning” (p. 150). The studies surveyed do not describe the experts in great detail with regard to qualifications and number of years experience in marking exit examinations other than to say that the experts have extensive teaching experience in their subject. The experts in this study were not subjects or participants and were utilized for their expertise in grade 12 physical sciences and NSC examinations. A group of 12 subject experts from the science education community voluntarily engaged in discussions with me to give their expert view during the development of the taxonomy. The number of experts in similar studies range from 2 to 27 (Näsström, 2009). Communication with these experts was done face to face, via telephone and via email. All of these experts each have more than twenty years of involvement in the area of physical sciences. Their experience includes teaching, lecturing, marking, examining and advising in physical sciences. Among these experts, one is a science education professor, three are former examiners, two are subject advisors, two are science education lecturers and four are school teachers currently involved in grade 12 physical sciences teaching and the marking of the NSC examinations.

I discussed my ideas regarding the creation of a suitable taxonomy for science skills and knowledge with the experts before embarking on the development of the cognitive demand taxonomy. When I completed the first draft, I emailed it for comments to the experts. I received feedback from some of the experts and adjusted the taxonomy after considering the inputs. Discussions with the experts were ongoing for subsequent drafts. Some of the discussions on the taxonomy follow. The experts advised me to describe "analysis and reasoning" in greater detail. One expert advised me to put in cognitive levels so as to compare with the DBE levels. Another advised to describe a hierarchy of complexity. I incorporated some of the suggestions in my second draft. My second draft had more detailed descriptions of the categories, a complexity continuum and cognitive levels one to four. My third draft had more detail per cognitive category as suggested by the experts. It was suggested that I separate reproductive and productive cognitive areas and work on the renaming and grouping of cognitive levels 3A and 3B which represented "routine application" and "extended application". One expert suggested that I remove complexity descriptors like elementary, basic, medium, slightly complex and very challenging and just keep to the category name. I was also reminded by another expert that "application" is associated with new connections made. My fourth draft had reproductive and productive thinking clearly distinguished, cognitive categories and levels one to five with renamed categories of "routine execution" (level three) and "application" (level four). Some minor changes and reshuffling took place in draft five. Thereafter, I discussed it with the expert who is a science education professor who agreed with me that the taxonomy was ready for piloting.

I decided to pilot the taxonomy with four of the experts who had agreed to classify questions from one of the examination papers. The four subject experts (expert one, expert two, expert three and expert four) that agreed to use their subject expertise to classify examination questions have 22, 20, 20 and 28 years of experience teaching physical sciences at grade 12 level respectively. Expert one and two each have 13 years of senior certificate marking experience. Expert one has served three of these years as a senior marker while expert two has served two years as senior marker. Expert three has a total of 16 years senior certificate marking experience of which six years were served as a marker, eight years were served as senior marker and two years were served as deputy chief marker. Expert four served as a senior certificate examiner for four years. He has a total of 20 years marking experience of which 11 years were served as a marker; five years were served as a senior marker, four years as a deputy chief marker. The literature review pointed to the need for raters to be well trained, experienced and familiar with the assessment items that need to be classified. All four subject expert raters are adequately qualified both academically and professionally and are familiar with the questions set in the NSC examinations during the period of this study. They were also trained in the use of the cognitive demand taxonomy through discussion sessions.

The pilot analysis was conducted to familiarise raters with the taxonomy, gauge the level of rater consistency and to make adjustments to the taxonomy. The researcher and the four raters initially used the taxonomy created by the researcher, to analyse the March 2013 physical science examination question paper (Department of Basic Education, 2013b), to provide their expert views on its effectiveness and also verify the researcher's analysis. Discussions between the researcher and individual experts took place on the classification of certain assessment items. The discussions were aimed at improving the consistency among the raters using the taxonomy.

After a common understanding of the categories between the researcher and the raters was reached, the taxonomy was ready for validation for this study. The raters thereafter analysed the November 2012 physical science examination paper one and paper two (Department of Basic Education, 2012a; 2012b). The data they provided was used to refine and validate the taxonomy. Once the taxonomy was validated, the researcher analysed the remaining question papers. The researcher has 20 years of subject experience and ten years of senior certificate marking experience. This shows that he is also adequately qualified to proceed with the remainder of the analysis thereby maintaining a level of consistency with the categorisation. A similar approach to generating an analysis taxonomy was used by Crowe (2012) to analyse biology examination question papers.

3.5.3 Structure of the TASSK cognitive demand typology.

The Taxonomy for the Assessment of Scientific Skills and Knowledge (TASSK) [Appendix G] is one that borrows ideas from Bloom's taxonomy (Bloom et al, 1956), the revised Bloom's taxonomy (Anderson et al., 2001), the TIMMS mathematics and science taxonomies (Reddy, 2006), Marzano's new taxonomy of educational objectives (Marzano & Kendall, 2007), the Department of Education (2008a) physical sciences assessment taxonomy, Umalusi (2009) cognitive demand instrument, the Department of Basic Education (2011a) physical sciences assessment taxonomy, the performance expectations taxonomy (Crowe, 2012) and the taxonomy of introductory physics problems (Teodorescu et al., 2013).

The TASSK typology categorises the cognitive demand of assessment items ranging from lower-order thinking to higher-order thinking along a continuum. The two major categories of the TASSK typology are reproductive and productive thinking. Reproductive thinking accommodates the lower order thinking skills of "retrieval", "basic comprehension" and "routine execution" while productive thinking accommodates higher order thinking skills of "application" and "analysis and reasoning". This distinction is deliberate in the TASSK typology, unlike in the revised Bloom's taxonomy and others preceding it. This distinction is clear as the literature surveyed points to researchers emphasising the need for the teaching and testing of higher order thinking skills and employers continually demanding employees to have such skills. When the assessment items are analysed and summarised they will indicate where the NSC physical sciences examination papers have their emphasis. The individual categories of the TASSK cognitive demand typology are shown in Table 3.1 below and explained thereafter.

Table 3.1
Categories of the TASSK cognitive demand typology

REPRODUCTIVE THINKING LOWER-ORDER THINKING SKILLS			PRODUCTIVE THINKING HIGHER ORDER THINKING SKILLS	
Retrieval Level 1	Basic Comprehension Level 2	Routine Execution Level 3	Application Level 4	Analysis and Reasoning Level 5

a) *Retrieval*: Retrieval is a lower order thinking skill as the learner is only able to recall, restate, remember, recognise, identify, list, write or name: facts, terms, definitions, phenomena, units, laws, diagrams, processes, concepts, routine equations, symbols and formulae. According to Anderson et al. (2001), this involves searching the long term memory for a piece of information that is identical or very similar to the one asked for in the test item. Schofield (1972) argues that memory correlates better with intelligence than the correlation between two tests of memory. However, while memory is an important aspect of the learning process, the retrieval of information in this categorisation does not necessarily require a demonstration of understanding or reasoning by the learner. The learner may learn answers by rote and regurgitate them in the examination. Rote learning has been described by Mayer (2002) as the inability to transfer knowledge to a new situation because the knowledge has not been understood.

b) *Basic Comprehension*: The learner demonstrates a basic understanding of standard familiar scientific concepts, practical work, graphs and diagrams by identifying, combining, separating, translating, discussing, explaining or interpreting the relevant information given in the question. It caters for that area that does not involve "retrieval" but at the same time is limited to the very few connections that have to be made to elicit a response.

c) *Routine Execution*: Not solely rote, this process demands some understanding of the question before selecting a learned process to provide a response. The learner is able to conduct routine processes or procedures to answer familiar questions and solve routine problems which may involve skills and algorithms (a set of specific steps performed in a particular order). The questions that fall in this category are familiar questions clearly described in the syllabus (Department of Education, 2006) and examination guidelines (Department of Education, 2009). They are generally found in a similar form in recommended text books, class exercises or question papers of the recent past. This process does not require the prepared learners to apply their knowledge to new situations.

d) *Application*: Bloom et al. (1956) argues that the application of knowledge is important as most of what we learn is intended to be used in problem situations in life. The learner has the ability to use the knowledge and skills learned and apply it to new contexts in this cognitive process. The learner has to demonstrate an understanding of the new situation before selecting a procedure from a range of options available. The questions in this category are more cognitively complex than the routine type, as the task has not been encountered before and modifications to procedures may have to be made before implementation. Anderson et al. (2001) maintains that implementation is frequently associated with use of techniques and methods rather than skills and algorithms. The questions that fall in this category require the learner to think beyond the routine type and apply what has been learnt to answer the questions which may require logical explanations, non-routine procedures or set of multiple steps not necessarily performed in a fixed order.

e) *Analysis and Reasoning*: The learner shows a deeper understanding and uses analytical reasoning to break down the question into its component parts. Higher order thinking has to be applied to answer these unfamiliar questions which may involve:

1) Problem solving - overcoming obstacles, barriers or limiting conditions to resolve the problem.
2) Decision making - selecting between two or more alternatives
3) Experimenting - generating and testing hypotheses for the purpose of understanding phenomena.
4) Investigating, generating and testing hypotheses using logical arguments and reasoning.

(Marzano & Kendall, 2007)

3.5.4 Comparison of the TASSK typology with Bloom's, the revised Bloom's and the Performance Expectations Taxonomy (PET).

This section discusses the conceptual similarities and differences within the categories of the Bloom et al. (1956), the revised Bloom's taxonomy (Anderson et al., 2001) and Crowe's (2012) performance expectations taxonomies and the TASSK. Bloom's

(1956) and the revised Bloom's taxonomy was chosen because of its dominance in the literature. Crowe's (2012) PET was chosen since it emancipated me to step out of the traditional hypnotic influence of Bloom et al. (1956) and Anderson et al. (2001) and design a taxonomy relevant to a subject area that seeks to measure cognitive demand as accurately as possible. Table 2.1 shows the relevant taxonomies surveyed in this study and table 3.1 shows the categories of the TASSK cognitive demand typology and may assist to contextualise the discussion below.

Neither Bloom et al. (1956) nor Anderson et al. (2001) separate the lower and higher order thinking skills in their respective taxonomies distinctively. They only indicate that cognitive complexity increases from knowledge/remember to evaluation/create respectively. The PET (Crowe, 2012) and TASSK separate these two major categories. The PET (Crowe, 2012) classifies the cognitive demand categories of "memorise" "perform routine procedures" and "explain" as lower order cognitive skills while the other categories of "analyse" and "apply" are classified as higher order cognitive skills. The TASSK typology groups these major categories distinctively to signify reproductive (lower order thinking skills) and productive thinking (higher order thinking skills). The paragraphs which follow will analyse each sub-category in each of the taxonomies in a chronological order from Bloom et al. (1956), Anderson et al. (2001), and Crowe (2012) to the TASSK.

The sub-categories of "knowledge", "remember" "memorize" and "retrieval" across the taxonomies closely resemble one another and all involve the retrieval of information from memory without a learner necessarily demonstrating understanding in the response. There appears to be no overlap from this category to the next within each of the above-mentioned taxonomies. From my experience, I have noticed that some teachers and examiners tend to classify complex recall as more cognitively demanding than "retrieval". This is not correct since we are measuring cognitive demand and the information that has to be retrieved could be simple, medium or complex but nonetheless remains "retrieval".

The sub-categories of "comprehension", "understanding", "perform rote and routine procedures" and "basic comprehension" convey somewhat different meanings across the taxonomies. Bloom et al. (1956) regard "comprehension" as the lowest level of understanding and corresponds well with the TASSK category of "basic comprehension".

Anderson et al. (2001) maintain that students understand when they make connections between new and prior knowledge. This meaning places it broadly within the higher order thinking skills, leaving no category in this taxonomy to encompass the "sense making" of familiar concepts or processes beyond rote. There appear to be more similarities with categories of "explain" (Crowe, 2012), "comprehension" (Bloom et al., 1956) and "basic comprehension" in the TASSK.

The sub-categories of "application" across the taxonomies differ significantly. Regarding science skills, Bloom et al., (1956), refers to application as "the ability to apply science principles, postulates, theorems or other abstractions in new situations" (p. 124). They argue further that if the situations presented to the student are old ones (familiar), the student does not apply. This corresponds well with the TASSK but differs with the revised Bloom's taxonomy (Anderson et al., 2001) which regards "application" as referring to both familiar and new situations. Crowe (2012) also regards "application" as referring to making connections in new situations. She however views "application" as the highest cognitive skill in the PET in line with the "analysis and reasoning" category of TASSK. The TASSK typology has categorised the ability to use knowledge and skills in a new situation as "application" and the solving of familiar procedures and algorithms as "routine execution", clearly separating this lower order thinking skill from the higher order skills.

The sub-categories of "analysis" refer to the breakdown of the question into its component parts and making connections between the parts by both Bloom et al. (1956) and the revised Bloom's taxonomy (Anderson et al., 2001). The PET (Crowe, 2012) explains this category as one which requires students to "make connections, not required by the syllabus or given in the question using memorized knowledge and routine procedures in familiar contexts" (p. 137). The "analyse" category in Crowe (2012) appears to correspond with the combined "application" and "analysis" categories of Bloom et al. (1956) and the revised Bloom's taxonomy (Anderson et al., 2001). The TASSK combines "analysis and reasoning" to make connections in new situations including problem solving, decision making, investigating, generating and testing hypotheses for the purpose of understanding and using logical arguments and reasoning. Bloom et al., (1956) and Anderson et al., (2001) also include "evaluate" and "synthesis" / "create" in their taxonomies. "Evaluate" refers to the making judgements and "synthesis" / "create" involves the putting together of elements to make a coherent whole (Bloom et al., 1956;

Anderson et al., 2001). The position of these two categories is interchanged in the revised Bloom's taxonomy as "create" is considered more cognitively complex than "evaluate" as it results in the construction of an original product (Anderson et al., 2001).

3.5.5 Alignment of the TASSK with the DBE taxonomy.

In this section, I compare the TASSK with the taxonomy used by the DBE. Since a modified version of Bloom's (1956) taxonomy was used by the Department of Education (2008a) in the study period of NSC examinations, I had to first align the TASSK with the Department of Education (2008a) taxonomy before making any comparisons of results.

In order to answer my research question:
How do the cognitive demands of the NSC physical sciences examination question papers compare with the weighting of cognitive levels recommended in the national curriculum statement?,
I first had to see how the two taxonomies can best be aligned. The suggested weighting of cognitive levels for tests and examinations in the physical sciences in grades 10-12 is given below in table 3.2.

Table 3.2
Suggested weighting of the cognitive levels for examinations and control tests

Cognitive Level Description	Weighting	
	Paper 1	Paper 2
Recall (Knowledge)	15	15
Comprehension	30	40
Analysis, Application	45	35
Evaluation, Synthesis	10	10

(Department of Education, 2008a, p. 9)

The matching of the TASSK with the Department of Education (2008a) physical sciences taxonomy which uses a version of Bloom et al. (1956) is given in table 3.3 below. The "retrieval" and "basic comprehension" are equivalent to the "recall" and "comprehension" categories respectively. The TASSK typology considers "application" to refer to new situations while the Department of Education (2008a) taxonomy combines routine, non-routine application and analysis in one cognitive level group. The “application” category in the Department of Education (2008a) appears to best fit a combination of “routine execution” and “application” in the TASSK typology. The “analysis and reasoning” category in TASSK is matched with analysis, synthesis and evaluation categories in the Department of Education (2008a).

Table 3.3
Comparison of TASSK with the Department of Education (2008a) taxonomy

TASSK		Higher order / Lower order thinking	Department of Education (2008a)	
Retrieval	(level 1)	Lower order thinking	Recall	(level 1)
Basic Comprehension	(level 2)	Lower order thinking	Comprehension	(level 2)
Routine Execution	(level 3)	Lower order thinking		
			Application	(level 3)
Application	(level 4)	Higher order thinking		
			Analysis	(level 3)
Analysis and Reasoning	(level 5)	Higher order thinking	Synthesis	(level 4)
			Evaluation	(level 4)

3.5.6 Examples of NSC Physical Sciences examination questions classified using the TASSK cognitive demand typology.

The categories of the TASSK typology arose as a need to classify questions effectively with respect to cognitive demand in an attempt to measure the content standards of the physical sciences examination. I have taken questions from past senior certificate examinations to illustrate the cognitive demand required to answer that question. The complete question is provided in order to place the different parts discussed into context. The reasons for each cognitive category classification are explained after each assessment item.

a) Reaction Rates Question (Department of Basic Education, 2012b)

6. Calcium carbonate chips are added to an excess dilute hydrochloric acid solution in a flask on a balance as illustrated below. The cotton wool plug in the mouth of the flask prevents spillage of reactants and products, but simultaneously allows the formed gas to escape. The balanced equation for the reaction that takes place is:

$$CaCO_3\,(s) + 2HCl\,(aq) \rightarrow CaCl_2\,(aq) + CO_2\,(g) + H_2O\,(l)$$

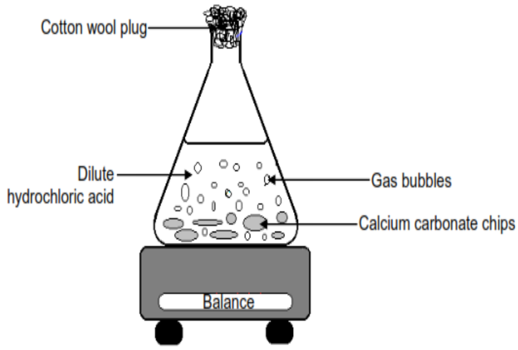

Figure 3.1 Diagram for reaction rates question

6.1 Write down the NAME of the gas that escapes through the cotton wool plug while the reaction takes place. *(1)*
Answer: Carbon dioxide

The above item is classified as "retrieval" as it requires the learner to recall the name of the gas produced. It does not demand any understanding or explanation.

The loss in mass of the flask and its contents is recorded in intervals of 2 minutes. The results obtained are shown in the graph below

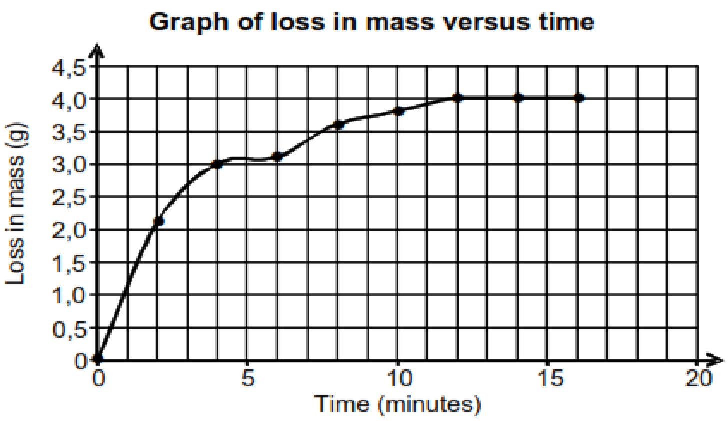

Figure 3.2 Graph for reaction rates question

6.2 From the graph, write down the following:
6.2.1 The coordinates of the point that represents results that were measured incorrectly. *(1)*
Answer: (6; 3,1)

The above item is classified as "basic comprehension" since it requires the learner to be able to read and interpret graphs in reaction rates.

6.2.2 How long (in minutes) the reaction lasts? *(1)*
Answer: 12 minutes

The above item is classified as "basic comprehension" since it requires the learner to be able to read and interpret graphs in reaction rates. The reaction stops when there is no further loss in mass.

6.2.3 How long (in minutes) it takes 75% (three quarters) of the reaction to occur? (1)
Answer: 4 minutes

The above item is classified as "basic comprehension" since it requires the learner to show simple computational ability in reaction rates.

6.3 The experiment is now repeated using hydrochloric acid of a higher concentration. It is found that the rate of the reaction INCREASES. Use the collision theory to explain this observation. (2)
Answer: A higher concentration implies that that there are more particles per unit volume. This leads to more effective collisions per unit time.

The above item is classified as "basic comprehension" since it requires the learner to demonstrate understanding of the familiar concepts of concentration and reaction rates.

6.4 How would a higher concentration of hydrochloric acid affect the following?
Loss in mass per unit time. (1)
Answer: Increases

The above item is classified as "basic comprehension" since it requires the learner to show some basic understanding of the familiar concepts of reaction rates.

Total loss in mass (1)
Answer: Remains the same

The above item is classified as "basic comprehension" since it requires the learner to show some basic understanding of the familiar concepts of reaction rates.

Time for the reaction to reach completion (1)
Answer: Decreases

The above item is classified as "basic comprehension" since it requires the learner to show some basic understanding of the familiar concepts of reaction rates.

6.5 Apart from concentration and temperature changes, write down TWO other changes that can be made to increase the rate of this reaction. (2)
Answer: Add a catalyst
Increase the surface area of the calcium carbonate/ use $CaCO_3$ powder.

The above item is classified as "retrieval" as it requires the learner to recall the factors that increase the rate of the reaction. It does not demand any understanding or explanation.

6.6 Calculate the mass of calcium carbonate used when the reaction is completed. Assume that all the gas that was formed, escaped from the flask. *(5)*

Answer:
$$n = \frac{m}{M}$$
$$= \frac{4}{44}$$
$$= 0{,}09\ mol$$
$$n\ (CaCO_3) = n\ (CO_2) = 0{,}09\ mol$$
$$0{,}09 = m/M$$
$$0{,}09 = m\ /\ 100$$
$$m\ (CaCO_3) = 0{,}09\ X\ 100$$
$$= 9\ g$$

The above item is classified as "routine execution" as it requires the learner to comprehend the question, recall the steps in a familiar situation and execute the algorithm.

b) Chemical Equilibrium Question (Department of Basic Education, 2012b)

7. A hypothetical reaction is represented by the balanced equation below.
A (g) + 2 B (g) → 2C (g)

Initially 3 moles of A (g) and 6 moles of B (g) are mixed in a 5 dm^3 sealed container. When the reaction reaches equilibrium at 25℃, it is found that 4 moles of B (g) is present.

7.1 Define the term chemical equilibrium. *(2)*

Answer: The stage in a reversible chemical reaction when the rate of the forward reaction equals the rate of reverse reaction.

The above item is classified as "retrieval" as it requires the learner to recall the definition of chemical equilibrium without demanding any understanding or explanation.

7.2 Show by calculation that the equilibrium concentration of C(g) is 0,4 $mol.dm^{-3}$ *(3)*

Answer:
n (B) reacted = 6 - 4 = 2mol
n (C) formed = n(B) reacted
= 2 mol
c (C) = $n\ /\ V$
= 2 / 5
= *0,4 $mol.dm^{-3}$*

The above item is classified as "routine execution" as it requires the learner to comprehend the question, recall the steps in a familiar situation and execute the algorithm.

7.3 How will an increase in pressure, by decreasing the volume of the container, influence the amount of C (g) in the container at 25℃? Write down INCREASES, DECREASES or REMAINS THE SAME. Explain the answer. *(3)*

Answer:
Increases
3mol / volumes (of gas) produce 2 mol / volumes (of gas).
The reaction which produces the smaller number of moles/volume is favoured.
Forward reaction is favoured.

The above item is classified as "routine execution" as it requires the learner to comprehend the question, recall the steps in a familiar situation and execute the procedure regarding Le Chatelier's principle.

7.4 The initial number of moles of B (g) is now increased while the initial number of moles of A(g) remains constant at 25 ℃. Calculate the number of moles B (g) that must be ADDED to the original amount (6 mol) so that the concentration of C (g) is 0,8 mol.dm^{-3} at equilibrium. The equilibrium constant (K_c) for this reaction at 25 ℃ is 0,625. *(9)*

Answer: Let ***x*** *represent the total initial amount of B (g) that must be used.*

	A	*B*	*C*
Initial quantity (mol)	*3*	***x***	*0*
Change (mol)	*-2*	*- 4*	*+4*
Quantity at equilibrium (mol)	*1*	***x*** *– 4*	*4*
Equilibrium concentration (mol.dm^{-3})	*0,2*	$\frac{(x-4)}{5}$	*0,8*

$$K_c = \frac{[C]^2}{[A][B]^2}$$

$$0{,}625 = \frac{(0{,}8)^2}{(0{,}2)\left(\frac{x-4}{5}\right)^2}$$

$$\mathbf{x} = 15{,}3 \text{ mol}$$

$$n\,(B)\text{ added} = 15{,}3 - 6$$

$$= 9{,}3 \text{ mol}$$

The above item is classified as "application" as it requires the learner to think beyond the routine steps in this situation, use the knowledge and skills learned to adopt a new approach to answer the question.

c) Momentum Question (Department of Basic Education, 2012a)

4. The diagram below shows a car traveling at a velocity of 20 $m.s^{-1}$ east on a straight level road and a truck of mass 2m traveling at 20 $m.s^{-1}$ west on the same road. Ignore the effects of friction.

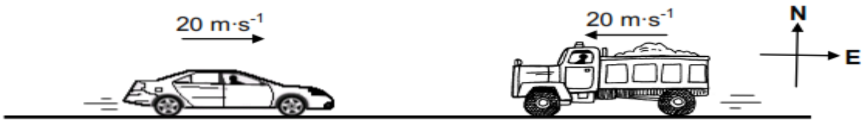

Figure 3.3 Diagram for momentum question

4.1 Calculate the velocity of the car relative to the truck. *(2)*
Answer: 40 $m.s^{-1}$ East

The above item is classified as "basic comprehension" as it requires a few familiar connections concerning relative velocity to be made.

The vehicles collide head-on and stick together during the collision.

4.2 State the principle of conservation of linear momentum in words. *(2)*
Answer: The total linear momentum of an isolated system is conserved / remains constant in both magnitude and direction.

The above item is classified as "retrieval" as it requires the recall of the statement of the principle of conservation of linear momentum.

4.3 Calculate the velocity of the truck-car system immediately after the collision. *(6)*

Answer: *Consider east as positive*

$$\Sigma p_i = \Sigma p_f$$
$$m(20) + 2m(-20) = (m + 2m)\,v_f$$
$$v_f = -6.67\ m.s^{-1}$$
$$v_f = 6.67\ m.s^{-1}\ \text{West.}$$

The above item is classified as "routine execution" as the learner would have to extract information from the familiar momentum situation given and routinely execute a procedure.

*4.4 On impact the car exerts a force of magnitude **F** on the truck and experiences an acceleration of magnitude **a**.*

*4.4.1 Determine, in terms of **F**, the magnitude of the force that the truck exerts on the car on impact. Give a reason for the answer.* *(2)*

Answer: F
Newton's Third Law of Motion

The above item is classified as "basic comprehension" as it requires connecting some familiar relationships involving Newton's third law.

*4.4.2 Determine, in terms of **a**, the acceleration that the truck experiences on impact. Give a reason for the answer.* *(2)*

Answer: $\frac{1}{2}\boldsymbol{a}$ *Same* F_{net}, $\boldsymbol{a} \ \alpha \ \frac{1}{m}$

The above item is classified as "routine execution" as the learner would have to extract information from the familiar Newton's second law situation and routinely execute a procedure.

4.4.3 Both drivers are wearing identical seatbelts. Which driver is likely to be more severely injured on impact? Explain the answer by referring to acceleration and velocity *(3)*

Answer: Car driver
Car-driver system has greater acceleration.
Car-driver system has greater change in velocity / greater Δv

The above item is classified as "routine execution" as the learner would have to extract information from the familiar situation involving Newton's second law and routinely execute a procedure.

d) Electric Circuit Question (Department of Basic Education, 2012a)

8. In the circuit represented below, an uncharged capacitor is connected in series with a 1000 Ω resistor. The emf of the battery is 12 V. Ignore the internal resistance of the battery and the ammeter.

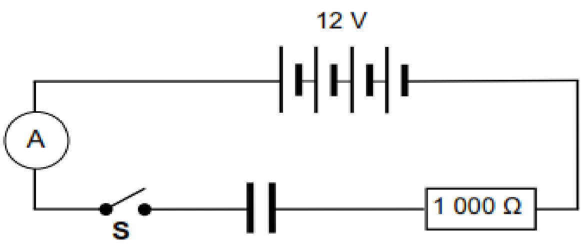

Figure 3.4 Diagram for electric circuit question

8.1 Calculate the initial current in the circuit when switch S is closed. *(3)*

Answer:

$$R = \frac{V}{I}$$

$$1000 = \frac{12}{I}$$

$$I = 1{,}012 \ A$$

The above item is classified as “routine execution” as the learner would have to extract information from the familiar electrical circuit situation given and routinely execute a procedure.

8.2 Write down the potential difference across the plates of the capacitor when it is fully charged. *(1)*

Answer: *12 V*

The above item is classified as “basic comprehension” as it requires the learner to make basic connections involving capacitance.

8.3 The capacitor has a capacitance of 120 μF and the space between the plates is filled with air. Calculate the charge stored on the plates of the capacitor when it is fully charged. *(3)*

Answer:

$$C = \frac{Q}{V}$$

$$120 \, X \, 10^{-6} = \frac{Q}{12}$$

$$Q = 1{,}44 \, X \, 10^{-3} \, C$$

The above item is classified as “routine execution” as the learner would have to extract information from the familiar capacitance situation given and routinely execute a procedure to determine the charge stored.

After discharging the capacitor, it is connected in the same circuit to a resistor of higher RESISTANCE and switch S is closed again.

8.4 How would this change affect each of the following?
(Write down INCREASES, DECREASES or REMAINS THE SAME)

8.4.1 The initial charging current *(1)*
Answer: Decreases

The above item is classified as “basic comprehension” as it requires the learner to make familiar connections involving capacitance.

8.4.2 The time it takes for capacitor to become fully charge *(1)*
Answer: Increases

The above item is classified as “basic comprehension” as it requires the learner to make basic connections involving capacitance.

8.5 The two parallel plates of the fully charged capacitor are 12mm apart.
8.5.1 Sketch the electrical field pattern between the parallel plates. *(3)*

Answer:

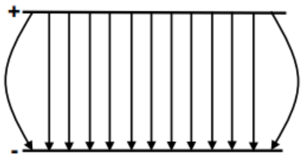

Figure 3.5 Diagram for electrical field pattern answer

The above item is classified as “retrieval” as it requires the learner to reproduce a seen electrical field pattern diagram involving oppositely charged parallel plates.

8.5.2 Calculate the magnitude of the electric field at a point midway between the plates. (3)
Answer:

$$E = \frac{V}{d}$$

$$= \frac{12}{12\,X\ 10^{-3}}$$

$$= 1000\ V.m^{-1}$$

The above item is classified as “routine execution” as the learner would have to extract information from the familiar electrical field situation given and routinely execute a procedure to determine the electric field.

3.6 CHAPTER SUMMARY

In this chapter, the research design and procedures of data collection in this study were first described. A post-positivist orientation to research was the approach used in this study. The analysis instrument in this study was considered reliable by the subject experts. The data was collected to analyse the physical sciences examination question papers in terms of cognitive demand. All NSC final examination papers from 2008 to 2013 were subjected to a process of deductive analysis to determine the cognitive demand. The next part of this chapter described the development of the cognitive demand taxonomy. The literature surveyed indicated that the revised Bloom’s taxonomy was often used as an instrument to determine the cognitive demand of examination items. I found it to be problematic in the categorisation and opted for the development of a more suitable taxonomy. The TASSK typology was developed and found to be reliable by four subject experts with extensive experience in physical sciences teaching and examinations. This taxonomy distinctively separates higher order and lower order cognitive skills into five unambiguous categories. A comparison of the TASSK cognitive demand typology with other relevant taxonomies was discussed. The chapter ends with showing how examples from past examination question papers were classified according to the TASSK typology. The data analysis and findings are discussed in chapter four.

Chapter 4

DATA ANALYSIS AND DISCUSSION OF FINDINGS

The previous chapter described the development of the cognitive demand taxonomy, the research design and the procedures of data collection. The data was analysed using the TASSK cognitive demand typology. This chapter presents, describes and discusses the data analysis and findings of the study. The verification and reliability of the cognitive demand taxonomy was first presented. The cognitive demand of the examination questions for each paper was then determined using the taxonomy that was considered reliable by the experts. This was aggregated to infer the cognitive demand of each paper. Thereafter the mean percentage per cognitive category for each year's paper was found. Comparisons were then made with the cognitive demand of the examination question papers and the recommendations made by the Department of Education (2008a) and discussions on the findings concluded the chapter.

4.1 VERIFICATION AND RELIABILITY OF THE INSTRUMENT

This section presents the findings from the analysis of papers by the five raters to verify the instrument and determine whether it was a reliable instrument to measure the cognitive demand. The complete analysis carried out by the expert group for the November 2012 paper one and paper two is shown in Appendix I and Appendix J respectively.

4.1.1 Rater analysis of examination question papers.

Each assessment item was categorised as one of five cognitive levels representing the cognitive processes of retrieval (level 1), "basic comprehension" (level 2), "routine execution" (level 3), "application" (level 4) and "analysis and reasoning" (level 5) using the TASSK cognitive demand typology. Table 4.1 which follows shows some items where differences in terms of cognitive levels for the November 2012 paper one (Department of Basic Education, 2012a) occurred.

Table 4.1

Rater analysis of November 2012 paper 1 (Physics) showing cognitive demand level by question.

Question No.	Researcher	Rater 1	Rater 2	Rater 3	Rater 4		Final level
3.1	2	1	1	1	2	X	1
3.3	4	3	4	3	4	X	4
4.4.3	2	2	3	3	3	X	3
8.5.1	1	1	1	2	2	X	1

The November 2012 paper one has 64 assessment items. The raters were consistent with 56 of the 64 assessment items that were categorised. A cross (X) in column seven represents a difference in categorisation among the raters. Where differences occurred, these items were discussed and thereafter consensus among the raters was reached. This final level after consensus was used in the analysis and is indicated in the last column. For example in question 3 below there was some disagreement.

> *3. An object is projected vertically upwards at 8 m.s^{-1} from the roof of a building which is 60m high. It strikes the balcony (at a height h above the ground) after 4 seconds. The object then bounces off the balcony and strikes the ground. Ignore the effects*
>
> > *3.1 Is the object's acceleration at its maximum height UPWARD, DOWNWARD or ZERO?* *(1)*

Two of raters classified the above item as “basic comprehension” (cognitive level 2). They argued that it was not retrieval of definition or term but required the learner to make a few simple connections by identifying that the acceleration due to gravity was directed downwards by initially eliminating the distracters UPWARD DOWNWARD or ZERO in the question which sought to describe velocity and not acceleration due to gravity. The other three raters classified it as “retrieval” as it was a familiar concept taught in lessons and features often in standard text book exercises and past examination papers. One of the raters who is a classroom teacher further said that when he explains this in his lessons, he drills the point that the acceleration due to gravity at any point of the object’s motion acts downwards. The raters eventually reached consensus and classified the item as “retrieval” (cognitive level 1).

Sub-question 3.3 was another example where there was some difference in the categorisation.

> *3.3 The object bounces off the balcony at a velocity of 27,13 m.s^{-1} and strikes the ground 6 s after leaving the balcony.*
> *Sketch a velocity-time graph to represent the motion of the object from the moment it is projected from the ROOF of the building until it strikes the GROUND. Indicate the following velocity and time values on the graph:*
> *the initial velocity at which the object was projected from the roof of the building*
> *the velocity at which the object strikes the balcony*
> *the time when the object strikes the balcony*
> *the velocity at which the object bounces off the balcony*
> *the time when the object strikes the ground* *(6)*

Two raters classified the above item as "routine execution" while three raters classified the item as "application". The two raters argued that this item requires a routine procedure to be followed and requires familiar connections to be made. The other three raters argued that while questions on this topic requiring the learner to draw velocity-time graphs is common in past papers, this question is somewhat different to those that appeared in past examination papers and requires new connections to be made. The situation of the projected ball bouncing off the balcony in the building to another height before striking the ground would make the learner now think and apply previous knowledge gained in this topic. It was agreed that this graph question has not appeared in this form in the recent past examinations and required new connections to be made and the raters eventually reached consensus and classified the item as an "application" type question (cognitive level 4).

The November 2012 paper two has 82 assessment items. The raters were consistent with 70 of the 82 assessment items that were categorised. Table 4.2 which follows shows some items where differences in categorisation in terms of cognitive levels for the November 2012 paper two (Department of Basic Education, 2012b) occurred. Where differences occurred, these items were discussed and thereafter consensus among the raters was reached. This final level reached after consensus was used in the analysis and is indicated in the last column.

Table 4.2
Rater analysis of November 2012 paper 2 (Chemistry) showing cognitive demand level by question

Question No.	Researcher	Rater 1	Rater 2	Rater 3	Rater 4		Final level
4.3	2	1	2	1	2	X	2
4.4	1	1	3	3	3	X	1
5.2.1	2	2	2	1	2	X	2
8.1.3	3	2	3	3	3	X	3

For example in question 4.4 there was some disagreement among raters:

> *Alkanes burn readily in oxygen. Write down a balanced equation, using molecular formulae, for the combustion of propane in excess oxygen.*

Two raters classified the above item as “retrieval” while three raters classified the item as “routine execution”. The two raters argued that this item requires recall and is clearly stated in the syllabus and given in the text books and has appeared in previous examination papers. They point out that it is further a seen equation and understanding is not necessarily required. The other three raters argued that the learner applies a rule to answer this question and has to further balance the complicated equation using a procedure. It was eventually settled that although the question was complicated it was not cognitively complex and required recall as it is a “seen” example in the syllabus, text books and past papers and was thus categorised as “retrieval” (cognitive level 1).

There was also some disagreement regarding sub-question 8.1.3 which is discussed below.

> *8.1 A strip of aluminium is placed in a beaker containing a blue solution of copper(II)salt. After a while the solution becomes colourless.*
>
> *8.1.3 Write down the balanced net ionic equation for the reaction that takes place. (3)*

Four raters classified the above item as “routine execution” and one rater classified it as “basic comprehension”. The question required the learner to comprehend the question, follow a routine procedure and use the half reaction table to determine the oxidation and reduction half reactions and thereafter find the balanced net ionic equation. It was agreed that this was a “routine execution” type question and categorised as cognitive level 3.

All other differences that occurred during the analysis were discussed and a common understanding among the raters was reached in a similar manner as discussed in the examples above. All the differences and discussions leading to consensus were noted by the researcher and the same reasoning was followed in the analysis of all the examination papers for example if any question required the learner to apply new connections to a new situation even though it resembled routine type of questions as discussed in the vertical projectile graph question 3.3 above, then it would be categorised as cognitive level 4. The rate of consistency among the raters to classify the assessment items is critical to the reliability of the cognitive instrument and is discussed below.

4.1.2 Inter-rater reliability.

Many educational and psychological studies require the use of independent raters in order to quantify some aspect of behaviour (Stemler, 2004). The value of inter-rater reliability has significant importance to the validity of the studies like this. Inter-rater reliability is used to assess the degree to which different raters make consistent estimates of the same phenomenon (Multon, 2010). In this study, this would imply the extent to which raters' classification of assessment items in terms of cognitive demand is consistent. Two statistical measures of inter-rater reliability are used in this study, namely the percentage agreement and kappa co-efficient. The traditional method of getting an estimate is obtained by using the percentage estimate of reliability. This popular method of calculating inter-rater reliability is conceptually simpler to understand, explain and easy to compute. Percentage agreement is calculated by adding up the number of cases that received the same rating by the raters and dividing that number by the total number of cases rated. The percentage inter-rater agreement is given by: $I = \frac{\mathrm{A}}{\mathrm{T}}$; where *I* represents the inter-rater agreement, A represents the number of items that raters reached agreement on and T represents the total number of items. In the above case of the November 2012 physics paper, the inter-rater reliability was calculated as follows:

$$\begin{aligned} I &= \frac{\mathrm{A}}{\mathrm{T}} \\ &= \frac{56}{64} \\ &= 0.87 \end{aligned}$$

In a similar manner, percentage estimate for inter-rater reliability for the November 2012 chemistry paper was calculated as 0.85. Stemler (2004) points out that a value above 0.70 is considered acceptable. While the inter-rater agreement for the two papers mentioned above is acceptable for this study, Grayson and Rust (2001) argue that percentage of agreement is not the ideal measure as it does not take into account chance agreement. Stemler (2004) adds that Cohen's kappa is a highly useful statistic when one is concerned that the percent-agreement statistic may be artificially inflated due to the fact that most observations fall into a single category. Multon (2010) concurs with the views of Grayson and Rust (2001) and Stemler (2004) that Cohen's kappa, represents a better estimate of inter-rater agreement corrected for chance. Thus I went ahead to calculate the kappa co-efficient which is given by:

$K = (p_a - p_c) / (1 - p_c)$, where

K represents the kappa co-efficient,
p_a represents the proportion of times raters agree and
p_c represents the proportion of agreement one would expect by chance

(Grayson & Rust, 2001)

Multon (2010) notes that the kappa co-efficient range from 0 to 1 and a value of 0.50 is considered acceptable. Stemler (2004) argues that the interpretation of the kappa statistic is slightly different than the interpretation of the percent-agreement figure. He adds that a value of zero on kappa does not indicate that the judges did not agree at all; rather, it indicates that the raters did not agree with each other any more than would be predicted by chance alone. Landis and Koch (1977) indicate that kappa values between 0.01 and 0.20 suggest slight agreement, values between 0.21 and 0.40 suggest fair agreement, values between 0.41 and 0.60 suggest moderate agreement and those above 0.60 represent substantial agreement. Using the kappa co-efficient, I calculated the inter-rater reliability for the November 2012 physics examination paper as follows:

$$
\begin{aligned}
K &= (p_a - p_c) / (1 - p_c) \\
&= (0.87 - 0.20) / (1 - 0.20) \\
&= 0.67 / 0.8 \\
&= 0.84
\end{aligned}
$$

Similarly the kappa co-efficient for the November 2012 chemistry paper was calculated as 0.81. These are acceptable values according to Landis and Koch (1977) and Multon (2010). Having obtained acceptable values for inter-rater reliability using two approved methods, indicated substantial consistency between raters and shows that the

TASSK typology can be used reliably. The NSC physical sciences examination question papers for the period 2008 to 2013 were analysed using the TASSK typology and are summarised in the section that follows.

4.2 ANALYSIS OF EXAMINATION QUESTION PAPERS

As described in Chapter 3, the researcher first conducted a pilot study together with four other raters by analysing the March 2013 paper one. After discussions and an agreement on how to conduct the analysis was reached, the raters then analysed the November 2012 paper one (physics) and thereafter analysed the November 2012 paper two (chemistry). The inter-rater reliability calculated in section 4.1.2 showed that the TASSK typology was a reliable instrument and the researcher used this instrument to analyse all the final NSC physical sciences examination papers set by the DBE.

Table 4.3 which follows shows the results of analysis of the physics examination question papers from 2008 to 2013. The first row indicates the percentage cognitive weighting that is recommended by the DBE. The DBE's "application" is matched with the combined "routine execution" and "application" categories of the TASSK. This comparison was discussed in section 3.5.5. The table below also shows the percentage per cognitive category for each year's question paper. The mean percentage per cognitive category was calculated for the period 2008 to 2013 and the variation between the DBE recommendation and the mean percentage is indicated on the last row of the table.

On average over the six year period, 93% of the total marks in question paper one (Physics) were allocated to questions requiring lower order thinking skills of "retrieval", "basic comprehension" and "routine execution". Only 7% of total marks were allocated to questions requiring higher order thinking skills of the "application" type. There were no marks allocated to test the higher order thinking skills of "analysis and reasoning" as described in the TASSK, the Department of Education (2008a) and the Department of Basic Education (2011a) taxonomies in any of the physics papers over the six year period from 2008 to 2013.

Table 4.3
Percentage of the total marks of Physical Sciences examination paper 1 (Physics) from 2008 to 2013 per cognitive level

<table>
<tr><th></th><th colspan="3">Lower order thinking skills</th><th colspan="2">Higher order thinking skills</th></tr>
<tr><th>Cognitive Domain</th><th>Retrieval
Level 1</th><th>Basic Comprehension
Level 2</th><th>Routine Execution
Level 3</th><th>Application
Level 4</th><th>Analysis & Reasoning
Level 5</th></tr>
<tr><td>DBE %</td><td>15</td><td>30</td><td colspan="2">45</td><td>10</td></tr>
<tr><td>2008</td><td>19</td><td>28</td><td>39</td><td>14</td><td>0</td></tr>
<tr><td>2009</td><td>17</td><td>31</td><td>43</td><td>9</td><td>0</td></tr>
<tr><td>2010</td><td>16</td><td>24</td><td>57</td><td>3</td><td>0</td></tr>
<tr><td>2011</td><td>18</td><td>24</td><td>52</td><td>6</td><td>0</td></tr>
<tr><td>2012</td><td>15</td><td>25</td><td>56</td><td>4</td><td>0</td></tr>
<tr><td>2013</td><td>19</td><td>26</td><td>48</td><td>7</td><td>0</td></tr>
<tr><td>Mean %</td><td>17</td><td>27</td><td>49</td><td>7</td><td>0</td></tr>
<tr><td>Variation %</td><td>+2</td><td>-3</td><td colspan="2">+11</td><td>-10</td></tr>
</table>

From table 4.3 it can be seen how the cognitive demand compares with the recommendations in the NCS subject assessment guidelines (Department of Education, 2008a). The “retrieval” category was over-emphasised in all question papers except the November 2012 paper. The “basic comprehension” was under-emphasised in all but the November 2009 question papers. The combined “application” to routine and new situations was over-emphasised in all the question papers while the higher order thinking questions that involve problem solving, decision making, experimenting and investigating according to the TASSK typology did not feature at all during the study period.

Table 4.4 which follows shows the analysis results of the chemistry examination question papers from 2008 to 2013. The first row gives the DBE percentage cognitive weighting recommendations for chemistry which differs from that of physics. This is followed by the percentage per cognitive category of the TASSK typology for each year’s paper. The mean per cognitive category is worked out for the period 2008 to 2013 and the variation between the DBE recommendation and the mean percentage is indicated on the last row of the table.

Table 4.4
Percentage of the total marks of Physical Sciences examination paper 2 (Chemistry) from 2008 to 2013 per cognitive level

Cognitive Domain	Lower order thinking			Higher order thinking	
	Retrieval Level 1	Basic Comprehension Level 2	Routine Execution Level 3	Application Level 4	Analysis & Reasoning Level 5
DBE	15	40	35		10
2008	30	41	29	0	0
2009	18	47	28	7	0
2010	30	29	32	9	0
2011	25	43	32	0	0
2012	15	25	56	4	0
2013	18	28	50	4	0
Mean %	22	36	38	4	0
Variation %	+7	-4	+7		-10

On average over the six year period, the findings of this study regarding paper two (Chemistry) reveal that 96% of the total marks were allocated to questions that tested skills requiring lower order thinking, while only 4% of the total marks were allocated to questions that tested skills requiring higher order thinking skills of the "application" type. The 2008 and 2011 chemistry question papers, however, did not include any of the "application" type questions. There were no marks allocated to test the higher order thinking skills of "analysis and reasoning" as described in the TASSK, the Department of Education (2008a) and the Department of Basic Education (2011a) taxonomies in any of the chemistry papers over the six year period from 2008 to 2013. In comparison with the physics papers, the chemistry papers for this period appear to have a lower cognitive demand. This, however, must be seen in the context of the DBE recommended weighting of cognitive level 3 which constitute 45 % for physics and 35 % for chemistry.

From table 4.4 it can be seen how the cognitive demand compares with the recommendations in the NCS subject assessment guidelines (Department of Education, 2008a). The "retrieval" category was over-emphasised in all question papers except the November 2012 paper. The 2008 and 2010 question papers doubled the recommendations for the "retrieval" category indicating an emphasis on lower order memory skills. The "basic comprehension" was under-emphasised in the November 2010, 2012 and 2013 chemistry question papers and over-emphasised in the 2009 and 2011 chemistry question papers. The combined “application” to routine and new situations was under-emphasised in the 2008 question paper and over-emphasised in all the other papers while the higher order thinking questions that involve problem solving, decision making, experimenting and investigating did not feature at all during the study period.

To answer the research question,
What are the cognitive demands of the NSC physical sciences examination question papers?
I combined the cognitive demand analysis of paper 1 and paper 2 since the DBE aggregates the NSC Physical sciences examination paper 1 (physics) and paper 2 (chemistry) and arrives at examination mark for the learner. The cognitive demand levels for the NSC physical sciences examination is reflected in table 4.5 below.

Table 4.5
Combined percentages of marks per cognitive level for paper 1 (Physics) and paper 2 (Chemistry)

	Lower order thinking skills			Higher order thinking skills	
Cognitive Domain	Retrieval Level 1	Basic Comprehension Level 2	Routine Execution Level 3	Application Level 4	Analysis & Reasoning Level 5
DBE Mean %	15	35	40		10
Mean P1 (Physics)	17	27	49	7	0
Mean P2 (Chemistry)	22	36	38	4	0
Mean % (P1 & P2)	19	32	43	6	0
Variation %	+4	-3	+9		-10

On average over the six year period, the findings of this study reveal 94% of the total marks were allocated to test lower order cognitive demand in the physical sciences NSC papers from 2008 to 2013 according to the TASSK cognitive demand typology. Application skills accounted for the remaining 6% of the total mark allocation, while 0% of total marks accommodated the testing of higher order thinking questions involving “analysis and reasoning” using this taxonomy.

4.3 COMPARISON OF THE TASSK AND DBE ANALYSIS

The search for the DBE cognitive demand analysis for the period of study in the public domain yielded one result for the November 2011 examination papers. This DBE analysis was extracted from the Department of Basic Education (2011e) that reported on the 2011 NSC learner performance. The complete TASSK analysis of the November 2011 paper 1 (Physics) is given in Appendix K and the TASSK analysis of paper 2 is given in Appendix L. The DBE analysis results for the November 2011 paper 1 are compared with the findings from the TASSK typology analysis in this study and are shown in Table 4.6 below.

Table 4.6
Difference in percentage of marks allocated to each cognitive level by the DBE and the TASSK for the November 2011 Physics examination

	Weighting of cognitive levels				
Cognitive Domain	Recall	Comprehension	Application/ Analysis		Evaluation/ Synthesis
% DBE recommendations	15 Level 1	30 Level 2	45 Level 3		10 Level 4
% DBE analysis	17	22	51		10
Cognitive Domain	Retrieval	Basic Comprehension	Routine Execution	Application	Analysis & Reasoning
TASSK levels	Level 1	Level 2	Level 3	Level 4	Level 5
% TASSK analysis	18	24	52	6	0

With regard to the November 2011 physics paper, the analysis of the “recall” / “retrieval” category in the DBE and the TASSK on par with each other within margin for error. The results for the “understanding” / “basic comprehension” categories are on par within margin for error to the DBE recommendations. The TASSK analysis shows that the “routine execution” appears to correspond better with the DBE’s “application” and “analysis” category. The TASSK analysis shows that none of the questions demanded higher order thinking as described in “analysis and reasoning” category while the DBE analysis showed that 10% of the questions were categorised as “evaluation” and “synthesis”.

I refer to Question 5 from November 2011 paper one (Department of Basic Education, 2011c) to discuss this difference in the categorisation.

> *5. A rescue helicopter is stationary (hovers) above a soldier. The soldier of mass 80kg is lifted vertically upwards through a height of 20m by a cable at a CONSTANT SPEED of 4 m.s^{-1}. The tension in the cable is 960 N.*
>
> *5.4 Use the WORK-ENERGY THEOREM to calculate the work done on the soldier by friction after moving through the height of 20m.* *(5)*

The DBE analysis classifies this question as having a cognitive demand of level 3/4 (Department of Basic Education, 2011e, pp. 124-125). This would imply that the question demands “application”, “analysis” (cognitive level 3), “synthesis” or “evaluation” (cognitive level 4). There is no creation of new ideas and no judgements based on evidence made to answer this question warranting it to be classified as cognitive level 4 on the DBE taxonomy. The question also directs the learner to use the work energy theorem hence it does not require the learner to identify patterns and analyse the question. The TASSK analysis categorises this question as a “routine execution” type question since it requires the learner to apply a set of known (routine) steps involving the work energy theorem. Based on the TASSK, this question cannot be classified as one that requires higher order reasoning as it does not demand new connections to be made by the prepared learner.

The DBE analysis results for the November 2011 paper 2 (Chemistry) are compared with the findings from the TASSK typology analysis in this study and are shown in table 4.7 below. The analysis of the chemistry paper showed that the “recall” / “retrieval” category in the DBE analysis and the DBE recommendations are on par with each other but the TASSK analysis indicates that this category is over-emphasised when compared to the DBE recommendations. The analysis of the “understanding” / “basic comprehension” categories of the DBE and the TASSK vary from each other and the DBE analysis shows that this category is under-emphasised relative to the DBE recommendations. The TASSK analysis shows that the “routine execution” appears to correspond better with the DBE’s “application” recommendations. The TASSK analysis showed that none of the questions demanded higher order thinking as described in the “analysis and reasoning” category, while the DBE analysis showed 7 % of the questions being categorised as “evaluation” and “synthesis”.

Table 4.7

Difference in percentage of total marks allocated to each cognitive level by the DBE and TASSK for the November 2011 Chemistry examination

Weighting of cognitive levels					
Cognitive Domain	Recall	Comprehension	Application/ Analysis		Evaluation/ Synthesis
% DBE recommendation	15 Level 1	40 Level 2	35 Level 3		10 Level 4
% DBE analysis	15	35	43		7
Cognitive Domain	Retrieval	Basic Comprehension	Routine Execution	Application	Analysis & Reasoning
TASSK levels	Level 1	Level 2	Level 3	Level 4	Level 5
% TASSK analysis	25	43	32	0	0

To show the cognitive demand categorisation of the TASSK and DBE taxonomies, I discuss sub-question 9.4.1 from the November chemistry paper below.

> *In the electrolytic cell, two CARBON RODS P and Q are used as electrodes and a concentrated copper (II) chloride solution is used as an electrolyte.*
>
> *9.4 The carbon rods in the above cell are now replaced with COPPER RODS.*
>
> *The following observations are made at electrode P.*
>
> *No gas is released*
>
> *Its surface appears rough and eroded.*
>
> *9.4.1 Refer to the RELATIVE STRENGTHS OF REDUCING AGENTS to explain this observation.* *(3)*

The question is classified by the DBE as involving higher order thinking skills of “synthesis” or “evaluation” and having a cognitive demand level 4 (Department of Basic Education, 2011e, p. 151). The TASSK typology considers this to be a “routine execution” of a familiar procedure. It involves lower order thinking as the question directs the prepared learner to refer to relative strengths of reducing agents and explain a familiar procedure. The learner has to use a given table and familiar procedures to deduce that the chloride (Cl^-) ion is a weaker reducing agent than copper (Cu) and will therefore not be oxidised. Copper (Electrode P) will be oxidised resulting in the plate becoming eroded. Similar questions of this nature have been asked previously and the learner is not faced with making new connections in an unfamiliar situation and therefore the question does not expose the learner to think critically.

4.4 DISCUSSION OF FINDINGS

Before the TASSK cognitive demand typology could be used, it was verified by subject experts and the inter-rater reliability calculated was found to be acceptable. Thus in the opinion of the subject experts who were very familiar with the teaching and learning of grade 10, 11 and 12 physical sciences the TASSK cognitive demand typology was a reliable instrument. The researcher used this taxonomy to analyse all NSC examination papers in terms of cognitive demand levels. The analysis revealed that higher order thinking questions involving “analysis and reasoning” were conspicuously absent in all the examination question papers.

The study by Govender and Moodley (2012) of first year physics students’ indicated that student’s scores dropped drastically in 2009, more so than in than previous years. These students were learners from the 2008 first NSC cohort. Their study concluded that the new NCS produced students less prepared for university study than the previous senior certificate examinations. It appears that the emphasis of lower order thinking skills in the new NSC physical sciences examinations may have allowed the students to gain entrance to university physics courses without having acquired the necessary higher order thinking skills to cope with university physics.

The chemistry papers of 2008, 2010, and 2011 over-emphasised the “retrieval” type questions. In contrast to this, Liu et al. (2008) point out that there is an overall shift to test higher order thinking by de-emphasising the “remember” skill and emphasising the “analyse” skill in China and Singapore. Muller (2005) warned that low cognitive demand threatened the health of the education system. It appears that NSC examiners have not heeded this warning. The Human Science Research Council (2012) has pointed out that South African top learners feature with the average learners of top performing participating TIMSS countries. The de-emphasizing of higher order thinking in favour of lower order thinking skills may have contributed to the poor standing of our top learners in international assessments like TIMSS.

On average 94% of the questions emphasised lower order thinking skills of “retrieval”, “basic comprehension” and “routine execution”. When compared to the recommendations on cognitive demand weightings made by the DBE, it was found that the

TASSK categories of “basic comprehension” was under-emphasised, “retrieval” and “routine execution” were over-emphasised while questions that involved the higher order thinking skills of “analysis and reasoning” were non-existent in all the examination question papers analysed.

The Department of Education (2011e) analysis, however, claimed that there were questions based on evaluation and synthesis in the November 2011 examination question papers. The discussions on the examples in section 4.3 challenge these claims for the 2011 papers. The findings of this study also throws into doubt the claim made by Umalusi (2009) that the 2008 examination question papers consist of 50% problem solving as the instrument used by Umalusi did not separate routine execution from the higher order thinking skills required for problem solving. This study found that the NSC examination question papers for the period 2008 to 2013 were of low cognitive demand and did not afford learners the opportunity to engage in analysis, scientific reasoning, problem solving, decision making, generating and testing of hypotheses.

The data in this study infers that prepared learners can score “A” symbols from the practising of examples from past papers and standard text book involving “retrieval”, “basic comprehension” and “routine execution” type questions since the findings reveal that 94 % of the NSC examination question papers test these lower order thinking skills. Some of these learners may not necessarily be exposed to the mastery of higher order thinking skills and still be classified as top learners.

Chapter 5

SUMMARY, IMPLICATIONS, RECOMMENDATIONS AND CONCLUSIONS

This chapter summarises, lists the limitations, indicates the implications, and gives recommendations and final words of the study. The purpose of this study was to determine the cognitive demand of the NSC physical sciences examination question papers. The researcher developed a TASSK typology to measure the cognitive demand of examination questions. Subject experts were utilised due to their vast experience in physical sciences teaching and examinations to validate the taxonomy. The taxonomy was first piloted to get the researcher and four subject experts familiar with the TASSK cognitive demand typology. Some adjustments to the taxonomy were made after discussions with the experts. The raters first analysed the November 2012 paper one (physics) and thereafter analysed the November 2012 paper two (chemistry) papers through a process of deductive analysis. Inter-rater reliability was calculated and found to be acceptable. After the cognitive instrument was found to be reliable by the subject experts, the researcher analysed all the NSC examination question papers from 2008 to 2013.

To answer my research question one, assessment items were analysed against the TASSK cognitive demand typology. Each assessment item was categorised as one of five cognitive levels representing the cognitive processes of “retrieval” (level 1), “basic comprehension” (level 2), “routine execution” (level 3), “application” (level 4) and “analysis and reasoning” (level 5) using the TASSK cognitive demand typology. The cognitive demand of the examination questions for each paper was first determined. This was aggregated to infer the cognitive demand of each paper. Thereafter the average percentage per cognitive category for each year’s paper was found. By combining the averages of the physics and chemistry papers, the cognitive demand of the physical sciences papers for the period was determined. To answer my research question two, comparisons were made with the cognitive demand of the examination question papers and the recommendations made by the Department of Education (2008).

5.1 SUMMARY OF FINDINGS

The summary to each of my research questions is found below. The findings are presented as answers to my research questions.

5.1.1 Research question 1.

What are the cognitive demands of the NSC physical sciences examination question papers?

The findings reveal that on average the NSC physical sciences examination question papers for the period 2008 to 2013 tested 19% “retrieval”, 32% “basic comprehension”, 43% “routine execution” and 6% “application” type thinking skills. The study revealed that there were no questions testing higher order thinking skills as described by the “analysis and reasoning” category of the TASSK typology. This implies that 94% of the examination question papers focused on lower order thinking skills. Apart from low cognitive demand, grade inflation can further boost the scores of learners with low raw scores into the 80% to 100 % category. Such a situation does not reveal the true value of symbols awarded to NSC learners and warrants further research into the standards of the examination question papers.

Thus if the NSC examination is a barometer of the success of the education system (Department of Basic Education, 2014a), we require urgent intervention to restore confidence in our NSC content standards. The findings of this study lend support to the claims of low senior certificate cognitive demand standards made by Muller (2005), Vinjevold (2005), Umalusi (2009), Nel and Kistner (2009), Edwards (2010), Mhlolo (2011) Govender and Moodley (2012), and Jansen (2012). The lack of opportunities for learners to demonstrate their higher order thinking skills in senior certificate examinations and the urgency regarding the necessity for national policy to be lucid about what different levels of cognitive demand mean (especially what constitutes higher order thinking skills) as pointed out by Crowe (2012) further supports the need to raise the levels of cognitive demand in the senior certificate examinations. It seems from the discussion above that the NSC examination papers in physical sciences are of low cognitive demand and needs to be addressed.

5.1.2 Research question 2.

How do the cognitive demands of the NSC physical sciences examination question papers compare with the weighting of cognitive levels recommended in the national curriculum statement?

I compared the findings from my research question one with the Department of Education (2008a) recommendations. The physical science paper one (physics) analysis reveals that the “retrieval” category was marginally over-emphasised, “basic comprehension” was under-emphasised, “routine execution” and “application” were over-emphasised while the “analysis and reasoning” category was non-existent in all papers. The physical science paper two (chemistry) analysis reveals that the “retrieval” category was over-emphasised, “basic comprehension” was under-emphasised, “routine execution” and “application” were over-emphasised while the “analysis and reasoning” category was once again non-existent in all the question papers. These findings indicate that although the recommendations are emphasised in policy documents, they are not compliant in practice. There appears to be a need to ensure that examiners comply with recommendations so that standards regarding NSC physical sciences examinations are maintained.

5.2 LIMITATIONS

This section addresses the limitations of this study. There is a generally accepted idea of cognitive levels of questions amongst teachers in the subject. The raters in this study were trained in the cognitive demand taxonomy and aspired to be as objective as possible. However, the analysis of examination papers is subjective to a marginal degree as every person conducting the analysis enters with his/her personal bias and experience and this may affect how some questions are categorised. There is a possibility that other raters may have varied in their analysis to some degree. In this study, this was addressed by having an experienced team of subject experts with more than twenty years of teaching experience and extensive marking experience to conduct the analysis. Inter-rater reliability that was calculated was good. This ensured rater consistency and the reliability of the cognitive demand taxonomy used to conduct the analysis.

The time period to conduct this study was short and resources were limited. Further resources and more time may have been beneficial to the development of the cognitive demand taxonomy and the understanding of the cognitive demand of examination assessment items. The researcher could have consulted with recent NSC examiners to improve the validity of the taxonomy. Candidates who scored “A” symbols in physical sciences could have shared their examination preparation techniques. This would have provided a new insight into the value of an “A” symbol. This study may serve to catalyse discussions by experienced researchers, science education academics, assessment experts, examiners and curriculum specialists. They may use this study as a starting point to engage in discussions and further research in this area to find common understanding of cognitive domain categories in taxonomies used to measure cognitive demand of the NSC high-stakes examination.

5.3 IMPLICATIONS

The findings of this study may be relevant to physical sciences teachers and all stakeholders in examination and assessment. The research problem identified in this study was that the NSC examination does not afford learners adequate opportunity to engage in higher order thinking. The findings have made us aware that the NSC physical sciences examination question papers were of low cognitive demand. This indicates that the idea of developing individuals to think critically and apply higher order reasoning skills that are valued by higher education institutions and the workplace are not encouraged at NSC level. An “A” symbol in NSC physical sciences examination is associated with the mastery of lower order thinking skills like “retrieval”, “basic comprehension” and “routine execution”. Learners are able to reproduce definitions, explain familiar concepts, observations and phenomena, execute routine procedures and answer routine problems. It does not necessarily imply that the “A” candidate can make new connections and apply the knowledge and skills gained to new situations or engage in problem solving, decision making, analysis and reasoning at an abstract level. The coaching methods employed by some teachers merely teach to the test and do not stimulate the cognitive faculties in learners. The ideals of scientific reasoning, problem solving, critical thinking, investigating, generating and testing of hypotheses that are enshrined in the curriculum and assessment policy documents appear to be neglected in the NSC physical sciences examinations.

The literature survey indicated that while some studies like Crowe (2012), Edwards (2010), Liu et al. (2008), Muller (2005), Näsström (2009) and Umalusi (2009), focused on standards, alignment and cognitive demand, there were limited in-depth studies on the cognitive demand of physical sciences examination questions. The findings of this study may stimulate discussions with relevant stakeholders to address the issue of the low cognitive demand of NSC physical science examination question papers. This study analysed all final public NSC examination question papers from 2008 to 2013. The findings may make future examiners and stakeholders tasked with the NSC examinations under the Curriculum and Assessment Policy Statement (Department of Basic Education, 2011a), aware of the low cognitive demand of NSC physical sciences examinations. The study may make stakeholders focus on the use of cognitive demand taxonomies to analyse question papers and ensure tests and examinations are keeping with recommended weightings regarding cognitive demands.

5.4 RECOMMENDATIONS

A critical component of standards of examinations is cognitive demand. From my experience as a physical sciences teacher and from this research study, I would recommend the following:

- Intensive training to be provided by experts in the area of cognitive demand to help subject advisors, teachers and other stakeholders to better understand this neglected but important aspect in the setting of examinations and assessments.
- A common understanding of categories like “routine execution”, “application” and “analysis and reasoning” to be developed and adopted by curriculum and assessment stakeholders.
- A rethink of cognitive weightings of examinations after consultation with higher education institutions, employers and assessment experts.

- Teachers and examiners must ensure that learners are afforded the opportunity to engage in higher order thinking at every stage of their development, be that in the classroom and or in the examination room.
- A balance of lower order and higher order skills and knowledge should be taught and tested at all junior grade levels. The cognitive demand of the NSC exit high school examination cannot be examined in isolation and the ANA question papers should contain an adequate amount of higher order questions so that learners learn to apply the knowledge and skills to new situations from junior grades.

By attending to the above, we will be in sync with the progressive economies that drive teaching for higher order thinking so that learners are able to cope with the high level challenges of the modern workplace.

5.5 CONCLUSION

High stakes examinations appear to increasingly dominate assessment methods globally. From the earliest recorded Chinese Mandarin civil service written tests to the current NSC exit examination in South Africa, the consequences of the results obtained for the candidates are life altering. It is therefore imperative that examiners, quality assurance moderators and other stakeholders ensure that the examination seeks out to measure what the curriculum and assessment policy purports.

Debates regarding the standards of the senior certificate examinations have been ongoing since the 1900s. The percentage pass rate of the NSC examinations has improved over the past six years. However, the improved percentages do not reveal whether the examination standards are improving or declining. The Department of Basic Education (2014b), while pleased with performance improvements states that it is worrisome that candidates are unable to answer physical sciences examination questions “which are somewhat outside the box although not too difficult as it points to a serious flaw in the thinking skills of candidates” (p. 185). This statement by the DBE lends support to evidence presented in this study with regard to low levels of cognitive skills tested in the examination. The literature reviewed has also pointed out that some scholars are of the view that grade inflation does not give learners a true indication as to the skills and knowledge that they possess and that actual standards may have fallen. Maintaining acceptable standards is essential and has implications on the learners' entrance to universities, other institutions of higher learning and the workplace. Alignment between assessment policy and the examinations is essential to ensure that the skills and knowledge tested are of acceptable cognitive demand standards.

Large-scale studies like TIMMS gives us an idea of how poorly South Africa compares with other countries in mathematics and science and this may stimulate educational reform to emphasising higher order thinking skills in assessment tasks like the NSC. Studies reviewed have indicated that teaching higher order thinking skills may lead to better performances and economic productivity. The findings in this study in contrast point to a de-emphasis on higher order thinking skills and an over-emphasis on lower order thinking skills in the NSC physical sciences examinations. The value of an "A" symbol in NSC physical sciences appears to be over-rated as the study findings have indicated that on average 94% of the questions emphasised lower order thinking skills of "retrieval", "basic comprehension" and "routine execution". When compared to the DBE recommendations on cognitive demand weightings, it was found that the TASSK categories of "basic comprehension" was under-emphasised, "retrieval" and "routine execution" were over-emphasised while questions that involved the higher order thinking skills of "analysis and reasoning" were not found in all the examination question papers analysed.

We can thus expect an "A" candidate to have mastered familiar definitions, laws in science, familiar explanations and routine problems and procedures. An "A" symbol in the NSC physical sciences examination does not guarantee that the learner can apply the knowledge and skills learned to new situations requiring critical thinking, non-routine problem solving, decision making, investigating and other higher order cognitive applications. The incorporating of higher order thinking skills as part of the written, taught and assessed curriculum is recommended as an area of further research. The study of cognitive demand in physical sciences examinations appears under-researched and this study together with further research in this area may add value to the existing pool of knowledge.

REFERENCES

Alexander, R. (2010). *The perils of policy: Success, amnesia and collateral damage in systematic educational reform*. Paper presented at the Miegunyah Distinguished Lecture in the Dean's Lecture Series, Melbourne Graduate School of Education, Melbourne, Australia. Retrieved May 6, 2014 from http://www.edfac.unimelb.edu.au/news/lectures/2010series.html

Allen, A. (2012). Cultivating the myopic learner: The shared project of high stakes and low stakes assessment. *British Journal of Sociology of Education, 33*(5), 641-659.

Amer, A. (2006). Reflections on Bloom's revised taxonomy. *Journal of Research in Educational Psychology, 4*, 213-230.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., . . . Wittrock, M. C. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

Black, P. (1998). Assessment by teachers and the improvement of students' learning. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 811-819). New York, NY: Kluwer.

Black, P. (2004). Purposes of assessment. In J. Gilbert (Ed.), *The RoutledgeFalmer Reader in science education* (pp. 189-199). London: RoutledgeFalmer.

Black, P., & Harrison, C. (2010). Formative assessment in science. In J. Osborne & J. Dillon (Eds.), *Good practice in science teaching. What research has to say* (pp. 183-210). Berkshire: McGraw Hill.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives. Handbook 1: Cognitive domains*. New York, NY: David McKay Company, Inc.

Braun, H., & Kanjee, A. (2006). Using Assessment to Improve Education in Developing Nations In H. Braun, A. Kanjee, E. Bettinger & M. Kremer (Eds.), *Improving education through assessment, innovation and evaluation* (pp. 1-46). Cambridge, MA: American Academy of Arts and Sciences.

Britton, E. D., & Schneider, S. A. (2007). Large-scale assessments in science education. In S. K. Abell & N. G. Lederman (Eds.), *Handbook on research in science education* (pp. 1007-1025). New York, NY: Routledge.

Caldwell, B. J. (2010). *The impact of high stakes test driven accountability*. Paper presented at the National Symposium, Australian Education Union, University of Technology, Sydney, Australia.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). London: Routledge.

Creswell, J. W. (2003). *Research Design: Qualitative, quantitative and mixed methods approaches* (pp. 2-23). London: Sage.

Crowe, A. A. (2012). *Standards of South African senior certificate biology examinations: 1994 to 2007.* Unpublished doctor of philosophy thesis, University of Cape Town, Cape Town, South Africa.

Department of Basic Education. (2009a). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2009*. Pretoria: Goverment Printer.

Department of Basic Education. (2009b). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2009*. Pretoria: Goverment Printer.

Department of Basic Education. (2010a). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2010*. Pretoria: Goverment Printer.

Department of Basic Education. (2010b). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2010*. Pretoria: Goverment Printer.

Department of Basic Education. (2011a). *Curriculum and assessment policy statements F.E.T. Grades 10-12*. Pretoria: Government Printer.

Department of Basic Education. (2011b). *Report on the qualitative analysis of Annual National Assessments 2011 results*. Pretoria: Government Printer.

Department of Basic Education. (2011c). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2011*. Pretoria: Goverment Printer.

Department of Basic Education. (2011d). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2011*. Pretoria: Government Printer.

Department of Basic Education. (2011e). *Report on the National Senior Certificate examination 2011: National diagnostic report on learner performance*. Pretoria: Government Printer.

Department of Basic Education. (2012a). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2012*. Pretoria: Goverment Printer.

Department of Basic Education. (2012b). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2012*. Pretoria: Government Printer.

Department of Basic Education. (2013a). *National Senior Certificate Examination Technical Report 2012*. Pretoria: Government Printer.

Department of Basic Education. (2013b). *National Senior Certificate Examination, Physical Sciences: Physics (P1) March 2013*. Pretoria: Government Printer.

Department of Basic Education. (2013c). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2013*. Pretoria: Goverment Printer.

Department of Basic Education. (2013d). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2013*. Pretoria: Government Printer.

Department of Basic Education. (2014a). *2013 National Senior Certificate Technical Report*. Pretoria: Government Printer.

Department of Basic Education. (2014b). *2013 National Senior Certificate examination, National diagnostic report*. Pretoria: Government Printer.

Department of Education. (2003). *National Curriculum Statement, Grades 10-12, Physical Sciences*. Pretoria: Government Printer.

Department of Education. (2006). *National Curriculum Statement, Grades 10-12, Physical Sciences Content*. Pretoria: Government Printer.

Department of Education. (2008a). *National Curriculum Statement, Grades 10-12, Physical Sciences, Subject Assessment Guidelines*. Pretoria: Goverment Printer.

Department of Education. (2008b). *National Senior Certificate Examination, Physical Sciences: Physics (P1) November 2008*. Pretoria: Goverment Printer.

Department of Education. (2008c). *National Senior Certificate Examination, Physical Sciences: Chemistry (P2) November 2008*. Pretoria: Goverment Printer

Department of Education. (2009). *Physical Sciences Examination Guidelines, Grade 12*. Pretoria: Government Printer.

Doran, R. L., Lawrenz, F., & Helgeson, S. (1994). Research on assessment in science. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning. A project of the national science teachers association* (pp. 388-411). New York, NY: Macmillan.

Edwards, N. (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.

Facione, P. (1990). *Critical Thinking: A statement of expert concensus for the purposes of educational assessment and instruction*. Newark, NJ: American Philosophical Agency.

Facione, P. (2013). *Critical Thinking: What it is and why it counts*. Retrieved June 10, 2013 from http://www.insightassessment.com/pdf_files/what&why2010.pdf

Foxcroft, C. D. (2006). Evaluating the school-leaving examination against measurement principles and methods. In V. Reddy (Ed.), *Marking matric: Colloquium Proceedings* (pp. 58-71). Pretoria Human: Sciences Research Council.

Govender, J., & Moodley, M. (2012). Efffect of national senior certificate maths on the pass rate of first-year engineering physics. *South African Journal of Science, 108*(7), 1-5.

Grayson, K., & Rust, R. (2001). Interrater reliability. *Journal of Consumer Psychology, 10*(1 & 2), 71-73.

Gronlund, N. E., & Waugh, C. K. (2009). *Assessment of student achievement* (9th ed.). Upper Saddle River, NJ: Pearson.

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: Sage Publications, Inc.

Harlen, W. (2010). Purposes and procedures for assessing science process skills. *Assessment in Education*, *6*(1), 129-144. Retrieved May, 10, 2013 from http://dx.doi.org/10.1080/09695949993044

Henningsen, M., & Stein, M. K. (1997). Mathematical tasks and student cognition: Classroom-based factors that support and inhibit high-level mathematical thinking and reasoning. *Journal for Research in Mathematics Education*, *28*(5), 524-549.

Herman, J. L., & Baker, E. L. (2009). Assessment policy. In G. Sykes, B. Schneider & D. N. Plank (Eds.), *Handbook of education policy research* (pp. 176-190). New York, NY: Routledge.

Hess, K. K. (2006). *Exploring cognitive demand in instruction and assessment*. Retrieved February 2, 2013 from http://www.nciea.org/publication_PDFs?DOK_ApplyingWebb_KH08.pdf

Hobden, P. A. (2002). *Reclaiming the meaning of problem solving: The need for a common understanding of the terms problem and problem solving.* Paper presented at the Annual meeting of South African Association of Research in Mathematics, Science and Technology Education, Durban.

Howie, S. (2001). *Mathematics and science performance in grade 8 in South Africa 1998/1999 TIMSS-R*. Pretoria: Human Sciences Research Council.

Human Sciences Research Council. (1998). *Media brief: SA matrics fared worst in international maths and science study*. Pretoria: Human Sciences Research Council.

Human Sciences Research Council. (2012). *Media brief: SA and Honduras at bottom of TIMSS rankings*. Retrieved June 3, 2013 from http://www.hsrc.ac.za/en/media-briefs/education-and-skills-development/Mathematics-and-science-levels-in-TIMSS-2011-shows-improvement-though-still-low

Human Sciences Research Council. (2013). *Improving teaching and learning through the South African Annual National Assessment: Challenges,possibilities and solutions*. Retrieved June 5, 2013 from http://www.hsrc.ac.za/uploads/pageContent/3324/04%20Teaching%20and%20Learning.pdf

Jansen, J. (2012). *Opinion analysis: Grade 12 results 2012*. Retrieved February 14, 2013 from http://www.pmi-sa.co.za/documents/Grade_12_Results_2012.pdf

Johnson, D., & Johnson, B. (2009). *High-stakes testing*. Retrieved May 20, 2014, from http://www.education.com/reference/article/high-stakes-testing1/#B

Kreitzer, A. E., & Madaus, G. F. (1994). Empirical investigations of the hierarchical structure of the taxonomy. In L. W. Anderson & L. A. Sosiak (Eds.), *Bloom's Taxonomy: A forty-year retrospective, Ninety-third yearbook of the National Society for the study of education, Part II* (pp. 64-81). Chicago, IL: University of Chicago Press.

Kubiszyn, T., & Borich, G. (2010). *Educational testing and measurement* (9th ed.). New York, NY: John Wiley & Sons Inc.

Kuhn, S. V. (2011). Exploring the use of statewide exit exams to spread innovation. The example of context in science tasks from an international comparitve perspective. *Studies in Education Evaluation, 37*, 189-195.

KwaZulu-Natal Department of Education. (2014). *National Senior Certificate 2013*. Durban: Government Printer.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lane, S. (2004). *Validy of high-stakes assessment: Are students engaged in complex thinking?* Paper presented at the Presidential Address to the National Council on Measurement in Education, San Diego, CA.

Leighton, J. P. (2013). Item difficulty and interviewer knowledge on effects on the accuracy and consistency of examinee response processes in verbal reports. *Applied Measurement in Education, 26*(2), 136-157.

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice, 32*(3), 131-137.

Liu, X., Zhang, B., Liang, L., Fulmer, G., Kim, B., & Yuan, H. (2008). *Alignment between the physics content standard and the standardized test: A comparison among the United States-New York State, Singapore, and China-Jiangsu.* Retrieved March 27, 2013 from doi:10.1002/sce.20330

Lolwana, P. (2006). The history of falling matric standards. In V. Reddy (Ed.), *Marking matric: Colloquium Proceedings* (pp. 18-32). Pretoria: Human Sciences Research Council.

Looney, J. W. (2009). Assessment and innovation in education. *OCED Education Working Papers* (Vol. 24). Paris: OECD publishing.

Madaus, G. F., & Kellaghan, T. (1996). Curriculum evaluation and assessment. In P. W. Jackson (Ed.), *Handbook of research on curriculum: A project of the American Educational Research Association* (pp. 119-154). New York, NY: Macmillan Library Reference.

Maree, K. (Ed.). (2007). *First steps in research*. Pretoria: van Schaik.

Marzano, R. J., & Kendall, J. S. (2007). *The new taxonomy of educational objectives* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory into Practice, 42*(4), 226-232.

McMillan, J. H. (2011). *Classroom assessment. Principles and practice for effective standards-based instruction* (5th ed.). New York, NY: Pearson.

Mertens, D. M. (2005). *Research methods in education and psychology: Integrating diversity with quantitaive and qualitative approaches*. Thousand Oaks, CA: Sage.

Mhlolo, M. K. (2011). *From coherence in theory to coherence in practice: A stock-take of the written, tested and taught national curriculum statement for mathematics (NCSM) at further education and training (FET) level in South Africa.* Unpublished doctor of philosophy thesis, University of Witwatersrand, Johannesburg, South Africa.

Muller, J. (2005). *The challenge of cognitive demand*. Paper presented at the Umalusi Seminar. Theme: "Matric: What is to be done?", UNISA Senate Hall. Pretoria.

Multon, K. D. (2010). Interrater reliability. In N. J. Salkind (Ed.), *Encyclopedia of Research Design* (pp. 627-629). Thousand Oaks, CA: Sage.

Naidoo, M. (2006). Transition from senior certificate to the further education and training certificate. In V. Reddy (Ed.), *Marking matric: Colloquium Proceedings* (pp. 10-17). Pretoria: Human Sciences Research Council.

Näsström, G. (2009). Interpretation of standards with Bloom's revised taxonomy: a comparison of teachers and assessment experts. *International Journal of Research & Method in Education, 32*(1), 39-51.

National Research Council. (2012). *A framework for K-12 Science Education: Practices, crosscutting concepts and core ideas.* Retrieved June 11, 2013 from http://www.nap.edu/catalog.php?record_id=13165.

Nel, C., & Kistner, L. (2009). The National Senior Certificate: Implications for access to higher education. *South African Journal of Higher Education, 23*(5), 953-973.

Orkin, F. M. (2006). Foreword. In V. Reddy (Ed.), *Marking matric: Colloquium Proceedings* Pretoria: Human Sciences Research Council.

Pithers, R. T., & Soden, R. (2000). Critical thinking in education. *Educational Research, 42*(3), 237-249.

Reddy, V. (2006). South African science achievement in an international context *Mathematics and Science achievement at South African Schools in TIMMS 2003*. Pretoria: Human Sciences Research Council.

Resnick, L. B., Rothman, R., Slattery, J. B., & Vranek, J. F. (2011). Benchmarking and alignment of standards and testing. *Educational Assessment, 9*(1), 1-27.

Schofield, H. (1972). *Assessment and testing: An introduction*. London: Allen & Unwin.

Schraw, G., & Robinson, D. R. (2011). Assessment of higher order thinking skills. In G. Schraw & D. R. Robinson (Eds.), *Conceptualizing and assessing higher order thinking skills*. Charlotte, NC: Information Age Publishing.

Squires, D. A. (2009). *Curriculum alignment. Research-based strategies for increasing student achievement* (pp. 3-60). London: Sage.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research & Evaluation, 9*(4) Retrieved March 11, 2014 from http://PAREonline.net/getvn.asp?v=9&n=4

Swain, J. (2010). Summative assessment. Gold or glitter? In J. Osborne & J. Dillon (Eds.), *Good practice in science teaching. What research has to say* (pp. 211-237). Berkshire: McGraw Hill.

Teodorescu, R. E., Bennhold, C., Feldman, G., & Medsker, L. (2013). *New approach to analyzing physics problems: A taxonomy of introductory physics problems*. Retrieved from doi:10.1103/PhysRevSTPER.9.010103

Trumpelmann, M. H. (1991). *The Joint Matriculation Board: Seventy five years achievement in perspective*. Cape Town: National Book Printers.

Umalusi. (2004). *Investigation into the Senior Certificate Examination*. Pretoria: Umalusi.

Umalusi. (2007). *Making educational judgements. Reflections on judging standards of intended and examined curricula*. Pretoria: Umalusi.

Umalusi. (2009). *From nated 550 to the new national curriculum: Maintaining standards in 2008. Part 3 Exam paper analysis*. Pretoria: Umalusi.

Vinjevold, P. (2005). *Improving the quality and cognitive demand of the National Senior Certificate: Making the case for the assessment tail wagging the teaching/learning dog*. Paper presented at the Umalusi Seminar. Theme: "Matric: What is to be done?" UNISA Senate Hall. Pretoria.

wa Kivilu, M. (2006). The matriculation examination: how can we find out if standards are falling? In V. Reddy (Ed.), *Marking matric: Colloquium Proceedings* (pp. 33-44). Pretoria: Human Sciences Research Council.

Webb, N. L. (1997). *WISE Brief: Determining alignment of expectations and assessments in mathematics and science education*. Retrieved June 4, 2013 from http://www,wcer.wisc.edu/archive/nise/publications/briefs/vol_1No_2.pdf.

Wheeler, G. F. (2000). The three faces of inquiry. In J. Minstrell & E. H. van Zee (Eds.), *Inquiring into Inquiry Learning and Teaching in Science*. Washington, DC: American Association for the Advancement of Science.

Wilson, L. D. (2007). High stakes testing in mathematics. In F. K. Lester Jr (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1099-1110). Charlotte, NC: Information Age Publishing.

APPENDICES

APPENDIX A: ETHICAL CLEARANCE CERTIFICATE

UNIVERSITY OF KWAZULU-NATAL
INYUVESI YAKWAZULU-NATALI

05 February 2014

Mr Nagesh Munsamy (212558711)
School of Education
Edgewood Campus

Protocol reference number: HSS/1189/013M
Project title: What does an "A" symbol in Physical Sciences represent about a learner's skills and knowledge in the subject? A study of the cognitive demand of the National Senior Certificate Physical Sciences examination question papers

Dear Mr Munsamy,

Full Approval – Expedited

In response to your application dated 12 November 2013, the Humanities & Social Sciences Research Ethics Committee has considered the abovementioned application and the protocol have been granted **FULL APPROVAL.**

Any alteration/s to the approved research protocol i.e. Questionnaire/Interview Schedule, Informed Consent Form, Title of the Project, Location of the Study, Research Approach and Methods must be reviewed and approved through the amendment/modification prior to its implementation. In case you have further queries, please quote the above reference number.

Please note: Research data should be securely stored in the discipline/department for a period of 5 years.

I take this opportunity of wishing you everything of the best with your study.

Yours faithfully

....
Dr Shenuka Singh (Chair)

/ms

Cc Supervisor: Professor Paul Hobden
cc Academic Leader Research: Dr MN Davids
cc School Administrator: Mr Thoba Mthembu

Humanities & Social Sciences Research Ethics Committee
Dr Shenuka Singh (Chair)
Westville Campus, Govan Mbeki Building
Postal Address: Private Bag X54001, Durban 4000
Telephone: +27 (0) 31 260 3587/8350/4557 **Facsimile:** +27 (0) 31 260 4609 **Email:** ximbap@ukzn.ac.za / snymanm@ukzn.ac.za / mohunp@ukzn.ac.za
Website: www.ukzn.ac.za

1910 - 2010
100 YEARS OF ACADEMIC EXCELLENCE

Founding Campuses: Edgewood, Howard College, Medical School, Pietermaritzburg, Westville

APPENDIX B: BLOOM'S TAXONOMY (Bloom et al., 1956)

Cognitive Domain

1.00 Knowledge

1.10 Knowledge of specifics
1.11 Knowledge of terminology
1.12 Knowledge of specific facts

1.20 Knowledge of ways and means of dealing with specifics

1.21 Knowledge of conventions
1.22 Knowledge of trends and sequences
1.23 Knowledge of classifications and categories
1.24 Knowledge of criteria
1.25 Knowledge of methodology

1.30 Knowledge of universals and abstractions in a field.

1.31 Knowledge of principles and generalizations
1.32 Knowledge of theories and structures

INTELLECTUAL ABILITIES AND SKILLS

2.00 Comprehension

2.1 Translation
2.2 Interpretation
2.3 Extrapolation

1.0 **Application**

2.0 **Analysis**

4.1 Analysis of elements
4.2 Analysis of relationships
4.3 Analysis of organizational principles

3.0 Synthesis

5.1 Production of a unique communication
5.2 Production of a plan, or proposed set of operations
5.3 Derivation of a set of abstract relations

4.0 Evaluation

6.1 Judgements in terms of internal evidence
6.2 Judgments in terms of external criteria

APPENDIX C: THE REVISED BLOOM'S TAXONOMY (COGNITIVE DOMAIN)

(Anderson et al., 2001)

1.0 **Remember:** Retrieving relevant knowledge from the long-term memory.
- 1.1 Recognizing
- 1.2 Recalling

2.0 **Understand:** Determining the meaning of instructional messages. Including oral and graphic communication.
- 2.1 Interpreting
- 2.2 Exemplifying
- 2.3 Classifying
- 2.4 Summarizing
- 2.5 Inferring
- 2.6 Comparing
- 2.7 Explaining

3.0 **Apply:** Carrying out or using a procedure in a given situation.
- 3.1 Executing
- 3.2 Implementing

4.0 **Analyze**: Breaking material into constituent parts and detecting how the parts relate to one another and to an overall structure or purpose.
- 4.1 Differentiating
- 4.2 Organizing
- 4.3 Attributing

5.0 **Evaluate:** making judgment based on criteria and standards.
- 5.1 Checking
- 5.2 Critiquing

6.0 **Create:** Putting elements together to form a novel, coherent whole or make an original product.
- 6.1 Generating
- 6.2 Planning
- 6.3 Producing

APPENDIX D: DEPARTMENT OF EDUCATION (2008a) PHYSICAL SCIENCES ASSESSMENT TAXONOMY

The following table provides a hierarchy of cognitive levels that can be used to ensure tasks include opportunities for learners to achieve at various levels and tools for assessing the learners at various levels. The verbs given in the fourth column below could be useful when formulating questions associated with the cognitive levels given in the first column.

COGNITIVE DESCRIPTION	EXPLANATION	SKILLS DEMONSTRATED	ACTION VERBS
EVALUATION	At the extended abstract level, the learner makes connections not only within the given subject area but also beyond it and generalises and transfers the principles and ideas underlying the specific instance. The learner works with relationships and abstract ideas.	Compares and discriminates between ideas. Assesses value of theories, presentations. Makes choices based on reasoned arguments. Verifies value of evidence Recognises subjectivity	Assess, decide, rank, grade, test, measure, recommend, convince, select, judge, explain, discriminate, support, conclude, compare, summarise, critique, appraise, interpret, justify
SYNTHESIS	The learner works at the extended abstract level but makes errors because he or she is insufficiently informed at more modest levels	Uses old ideas to create new ones. Generalises from given facts. Relates knowledge from several areas. Predicts and draws conclusions.	Combine, integrate, modify, rearrange, substitute, compare, formulate, prepare, generalise, rewrite, categorise, combine, compile, reconstruct, organise, revise, what if?
ANALYSIS	The learner appreciates the significance of the parts in relation to the whole. Various aspects of the knowledge become integrated; the learner acquires deeper understanding and the ability to break down a whole into its component parts. Elements	Sees patterns and the organisation of parts. Recognises hidden meanings. Identifies parts of components	Analyse, separate, order, explain, connect, classify, arrange, divide, compare, select, infer, breakdown, contrast, distinguish, diagram, illustrate, identify, outline, point out, relate

	embedded in a whole are identified and the relations among the elements are recognised.		
APPLICATION	The learner establishes a relational construct but which has errors. The learner has the ability to use (or apply) knowledge and skills in new situations.	Uses information, methods, concepts and theories in new situations. Solves problems using required skills or knowledge	Apply, demonstrate, calculate, complete, illustrate, show, solve, examine, modify, relate, change, classify, experiment, discover, construct, manipulate, prepare, produce
COMPREHENSION	A number of connections may be made but the meta-connections are missed, as is the significance for the whole. The learner has first level of understanding, recalls and understands information and describes meaning.	Understands information and grasps meaning. Translates knowledge into new contexts and interprets facts. Compares, contrasts, orders, groups and infers causes and predicts consequences	Summarise, describe, interpret, contrast, predict, associate, distinguish, estimate, differentiate, discuss, extend, comprehend, convert, defend, explain, generalise, give example, rewrite, infer.
RECALL	Simple and obvious connections are made. The learner recalls and remembers facts.	Observes and recalls information	List, define, tell, describe, identify, show, know, label, collect, select, reproduce, match, recognise, examine, tabulate, quote, name

APPENDIX E: DEPARTMENT OF BASIC EDUCATION (2011a) PHYSICAL SCIENCES ASSESSMENT TAXONOMY

The following table provides a hierarchy of cognitive levels that can be used to ensure tasks include opportunities for learners to achieve at various levels and tools for assessing the learners at various levels. The verbs given in the fifth column below could be useful when formulating questions associated with the cognitive levels given in the first column.

COGNITIVE DESCRIPTION	LEVEL	EXPLANATION	SKILLS DEMONSTRATED	ACTION VERBS
CREATING	**4**	The learner creates new ideas and information using the knowledge previously learned or at hand. At the extended abstract level, the learner makes connections not only within the given subject area but also beyond it and generalises and transfers the principles and ideas underlying the specific instance. The learner works with relationships and abstract ideas.	Generating Planning Producing Designing Inventing Devising Making	Devise, predict, invent, propose, construct, generate, make, develop, formulate, improve, plan, design, produce, forecast, compile, originate, imagine
EVALUATING	**4**	The learner makes decisions based on in-depth reflection, criticism and assessment. The learner works at the extended abstract level.	Checking Hypothesising Critiquing Experimenting Judging Testing Detecting Monitoring	Combine, integrate, modify, rearrange, substitute, compare, prepare, generalise, rewrite, categorise, combine, compile, reconstruct, organise, justify, argue, prioritise, judge, rate, validate, reject, appraise, judge, rank, decide, criticise
ANALYSING	**3**	The learner appreciates the significance of the parts in relation to the whole. Various aspects of the knowledge become integrated, the learner shows a deeper understanding and the ability to break down a whole into its component parts. Elements embedded in a whole are identified and the relations among the elements are	Organising Comparing Deconstructing Attributing Outlining Finding Structuring Integrating	Analyse, separate, order, explain, connect, classify, arrange, divide, compare, select, infer, break down, contrast, distinguish, draw, illustrate, identify, outline, point out, relate, question, appraise, argue, defend,

		recognised.		debate, criticise, probe, examine, investigate, experiment.
APPLYING	**3**	The learner has the ability to use (apply) knowledge and skills in other familiar situations and new situations.	Implementing Carrying out Using Executing	Apply, demonstrate, calculate, complete, illustrate, show, solve, examine, modify, relate, change, classify, experiment, discover, construct, manipulate, prepare, produce, draw, make, compile, compute, sequence, interpret
UNDERSTANDING	**2**	The learner grasps the meaning of information by interpreting and translating what has been learned	Interpreting Exemplifying Comparing Explaining Inferring Classifying	Summarise, describe, interpret, contrast, associate, distinguish, estimate, differentiate, discuss, extend, comprehend, convert, explain, give example, rewrite, infer, review, observe, give, main idea
REMEMBERING	**1**	The learner is able to recall, remember and restate facts and other learned information	Recognising Listing Describing Identifying Retrieving Recalling Naming	List, define, tell, describe, identify, show, know, label, collect, select, reproduce, match, recognise, examine, quote, name

APPENDIX F: THE PERFORMANCE EXPECTATIONS TAXONOMY (Crowe, 2012)

LOWER-ORDER COGNITIVE SKILLS Content, procedure or connections explicitly listed in the syllabus or given in the question.			**HIGHER–ORDER COGNITIVE SKILLS** Use learned content, procedures or make connections in ways not explicated in the syllabus or given in the question	
ROTE–AND–ROUTINE (i.e., no demonstration of understanding required)		DEMONSTRATE UNDERSTANDING (i.e., using acquired knowledge and skills)		
ACQUIRE important information and skills	MAKE MEANING of important information and skills		TRANSFER meaning to new situations	
A Memorise (i.e., knowing)	B Perform Routine Procedures (i.e., doing)	C Explain- demonstrating a basic understanding of memorised knowledge and routine procedures	D Analyze information / make connections, not required by the syllabus or given in the question, using memorized knowledge and routine procedures in familiar contexts (i.e., of the syllabus or given in the question.)	E Apply (use) concepts / analyse information / make connections in new contexts (i.e., outside of syllabus and not given in the question)
1. Recall / recognize science terms, facts, definitions, concepts. 2. Recall / recognize scientific formulae	3. Make measurements 4. Make a scientific drawing. 5. Make observations / describe objects, processes, results. 6. Read values / information from graphs. 7. Compute. 8. Use given formulae. 9. Use / assemble / handle apparatus 10. Conduct routine / explained experiments	15. Explain / show understanding of learned concepts/ routine procedures / processes.	16. Observe and explain student / teacher / given demonstrations. 17. Explain methods of science and inquiry. 18. Classify and compare data (similarities and differences) 19. Analyze data, recognize patterns / trends. 20. Reason inductively / deductively. 21. Draw conclusions 22. Identify faulty	25. Generate questions / hypotheses or make predictions from unlearned / experimental data 26. Select, use and integrate science concepts / formulae / routines 27. Test the effect of different variables. 28. Recognise experimental design, errors / appropriate use of controls 29. Synthesize content and ideas from several sources. 30. Plan and design an investigation/ experiment to address a given/ generated problem or question or

	11. Test the effects of different variables in routine experiments. 12. Collect and record data. 13. Organize and display data in tables / graphs/ charts as instructed. 14. Neatness and presentation of work.		arguments or misrepresentations of data 23. Generate questions or make predictions from prescribed knowledge or routine procedures 24. Present analysed information / results	hypothesis. 31. Organize and display data in tables, graphs or charts of own design. 32. Reason inductively / deductively 33. Apply and adapt science information to real-world situations 34. Build or revise a plan/ theory 35. Present applied concepts and connections 36. Construct an argument

APPENDIX G: TAXONOMY FOR THE ASSESSMENT OF SCIENTIFIC SKILLS AND KNOWLEDGE (TASSK)

REPRODUCTIVE THINKING			PRODUCTIVE THINKING	
LOWER-ORDER THINKING SKILLS			HIGHER ORDER THINKING SKILLS	
Retrieval	Basic Comprehension	Routine Execution	Application	Analysis / Reasoning
Level 1	Level 2	Level 3	Level 4	Level 5
The learner is only able to recall, restate, remember, recognise, identify, list, write or name: facts, terms, definitions, phenomena, units, laws, diagrams, processes, concepts, routine equations, symbols, formulae. A demonstration of understanding or reasoning is not necessarily required of the learner.	The learner demonstrates a basic understanding of standard familiar concepts, practical work, graphs and diagrams by identifying, combining, separating, translating, discussing, explaining or interpreting the relevant information given in the question.	The learner is able to conduct routine processes, procedures, answer familiar questions and solve routine problems which may involve algorithms (a set of specific steps performed in a particular order)	The learner has the ability to use the knowledge and skills learned and apply it to new contexts. The problems / questions that fall in this category require the learner to think beyond the routine type and apply greater understanding and reasoning to solve the problems / questions which may require logical explanations, procedures or set of multiple steps not necessarily performed in a certain (routine) order. Note: the questions in this category are more cognitively complex than the routine type, not necessarily more difficult.	The learner shows a deeper understanding and analytical reasoning to break down the very complex problem / question into its component parts. Higher order thinking has to be applied to make connections to answer /solve these unfamiliar questions or problems which may involve:- a) Problem solving - overcoming obstacles, barriers or limiting conditions to resolve the problem. b) Decision making - selecting between two or more alternatives c) Experimenting - generating and testing hypotheses for the purpose of understanding phenomena. d) Investigating, generating and testing hypotheses using logical arguments and reasoning

APPENDIX H: INFORMED CONSENT FORM

INFORMED CONSENT FORM **12 November 2013**

Dear Subject Expert

Thank you for offering to participate as a member of my expert group. I am a Master of Education (M Ed.) student in the School of Education, University of KwaZulu-Natal and my research focuses on the cognitive demand of the National Senior Certificate (NSC) examination papers. The aims of my study are to determine the cognitive demand of the NSC physical sciences examination question papers from 2008 to 2012. My research study topic is:-

What does an "A" symbol in Physical Sciences represent about a learner's skills and knowledge in the subject? A study of the cognitive demand of the National Senior Certificate (NSC) Physical Sciences examination question papers.

You were identified as a possible subject expert because of your extensive experience in teaching physical sciences at grade 12 level and your vast experience in marking external NSC scripts. You have indicated your interest in participating as a member of my subject expert group. Should you still agree to participate in this study, you will be required to sacrifice about 3 hours of your time over a period of one month in classifying NSC examination questions into various cognitive levels according to a taxonomy given to you. All examination paper analysis for this study should be done in your personal capacity as an experienced physical sciences educator and you may not use any of your employer's time and resources. Your agreement to participate in this study should not impact in any way on your responsibilities to your current job duties.

The potential benefits of your participation are that you may possibly be introduced to a new cognitive demand taxonomy and may gain more experience to classify questions in the different cognitive domains recommended by the NCS. There will be no financial expenses incurred as a result of your participation as all telephone calls and/or visits to you will be at my expense. Copies of question papers, marking memoranda, data grids and other information will be emailed to you. This study will not involve the use of any audio or video recordings. Your time as an expert in the project is voluntarily required and no monetary amounts will be paid to you in this regard.

In summary I am required to inform you that:-

- your participation in this study is voluntary
- your responses will be treated in a confidential manner and no limits apply
- your anonymity will be ensured
- a decision not to participate in this study will not result in any negative or undesirable consequences to you.
- a decision to participate allows you the freedom to withdraw from the research at any stage for

any reason and without any negative or undesirable consequences to you.

- the findings of the study will be made available to you on request.

I am conducting this research study under the supervision of Professor Paul A. Hobden and the contact details for me, my supervisor and the officer at research office for Human and Social Sciences at the College of Humanities appear below

------------------------ ------------------------------------

Nagesh Munsamy Prof. Paul A. Hobden, PhD

CONTACT DETAILS

Mr N. Munsamy
P.O. Box 696
Umzinto
4200
Tel 039 9741490
Cell: 0827385589
Email: nagz@telkomsa.net

Prof. Paul A. Hobden
School of Education
University of KwaZulu-Natal
P. Bag X03, Ashwood
3605
Office: 27-31-2603447
Cell: +27 825474031
Email: hobden@ukzn.ac.za

CONTACT DETAILS FOR HSSREC RESEARCH OFFICE:

Ms P. Ximba
Tel. 031 260 3587
Email ximbap@ukzn.ac.za

Please completed the section below and return to the researcher

The Researcher
c/o School of Education, UKZN

I, ... (full name of subject expert) hereby confirm that I understand the contents of this document and the nature of the research project, and I consent / do not consent to participating in this research study.

I further understand that, should I so desire, I am at liberty to withdraw from the research project at any time and for any reason whatsoever.

______________________________ _______________

SIGNATURE OF SUBJECT EXPERT DATE

APPENDIX I: RATERS' ANALYSIS OF NOVEMBER 2012 PAPER 1 BY COGNITIVE DEMAND LEVEL

Question Number	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Level						
1.1	1	1	1	1	1	**1**	1
1.2	1	1	1	1	1	**1**	1
1.3	1	1	1	1	1	**1**	1
1.4	1	1	1	1	1	**1**	1
1.5	1	1	1	1	1	**1**	1
2.1	2	2	2	2	2	**2**	2
2.2	2	2	2	2	2	**2**	2
2.3	2	2	3	3	2	**X 2**	2
2.4	3	3	3	3	3	**3**	2
2.5	2	2	2	2	2	**2**	2
2.6	2	2	2	2	2	**2**	2
2.7	2	2	2	2	2	**2**	2
2.8	1	1	1	1	1	**1**	2
2.9	3	3	3	3	3	**3**	2
2.10	2	2	2	2	2	**2**	2
3.1	1	1	1	2	2	**X 1**	1
3.2.1	3	3	3	3	3	**3**	4
3.2.2	3	3	3	3	3	**3**	5
3.3	3	4	3	4	4	**X 4**	6
4.1	2	2	2	2	2	**2**	2
4.2	1	1	1	1	1	**1**	2
4.3.	3	3	3	3	3	**3**	6

APPENDIX I: CONTINUED

Question Number.	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Levels						
4.4.1	2	2	2	2	2	**2**	2
4.4.2	3	3	3	3	3	**3**	2
4.4.3	2	3	3	3	2	**X 3**	3
5.1	2	2	2	2	2	**2**	3
5.2	1	1	1	1	1	**1**	2
5.3.1	3	3	3	3	3	**3**	4
5.3.2	3	3	3	3	3	**3**	4
6.1	1	1	1	1	1	**1**	1
6.2	2	2	2	2	2	**2**	1
6.3	2	2	2	2	2	**2**	1
6.4.1	3	3	3	3	3	**3**	3
6.4.2.	3	3	3	3	3	**3**	5
7.1	1	1	1	1	1	**1**	1
7.2	2	2	2	2	2	**2**	2
7.3.1	3	3	3	3	3	**3**	3
7.3.2	3	3	3	3	3	**3**	5
7.4	3	3	3	3	3	**3**	2
8.1	3	3	3	3	3	**3**	3
8.2	2	2	2	2	2	**2**	1
8.3	3	3	3	3	3	**3**	3
8.4.1	2	2	2	2	2	**2**	1
8.4.2	2	2	2	2	2	**2**	1

APPENDIX I: CONTINUED

Question Number.	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Level						
8.5.1	1	1	1	2	2	**X 1**	3
8.5.2	3	3	3	3	3	**3**	3
9.1.1	3	3	3	3	3	**3**	3
9.1.2	3	3	3	3	3	**3**	5
9.1.3	3	3	3	3	3	**3**	3
9.2.1	2	2	2	2	2	**2**	1
9.2.2.	3	3	3	3	3	**3**	3
9.2.3	2	2	2	2	2	**2**	2
9.2.4	2	2	3	3	3	**X 3**	3
10.1	1	1	1	1	1	**1**	1
10.2	1	1	1	1	1	**1**	2
10.3	3	3	3	3	3	**3**	5
10.4	2	1	2	2	2	**X 2**	1
11.1.1	2	2	2	2	2	**2**	1
11.1.2	2	2	2	2	2	**2**	1
11.1.3	2	2	2	2	2	**2**	1
11.2	1	1	1	1	1	**1**	2
11.3	2	2	3	2	2	**X 2**	1
11.4	3	3	3	3	3	**3**	4
11.5	2	2	2	2	2	**2**	1

APPENDIX J: RATERS' ANALYSIS OF NOVEMBER 2012 PAPER 2

Question Number	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Level						
1.1	1	1	1	1	1	**1**	1
1.2	1	1	1	1	1	**1**	1
1.3	1	1	1	1	1	**1**	1
1.4	1	1	1	1	1	**1**	1
1.5	1	1	1	1	1	**1**	1
2.1	2	2	2	2	2	**2**	2
2.2	3	3	3	3	3	**3**	2
2.3	2	2	2	2	2	**2**	2
2.4	2	2	2	2	2	**2**	2
2.5	3	3	3	3	3	**3**	2
2.6	3	3	3	3	3	**3**	2
2.7	2	2	2	2	2	**2**	2
2.8	2	2	2	2	2	**2**	2
2.9	1	1	2	1	1	**X 1**	2
2.10	1	1	1	1	1	**1**	2
3.1.1	1	1	1	1	1	**1**	1
3.1.2	3	3	3	3	3	**3**	2
3.1.3	2	2	2	2	2	**2**	1
3.1.4	1	1	1	1	1	**1**	1
3.1.5	1	1	1	1	1	**1**	1
3.2.1	3	3	3	3	3	**3**	2
3.2.2	3	3	3	3	3	**3**	2
3.3.1	2	1	1	2	2	**X 2**	1
3.3.2	3	3	3	3	3	**3**	2
3.3.3	3	3	3	3	3	**3**	2
4.1.1	1	1	1	1	1	**1**	1
4.1.2	1	1	1	1	1	**1**	1
4.2.1	2	2	2	2	2	**2**	1
4.2.2	2	2	2	2	2	**2**	1
4.2.3	2	2	2	2	2	**2**	2
4.3	1	2	1	2	2	**X 2**	1
4.4	1	3	3	3	1	**X 1**	3
4.5	2	2	2	2	2	**2**	4

APPENDIX J: CONTINUED

Question Number	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Level						
5.1.1	2	2	2	2	2	**2**	1
5.1.2	3	3	3	3	3	**3**	2
5.2.1	2	2	1	2	2	**X 2**	1
5.2.2	3	3	3	3	3	**3**	2
5.3.1	1	1	1	1	1	**1**	1
5.3.2	2	2	2	2	2	**2**	1
5.3.3	3	3	3	3	3	**3**	2
5.4	3	3	3	3	3	**3**	2
6.1	1	1	1	1	1	**1**	1
6.2.1	2	2	2	2	2	**2**	1
6.2.2	2	2	2	2	2	**2**	1
6.2.3	2	2	2	2	2	**2**	1
6.3	2	2	2	2	2	**2**	2
6.4.1	2	2	2	2	2	**2**	1
6.4.2	2	2	2	2	2	**2**	1
6.4.3	2	2	2	2	2	**2**	1
6.5	1	1	1	1	1	**1**	2
6.6	3	3	3	3	3	**3**	5
7.1	1	1	1	1	1	**1**	2
7.2	3	3	3	3	3	**3**	3
7.3	3	3	3	3	3	**3**	3
7.4	4	4	4	4	4	**4**	9
8.1.1	2	2	2	2	2	**2**	2
8.1.2	2	2	2	2	2	**2**	2
8.1.3	2	3	3	3	3	**X 3**	3
8.2.1	2	2	2	2	2	**2**	1
8.2.2	2	3	2	3	3	**X 3**	3
8.2.3	3	3	2	3	3	**X 3**	4
8.2.4	2	2	2	2	2	**2**	2
9.1	1	1	1	1	1	**1**	1
9.2	2	2	2	2	2	**2**	1
9.3	2	2	2	2	2	**2**	1
9.4.1	2	2	2	2	2	**2**	2
9.4.2	2	2	2	2	2	**2**	1
9.5	2	2	2	2	2	**2**	2
9.6	1	1	1	1	1	**1**	2
9.7	1	1	1	1	1	**1**	1

APPENDIX J: CONTINUED

Question Number	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5	Final Rating	Marks
	Cognitive Level						
10.1	1	1	1	1	1	**1**	1
10.2	2	2	2	1	2	**X 2**	2
10.3	2	1	2	2	2	**X 2**	1
10.4.1	3	3	3	3	3	**3**	4
10.4.2	3	3	3	3	3	**3**	3
11.1.1	1	1	1	1	1	**1**	1
11.1.2	1	1	1	1	1	**1**	3
11.1.3	1	2	2	2	2	**X 2**	1
11.1.4	1	3	1	3	1	**X 1**	3
11.1.5	1	1	1	1	1	**1**	1
11.2.1	1	1	1	1	1	**1**	1
11.2.2	1	1	1	1	1	**1**	2

APPENDIX K: RESEARCHER'S ANALYSIS OF NOVEMBER 2011 PAPER 1 BY COGNITIVE LEVEL

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
1.1	1					1
1.2	1					1
1.3	1					1
1.4	1					1
1.5	1					1
2.1	2					2
2.2		2				2
2.3			2			2
2.4		2				2
2.5		2				2
2.6		2				2
2.7		2				2
2.8		2				2
2.9	2					2
2.10	2					2
3.1		1				1
3.2			4			4
3.3			4			4
3.4			5			5
4.1		2				2
4.2			2			2
4.3	2					2
4.4			6			6
4.5			5			5

APPENDIX K: CONTINUED

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
5.1	2					2
5.2		3				3
5.3	1					1
5.4			5			5
6.1	1					1
6.2			4			4
6.3		1				1
6.4		2				2
7.1	2					2
7.2		2				2
7.3		2				2
7.4			2			2
7.5			5			5
8.1		1				1
8.2			3			3
8.3			2			2
8.4				6		6
9.1		1				1
9.2.1		2				2
9.2.2		2				2
9.3			4			4

APPENDIX K: CONTINUED

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
10.1		1				1
10.2.1			3			3
10.2.2			3			3
10.3			5			5
10.4				3		3
11.1.1	1					1
11.1.2	1					1
11.1.3	1					1
11.1.4	1					1
11.2		2				2
11.3			6			6
12.1	1					1
12.2	1					1
12.3			4			4
12.4			4			4
12.5.1		1				1
12.5.2		1				1
12.6.1	1					1
12.6.2	1					1
Marks	27	36	78	9	0	150
%	18	24	52	6	0	

APPENDIX L: RESEARCHER'S ANALYSIS OF NOVEMBER 2011 PAPER 2

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
1.1	1					1
1.2	1					1
1.3	1					1
1.4	1					1
1.5	1					1
2.1	2					2
2.2	2					2
2.3		2				2
2.4			2			2
2.5		2				2
2.6		2				2
2.7		2				2
2.8	2					2
2.9		2				2
2.10		2				2
3.1.1		1				1
3.1.2		1				1
3.2.1			2			2
3.2.2			2			2
3.3	2					2
3.4.1	1					1
3.4.2	1					1
3.4.3		2				2
3.4.4			2			2

APPENDIX L: CONTINUED

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
4.1	1					1
4.2.1		1				1
4.2.2		1				1
4.2.3		1				1
4.3		3				3
4.4.1		1				1
4.4.2			2			2
4.5.1			2			2
4.5.2		3				3
4.6		2				2
5.1.1		1				1
5.1.2		1				1
5.1.3		1				1
5.2			4			4
5.3		2				2
5.4		2				2
6.1	1					1
6.2		2				2
6.3		2				2
6.4		1				1
6.5		2				2
6.6		2				2
6.7			2			2

APPENDIX L: CONTINUED

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
7.1.1	3					3
7.1.2			3			3
7.1.3		2				2
7.2.1			8			8
7.2.2			3			3
8.1	1					1
8.2	1					1
8.3			2			2
8.4		1				1
8.5			3			3
8.6		1				1
8.7			4			4
8.8		4				4
9.1	2					2
9.2.1		2				2
9.2.2		2				2
9.3		2				2
9.4.1			3			3
9.4.2			1			1
10.1	1					1
10.2		2				2
10.3.1		1				1
10.3.2		1				1

APPENDIX L: CONTINUED

Question	Level 1	Level 2	Level 3	Level 4	Level 5	Marks
10.3.3		1				1
10.4	3					3
10.5		2				2
11.1.1	1					1
11.1.2	3					3
11.2	1					1
11.3	2					2
11.4			3			3
11.5	2					2
	37	65	48	0	0	150

APPENDIX M: TURNITIN RECEIPT

turnitin

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Paul Hobden
Assignment title:	MEd dissertation
Submission title:	MEd Thesis from Munsamy
File name:	1803_1412174344_60720.docx
File size:	586.39K
Page count:	130
Word count:	37,984
Character count:	208,516
Submission date:	01-Oct-2014 04:39PM
Submission ID:	459028137

What does an "A" symbol in Physical Sciences represent about a learner's skills and knowledge in the subject?
A study of the cognitive demand of the National Senior Certificate Physical Sciences examination question papers.

Nagesh Munsamy

Submitted in partial fulfilment of the academic
requirements for the degree of
Master of Education in the
School of Education
University of KwaZulu-Natal

September 2014

MEd Thesis from Munsamy

ORIGINALITY REPORT

16%	14%	6%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

I have checked the full Turnitin report and there is no substantive copying without atribution. Majority of highlighted excerpts were from the tables of taxonomies which have citations. Prof Paul Hobden - Supervisor

APPENDIX N: GLOSSARY

Some of the terminology used in this study are discussed in this section while others were discussed in detail in the chapters and do not appear in the list below.

Algorithms	A set of rules performed in a particular order.
Assessment	A measure of a candidate's skills and knowledge in a particular task.
Assessment item	The smallest measurable unit in an assessment task.
Assessment standards	These are criteria that describe what a learner should know and be able to demonstrate at a specific grade (Department of Education, 2003).
NSC with bachelor's entrance	One of four levels of NSC passes that allows a candidate to apply to a higher education institution to pursue a bachelor's degree. The other levels of passes are: NSC, NSC with Higher Certificate entrance and NSC with Diploma entrance.
Candidate	An individual who is allowed to sit for an examination or assessment task.
Cognitive demand	The mental demand made on an individual in the answering of an assessment item.
Content standards	Content standards clarify what a learner should know, understand and apply in a particular learning area (McMillan, 2011).
Electives	A minimum of three subjects listed in the NCS document that are chosen by a learner in the Further Education and Training band starting at grade 10.
Examination	An assessment task administered in a highly controlled procedural and structured manner.
Fail	Candidates who have not met the requirements for promotion into the next grade or for an exit certificate are deemed to have failed. This term was used in the old system and has been replaced by "not achieved" in individual subjects and "not ready to progress" for grade progression.
Grade	The level of schooling, starting with grade R (reception year), followed by grade 1 to grade 12 (final year of basic education) in the South African education system.

Higher Grade (HG)	The higher of the two levels at which a subject could have been taken in the old senior certificate examination prior to 2008.
Level of competence	This is referred to as grades in some countries. Previously in South Africa a symbol (A to H) indicating the level of competence was written on the candidate's senior certificate. This has now been replaced with a rating code (1 to 7) indicating the level of competence.
Marks	Points or scores given per assessment item and totalled for the assessment task.
Marking	The process of awarding marks or points to candidates' responses in an assessment task using a marking guideline, memorandum or rubric.
Matric	This term commonly used in South Africa refers to grade 12 (previously referred to as standard 10). A candidate who passes this grade exit exam is said to have matriculated.
Matriculation endorsement	A level of passing the old Senior Certificate (matriculation/matric) examination that allows the candidate to apply to study for a bachelor's degree at a University.
National Senior Certificate	The newly named grade 12 exit examination qualification introduced in 2008 that replaced the old senior certificate.
National Qualifications Framework	The National Qualifications Framework (NQF) in South Africa is a qualifications' system that categorises qualifications or training from level one (grade 9) to level 10 (doctorate degree).
Pass	A term used when a candidate meets the minimum requirement for competency in an assessment task. This is currently 30% for a subject like physical sciences at NSC level in South Africa.
Pass rates	The rate at which a cohort of learners pass an examination or assessment task. The grade 12 NSC pass rate announcement generates much debate early every year in South Africa.
Quality	Describes a worthy product or service of value to society and associated with high standards.
Routine Execution	The execution of a familiar task using familiar procedures and algorithms.
Score	The total number of marks or points achieved in an assessment task.

Senior Certificate	The old grade 12 (matric) certificate obtained prior to 2008. This has been replaced by the NSC.
Standard Grade	The lower of the two levels that a subject could be offered in the old senior certificate examination. This has been replaced with all subjects in the NSC written at one level since 2008.
Subject	A body of skills and knowledge integrated with attitudes and values in a particular academic discipline.
Symbol	Referred to as grades in some countries, a symbol (A to H) was used in the old senior certificate examination to represent different levels of competence in a subject. This use of term "symbols" continues to be used orally with the NSC; however policy documents refer to it as rating codes (1 to 7).
Task	An informal or formal learning activity given to learners in order to gauge their understanding of a topic.
TASSK	The taxonomy for the assessment of scientific skills and knowledge developed in this study and used to measure the cognitive demand of assessment items
Taxonomy	An instrument developed to categorise assessment items in terms of cognitive demand.
Test	A structured assessment task used to make judgements on a learner's understanding of a topic(s) in a subject.

M. Govender, M.A. (English)

P.O. Box 223, Umzinto, 4200, South Africa cell 082 940 19 50
Email: mickgovender@gmail.com

23 September 2014

DECLARATION OF EDITING

Master of Education dissertation by Nagesh Munsamy (212 558 711)

Topic: What does an "A" symbol in Physical Sciences represent about a learner's skills and knowledge in the subject? A study of the cognitive demand of the National Senior Certificate Physical Sciences examination question papers.

Submitted in partial fulfilment of the academic requirements for the degree of Master of Education in the School of Education, University of KwaZulu-Natal.
Supervisor: Professor Paul Hobden

I hereby declare that I carried out the language editing of the above dissertation by Nagesh Munsamy. I am a professional English Home Language educator and freelance editor with more than 40 years of work experience.

Yours sincerely

MOONSAMY GOVENDER