Alternative assessment strategies within a context-based science teaching and learning approach in secondary schools in Swaziland

Victoria Louise Kelly

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophiae in the Faculty of Education, University of the Western Cape**.**

Supervisors

Professor Dirk Meerkotter

Dr. Lorna Holtman

Professor Øyvind Mikalsen

April 2007

ABSTRACT

Assessment in education systems serves many functions that influence teaching and learning in significant ways. Teachers' implementation of curricula and evaluation of students' learning, as well as the effort students invest in their learning and assessment activities are influenced by how they view assessment. As such, assessment is undergoing a transformation from psychometric models to educational assessment models, where assessment and instruction are integrated to support learning. Alternative assessment models and a transformed view of assessment are central to the success of the educational assessment models.

The aim of this study was to use a case study approach to explore and describe how students and teachers perceived performance assessment and context-based assessment models that were used within a real world context teaching and learning approach. The topics Electricity and Air and Living Things formed the science knowledge base for the study. Four junior secondary school science teachers and their students in four schools participated. Participants' experiences of the assessment models were achieved through teachers administering and scoring performance assessment tasks and context-based unit tests to their students. Perceptions were obtained through questionnaires and interviews from students. Interviews and informal discussions were used to elicit teachers' perceptions. Observations during the administration of performance assessment tasks were also used for triangulation.

Based on the findings of the research, I concluded that:

- Students and teachers held similar and varied, as well as positive and negative perceptions about the assessment models used. Perceptions were, however, influenced by the use of a group assessment mode of administration and teacher mentorship in the case of performance assessment.

- Both performance assessment tasks and context-based questions were perceived to engage students in deep thinking while addressing the tasks/questions. Performance assessment tasks motivated students to increase their participation and work with greater commitment when performing the tasks and other lesson activities.

- Assessment models were also perceived to improve knowledge and understanding, develop useful skills and increase awareness of real world occurrences and practices. These models also developed students' eagerness to be more observant in their surroundings.

In addition to the above, it became evident to me that performance assessment tasks were made easier by their practical nature, familiar contexts, teacher mentorship and group work. Furthermore, group assessment improved the quality of answers and scores. However, allocation of equal marks to individual group members was perceived to be unfair because of differentiated participation. Participants also attached much significance to scores as a motivating factor for students. This perceived importance of scores has implications for the success of a view of assessment that emphasises learning and active engagement rather than scores.

I also found that context-based questions undoubtedly invited the use of "informal" knowledge in answers. Teachers perceived such answers as an indication that students believed in the acceptability of these "informal" answers, used the context in questions as a source of information in arriving at answers, and had difficulty linking scientific concepts used in the context to scientific concepts learned in the classroom. Teachers felt challenged to establish the degree of correctness of students' answers in view of expected "scientific" answers.

Recommendations from this study are that students and teachers need training and more experience in the use of these assessment models so that interferences of lack of practice in working with the assessment models may be reduced.

KEYWORDS

Alternative; Assessment; Context-based; Education; Perceptions; Performance; Secondary; School; Science; Swaziland

DECLARATION

I declare that ***Alternative assessment strategies within a context-based science teaching and learning approach in secondary schools in Swaziland*** is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been acknowledged by complete references.

Signed: ………………………………..

Victoria Louise Kelly April 2007

ACKNOWLEDGEMENTS

I would like to acknowledge individuals and institutions for their contribution towards making my study a success.

My thanks go to:

My three supervisors, Professor Dirk Meerkotter, Dr. Lorna Holtman and Professor Øyvind Mikalsen, for the valuable support and guidance throughout the study.

The Graduate Studies in Science and Mathematics Education (GRASSMATE) Project anchored at UiB, Norway and UWC, South Africa for the financial support it provided for me to pursue this study.

The coordinators of the project and the Postgraduate Enrolment and Throughput (PET) Project for organising our stay at UWC and other venues where we held our meetings with the supervisors, and Professor Stein Dankert Kolstoe for showing me how to use the ATLAS.ti programme.

The schools, teachers and students who participated in this study. Their commitment, patience and loyalty throughout the study were appreciated.

The University of Swaziland for granting me training leave in the last stages of the project. Thanks also go to the In-Service Education Department for allowing me to use the laboratories for workshops and meetings and for making available the science equipment I used for the performance assessment tasks.

Colleagues from the University of Swaziland who provided assistance of whatever nature. Special thanks go to Dr. Betty Dlamini and Dr. Sabelo Manyatsi for proofreading the thesis and for encouraging me all the way.

Members of my family for their patience, understanding, support and encouragement throughout the duration of this study.

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS

A level	Advanced level
AfL	Assessment for learning;
AoF	Assessment of learning;
ECOS	Examinations Council of Swaziland
GCE	General Certificate of Education
INSET	In-service Education and Training
IGCSE	International General Certificate of Secondary Education
LTA	Learning through assessment
LISSIT	Linking School Science with Industry and Technology
NCC	National Curriculum Centre
NSTP	National Science Teaching Panel
O level	Ordinary level
SSTA	Swaziland Science Teachers' Association
TTC	Teacher Training College
UCLES	University of Cambridge Local Examinations syndicate
UNISWA	University of Swaziland
SWISP	Swaziland Integrated Science Programme
WISCIP/C	West Indies Science Curriculum Innovation Project, "C" version
FGI	Focus Group Interview

1. CHAPTER 1

INTRODUCTION TO THE THESIS

1.1 INTRODUCTION AND MOTIVATION

> Assessment "is often poorly understood, its purpose confused and its design inadequate" (Goldstein and Lewis 1996:ix).

In my experience of being a teacher and working with teachers through in-service and pre-service teacher education, I observed practices that were of interest in the way assessment was used and its relationship to teaching and learning. I have observed, for example during teaching practice, experienced teachers pursuing the bachelor of education programme setting tests and revising these tests with their students. Of interest in the practice of these teachers was that when scripts were returned to students the teachers revised the tests by simply discussing the answers to the questions without any attempts to identify sources of the errors students had made.

My interest was further developed through the experience of setting and marking of the Integrated Science national examinations for the junior secondary school level. The examinations I set in the two years of my contract as external examiner were inclined towards getting students to show their understanding of scientific concepts and procedures, by including questions on the uses of scientific information in everyday life. The examination papers were checked and accepted by a team of moderators as required by the Examinations Council of Swaziland (ECOS). Teachers, however, commented that the examinations were difficult because of the number of application level questions contained in the papers. Teachers who participated in the marking of candidates' scripts also complained that they had not taught their students some of the things asked in the questions. Now that Swaziland has changed to a context-based teaching and learning approach my interest in the uses of scientific information in everyday life has developed even further.

I chose to conduct this research in the area of assessment in order to develop a deeper understanding of assessment in education. The focus on the topic for this study came as a result of preliminary readings around issues on context-based teaching and learning and their link to assessment. My reading in the area provided information about the possible function of assessment in promoting students' learning as a direct outcome of engaging in assessment tasks. This promotion of learning was in addition to the mastery of subject content by students from revision exercises and practice of the skill of answering assessment questions. It is clear that assessment activities could also serve as learning activities during which students could develop new knowledge and skills.

New developments in assessment practices have taken place that draw attention to the weaknesses of traditional assessment models such as multiple choice questions and other content-based assessment tasks. These developments have given rise to alternative assessment models that go beyond assessing students' recall of content knowledge to assessing more complex abilities. I came across alternative assessment models, such as performance assessment, the assessment of portfolios and classroom embedded assessment while reviewing the literature on assessment. An idea that appealed to me was that alternative models of assessment could be integrated with instruction to support learning. I became interested in two alternative assessment models: performance assessment and context-based assessment because of their link to practical work and real world contexts. I was also interested in how teachers and students would view the use of these models in their classrooms. These assessment models are introduced in section 1.3 below.

In view of the experiences and interest noted above, this chapter presents the aim of the study, context and background to the study, statement of the problem, research questions and significance of the study. In addition, background information on the significance of assessment in education, characteristics of educational assessment, assessment and perceptions about assessment, and the education system in Swaziland are provided. The conceptual framework and some limitations of the study are presented in this chapter.

1.2 AIM OF THE STUDY

The aim of this study was to explore and describe the perceptions and experiences of science students and teachers regarding the use of alternative strategies for assessing learning in secondary school science. Performance assessment and context-based assessment models were conducted in a context-based teaching situation to provide the context in which the perceptions and experiences were explored.

1.3 ASSESSMENT IN EDUCATION

Educational assessment influences education systems and their endeavours in profound ways. Goldstein and Lewis (1996) note that assessment affects various aspects of an education system such as curriculum design practices, education policies, institutions of learning, teachers, their teaching and assessment styles, as well as students and their learning and studying styles. Educational assessment is thus a fundamental component of the education process, in particular teaching and learning as activities of curriculum implementation (Gipps and Stobart 2003; Shepard 2000). Summarising the importance of assessment, Broadfoot said:

> Assessment is arguably *the* most powerful policy tool in education. … and will probably continue to be the single most significant influence on the quality and shape of students' educational experience and hence their learning (1996:21 and 22, original emphasis).

As important as assessment is perceived to be, different people view it in different ways. For example, Kotzé (2003) cites assessment as an encompassing concept that is associated with concepts like measurement, evaluation, testing, standards and criteria. Another view is that assessment encompasses a wide range of procedures that are used to obtain information about students' learning such as observations, ratings of performance or projects, as well as paper-and-pencil tests that is used to form value judgements about students' learning progress. According to these views assessment encompasses the process and the tools used to collect information about students' learning and the judgements made from that information.

Measurement and evaluation are important processes in assessment. According to Popham (1999) measurement is traditionally used in reference to the process of assigning numbers to test results or other performance and establishing how much of the quality or attribute being measured is possessed by an individual. Evaluation on the other hand involves the interpretation of, or inferences made from, the result to judge that quality. These views of assessment are grounded in the traditional standardised models of assessment. Sanders and Horn (1995) note that traditional standardised assessment models rely on tests like multiple choice and/or closed questions that are administered, scored and interpreted in a standard manner. Their preparation is a rigorous process of validation by content and assessment experts, reliability testing and other standardisation procedures and criticism before they are used in assessing students' competencies (Popham 1999). Standardised tests are thus more appropriate for use in high stakes assessment activities where their outcomes are useful when comparing students, generalising about their achievement, indicating levels of attainment based on set standards and informing education policy decisions.

Methods of assessment that are based on traditional testing models have been under intense criticism for their inadequacy in measuring higher order thinking skills, and intellectual and manipulative processes (Sanders and Horn 1995). They have also been criticised for their effects on instruction and the curriculum (Gipps and Stobart 2003). Moorcroft, Desmrais, Hogan and Berkowitz (2000) also observed that traditional models of testing encourage a view that there is one correct answer to every question and a view that learning by rote memorisation and recall of memorised information in tests is a good way of learning. Objective type questions like those often used in multiple choice items tend to emphasise rote recall. Such weaknesses of traditional models of assessment have led to a search for other ways of appraising students' learning, commonly referred to as alternative assessment.

A new way of viewing assessment is emerging and is calling for a paradigm shift in assessment from a "testing culture to an assessment culture", from "psychometrics to educational assessment" (Gipps 1994:158). According to Sambell, McDowell and Brown (1997) an assessment culture is demonstrated in classroom practices that favour the integration of assessment, teaching and learning processes, as well as students' active and informed participation in the assessment. In an assessment culture tasks that are authentic, meaningful and engaging are used. The tasks mirror realistic contexts in contrast with the artificial time constraints and limited access to support experienced in conventional examinations. Furthermore, assessment focuses on both the process and products of learning and not on the psychometric use that emphasises test scores. Gipps (1994) describes several qualities and characteristics of educational assessment. These are elaborated on in Section 1.3.1 below.

1.3.1 Educational assessment

The characteristics of educational assessment pertinent to this study are described by Gipps (1994). According to Gipps, educational assessment:

1. Requires clear standards or criteria against which students' levels of performance are measured. These standards are shared with students to encourage them to reflect on their performance and become self-monitoring learners.
2. Encourages students to think rather than pick alternative responses or reiterate facts, although a significant amount of teacher input to help students organise and integrate their ideas may be provided. They need good quality tasks, which are also anchored in subject matter perceived to be relevant and important.
3. Depends on appropriate tasks, which are also anchored in subject matter perceived to be relevant and important.
4. Uses concrete assessment tasks that are embedded in students' experiences to elicit their best performance. The tasks are administered in a non-threatening environment to reduce stress in students.
5. Involves grading by teachers who are well prepared for the tasks and understand the scoring criteria in relation to the levels of performance.
6. Involves interactive assessment which allows teachers to engage fully with students to gauge their understanding, scaffold the learning process and evaluate performance in a range of contexts.
7. Does not favour publicising the results at classroom or school level.

The above characteristics show another view of implementing assessment in education. That is, assessment that uses concrete and interactive tasks that are aligned to the curriculum and students' experiences. Assessment models that conform to the characteristics enumerated above are likely to capture a wider range of attributes than has been the case in traditional assessment models.

1.3.2 Alternative assessment models

The call for new assessment models is a call to expand assessment practices to include the assessment of skills and competencies that are important for daily living and that reflect ways in which knowledge and skills are used in real world contexts (Sharikzadeh 2003; Dietel, Herman and Knuth 1991). Assessment practices which require students to demonstrate their competences in more comprehensive ways are needed to complement traditional assessment practices currently in use. Gipps and Stobart (2003) and Veronesi (1997) note that alternative models of assessment can focus on both the learning processes and learning outcomes. They also possess other advantageous properties such as:

- Using challenging tasks that elicit higher order thinking.
- Allowing students to evaluate their own work.
- The possibility of integration into classroom instruction.

According to Linn and Gronlund (2000) and Moorcroft *et al.* (2000) alternative assessment models assume different formats, such as portfolio assessment and performance assessment. They range in complexity from simple extended multiple choice questions - restricted response format, to an in-depth scientific investigation - extended response format. Performance assessment and context-based assessment models are alternative assessment models that were of interest in the study.

Performance assessment requires of students to construct responses, create products or perform "live" practical activities to demonstrate their knowledge and skills (Linn and Gronlund 2000). They also assess procedural and conceptual knowledge, enabling students to show what they know and can do intellectually and physically. This study adopted the hands-on practical investigations model of performance assessment in which students manipulate scientific instruments to generate data from which they drew relevant conclusions (see Linn and Gronlund 2000; Brualdi 1998; Elliot 1995).

Context-based assessment models use open, interpretive items that describe scenarios of real life events in which scientific concepts are embedded. The assessment tasks require of students to use scientific knowledge to interpret and explain occurrences in real world situations presented in the questions. The assumption is that if students succeed in providing a scientifically sound answer to a context-based question then they have understood the scientific concepts embedded in the questions (Ahmed and Pollitt 2001).

1.4 ASSESSMENT AND PERCEPTIONS

Not only is there a call for more inclusive assessment models, but there is also a need to recognise the importance of students' and teachers' perceptions about assessment practices and formats. Perceptions about classroom processes of teaching and learning held by teachers and students affect implementation of new reforms in curricula and assessment in significant ways (Aschbacher 1993). These perceptions influence and maintain students' motivation in preparing, performing and persevering with tasks they see as important, useful and of value in their learning, development of skills and achievement (Ames and Archer 1988). Teachers' perceptions influence the way they use assessment to guide teaching and learning and the form of assessment they use to assess such learning.

Formal education in Swaziland uses assessment practices that serve different functions and which may influence or be influenced by classroom processes and perceptions of students and teachers. The following section gives the background to the education system, the curriculum and the organisation of assessment in Swaziland.

1.5 SWAZILAND EDUCATION SYSTEM

School education in Swaziland is organised into three levels of varied duration. These are the seven year primary school level with students whose age ranges from 6 to 12 (The Government of Swaziland 1985); the three year junior secondary school level with students aged from 13 to 15 and the two year senior secondary school (high school) level where students' ages range from 16 to17. These age ranges are not strictly adhered to as factors such as repetitions and financial constraints may lead to students of ages above these ranges remaining in a given level.

Post secondary school education is offered in various colleges and the University of Swaziland (UNISWA). Teachers receive training in any of four colleges, two of which specialise in primary school teacher training and one in secondary school teacher training. The fourth is a technical college. UNISWA trains teachers through the diploma, bachelor of education degree programmes and the post graduate certificate in education.

Financing education at school level is borne largely by the parents, while at the tertiary level funding is predominantly by the Government of Swaziland. Parents pay school fees and for other education needs of their children. The Government is responsible for all financial needs in State schools whilst the responsibilities of constructing school infrastructure (classroom, laboratories, and teachers' quarters) in non-Government schools are borne by the local communities whose children attend that school. The involvement of communities in the construction of school structures was and continues to be, an endeavour to provide the much needed teaching rooms and to curb the acute shortage of accommodation for teachers. The major expense for the Government in schools is the payment of teachers' salaries in all public schools.

Funding of school education has, in addition, been affected by the effects of HIV and AIDS. The Government of Swaziland, through the Ministry of Education has had to reconsider her responsibility of funding formal education as many children of school going age leave school after losing their parents to HIV and AIDS related diseases or other causes. The Government of Swaziland now provides bursaries for orphaned and vulnerable children at all levels of schooling, and also pays their national examinations fees across the different levels of schooling. As a signatory to the Dakar Convention on Education for All the Government of Swaziland introduced the Basic Education for All policy that is aimed at providing formal education and improving the access and retention of students in schools for the first ten years (The Government of Swaziland 2000a). Thus, subsidised education is offered to primary school students by providing school textbooks and stationery for all students from Grade 1 to Grade 4 (The Government of Swaziland 1995).

1.5.1 School science curriculum

a) Primary school science

At primary school level Science is taught as a separate subject from Grade 1 to Grade 7, since 1985. Prior to this it was taught as part of a subject called General Knowledge. The primary school curriculum is developed and managed at The National Curriculum Centre (NCC). The NCC drafts syllabi, develops teaching and assessment material and organises induction workshops for the teachers. The Centre also works in collaboration with and approval of a committee of science educators, the Primary Science Panel, which reports to the National Science Teaching Panel (NSTP) (Dlamini 2000). The NSTP is another committee that is responsible for school level issues that are related to science education. This committee is chaired by the Senior Inspector for Science and comprises representation from teacher training institutions, the University of Swaziland (UNISWA), the Swaziland Science Teachers' Association (SSTA), the NCC, and the Examinations Council of Swaziland (ECOS). The SSTA also organises the setting of mid-year examination papers for Grade 7, while ECOS is responsible for the primary school leaving national examinations.

b) Science at the junior secondary school level

At the junior secondary school level students take Integrated Science (until 2007), which is generally referred to as the Swaziland Integrated Science Programme (SWISP). SWISP is a three-year programme that is supported by a teaching and learning package that integrates content in Biology, Chemistry, Physics and some elements of Earth Science. SWISP was developed between 1971 and 1975 through the collaboration of the Ministries of Education and science teachers from Botswana, Lesotho and Swaziland (Putsoa, Manyatsi and Dlamini 2000). At the end of the three years students write the national examination. The examinations are set by examiners who are selected from teachers and other professionals from the tertiary institutions and the Ministry of Education by the National Science Teaching Panel (NSTP) and are forwarded to the Examinations Council of Swaziland (ECOS).

According to Lichtenstein (1980) the development of SWISP was motivated by dissatisfaction of the teachers with the Science that was being taught in schools at the time (i.e. prior to 1971). That programme was presented in two different syllabi: Syllabus A and Syllabus B. Syllabus A was intended for schools that had no laboratory facilities while Syllabus B was used by schools that had science laboratories. The displeasure in these syllabi arose from the following observations:

i) the syllabi had been written by non-local curriculum developers and without consultations with local educators. Teachers did not identify well with the material.
ii) the programme was never piloted prior to introduction, so teachers felt that the curriculum did not work.
iii) the teachers were also not happy that no laboratory equipment was provided for Syllabus B and schools offering that syllabus had very little or no science equipment. Teachers had to write the activities on the board for pupils to copy, and reproduce results when required to. Pupils learned Science without any experience of hands-on practical activities.
iv) The programme had no textbooks to guide the teachers.

Thus, the Swaziland Integrated Science Programme (SWISP) was developed with the aim of enabling

> … teachers in this country (Swaziland) to participate in the world-wide movement for the reform of science teaching by providing them with appropriate material and guidance (The Government of Swaziland (not dated):0-1 *sic*).

SWISP was developed by adapting the West Indies Science Curriculum Innovation Project, "C" version (WISCIP/C). The adaptation process incorporated local examples into the curriculum resulting in a localised science curriculum. In the early 1980s SWISP was revised to reduce teacher material and increase learner material (Manyatsi 1996). The programme has high orientation to practical work. SWISP materials come as a package consisting of pupil workbooks and teacher guides for each of the three years of the programme. All public schools offer SWISP.

c) Science at senior secondary school level

Subjects offered at the senior secondary school level come from the University of Cambridge Local Examinations Syndicate (UCLES) and lead to the General Certificate of Education (GCE) Ordinary (O) level certificate. The use of the Cambridge programme has been reviewed and has been replaced by the International General Certificate of Secondary Education (IGCSE).

The General Certificate of Education (GCE) offered Chemistry, Physics, Science (Biology and Chemistry), Science (Physics and Biology), Human and Social Biology, as well as the following subjects, which were offered in most of the schools in Swaziland.

a) Biology
b) Science (Physics and Chemistry): where the Chemistry and Physics content is contained in one syllabus although in separate sections;
c) Combined Science with all three sciences: Biology, Chemistry and Physics content in separate sections.
d) Additional Combined Science: with all three sciences: Biology, Chemistry and Physics content in separate sections.

Biology was generally taken together with Science (Physics and Chemistry). Additional Combined Science was taken in combination with Combined Science. It comprised a wider scope and depth of content from the three science subjects than did Combined Science. Biology and Science were not to be taken together with Combined Science because their content overlapped. Examinations for all options came from the University of Cambridge Local Examinations Syndicate (UCLES) but were supervised by Examinations Council of Swaziland (ECOS).

1.5.2 Curriculum reform

In the mid-1980s a National Curriculum Review Commission was set up and was assigned the task to review education in Swaziland. Among the recommendations from the commission was the re-organisation and diversification of the school curriculum to emphasise vocational subjects that would promote self-employment amongst school leavers. Since then the Ministry of Education has attempted to address problems of educational relevance, quality and accessibility as required by the Swazi society (The Government of Swaziland 1999). This arises from the realisation that the education system should respond to societal changes and address issues of quality, relevance, accessibility and variety in school curriculum options in addition to aligning it to regional and global developments in education (The Government of Swaziland 1995).

To this end the Government of Swaziland has started to pay attention to the provision of quality teaching resources and facilities, as well as ensuring affordable education.

As part of the curriculum review, syllabi were drawn up and teaching materials developed for nearly all subjects offered in primary and secondary school levels. The new science curriculum for primary school and junior secondary school levels has no indications of utilising assessment through practical work.

At the senior secondary school level the curriculum is also changing to a localised examinations system, which would also be cost effective (The Government of Swaziland 1999). According to The Government of Swaziland (2000b) the Ordinary (O) level certificate, although internationally recognised, was not a recognised entry requirement for South African tertiary institutions where a number of school leavers from Swaziland pursue further studies. Meanwhile, the local institutions of higher learning are no longer able to absorb all the school leavers who qualify for entry into these institutions. As a result many O level candidates have to study for the matriculation certificate or apply for matriculation exemption before they can enrol in certain South African tertiary institutions. The O level examinations system does not only constrain the education progress of many students but is also an economic strain for the country since students need to study for a matriculation certificate in addition to the O level certificate. Therefore, the Ministry of Education felt a need to explore other more affordable examining routes of acceptable standards to institutions of higher learning in the region. Such a move would thus broaden opportunities for further studies for secondary school graduates from Swaziland (The Government of Swaziland 2000b).

Another justification for the curriculum review was the need to incorporate aspects of scientifically based local knowledge and technologies into the school curriculum to promote awareness of these and their potential in developing the Swazi society. Local knowledge and technology had also been sidelined in the school curriculum in the past (The Government of Swaziland 1999).

Swaziland has thus changed from the GCE examinations to the International General Certificate of Secondary Education (IGCSE) examinations - also from the University of Cambridge. According to The Government of Swaziland (2005) GCE examinations were written for the last time in November 2006. The IGCSE programme was introduced in Form 4 (Grade 11) in January 2006 and examinations will be written for the first time in November 2007.

The use of the IGCSE syllabi and examinations will continue until the local syllabi are ready for implementation at the senior secondary school level. The senior secondary school level science subjects (Physical Science, Biology and Combined Science), selected for Swaziland, offer examination of practical work as part of the school leaving examinations.

1.5.3 Science curriculum review

Contextualising the school science curriculum in Swaziland is guided by the policy for Science from Grade 1 to Form 5 (Grade 12). The policy clause focusing on the contextualisation of science teaching is presented as follows:

> Relevance: Science education should draw extensively on the everyday experience of pupils, (Contextualised) and should be aimed at preparing pupils as effectively as possible for adult and working life. (The Government of Swaziland 1997a:3)

Such a curriculum shows the relationship and role of scientific knowledge acquired in school to everyday life and its relevance to the lives of the students. This curriculum is also intended to promote awareness of the potential of indigenous and local technologies in developing local Swazi societies (The Government of Swaziland 1997b).

A curriculum that integrates students' experiences and contexts into the teaching and learning of school science has been trialled at the junior secondary school level through two projects known as the Matsapha Lessons Project (1993 to 1995) and the Linking School Science to Industry and Technology (LISSIT) project (1996 to 1999). The main purpose of the two projects was to design and implement alternative teaching materials for the Swaziland Integrated Science Programme (SWISP) at Form II (Grade 9). These materials followed a context-based teaching approach and addressed the same learning outcomes as those of SWISP. Materials also emphasised the development of problem-solving skills and abilities to recognise scientific concepts that apply in a given everyday situation or industrial setting (Putsoa and Maphalala 1999). However, assessment has continued to bear the traditional content-based format. One study by Putsoa, Dlamini, Dlamini, Dlamini, Dube, Khumalo, Masango, Ndlela, Nhlengethwa and Tsabedze (2003) investigated the use of context-based assessment but the results were inconclusive. There is still insufficient information about what format of assessment should be used in the case of context-based learning.

1.5.4 Assessment in Science

In Swaziland assessing learning in science subjects takes place through two main approaches: internal school assessment and national exit examinations. Internal school assessment utilises teacher made monthly or regular topic tests, as well as end of year examinations given to non-external classes. External classes write an exit national examination as discussed in Section 1.5.1 above. Questions used in the internal examinations and tests are generally structured in a similar way to external examinations (and in many instances are simply prepared by picking relevant questions from past examinations).

With the exception of GCE Biology at the senior secondary school level, tests and examinations in the science subjects did not include assessment through practical work, although questions that were based on practical activities were sometimes included. In Biology a hands-on practical test or a written alternative to the practical test were two approaches used to assess students' practical competencies. The alternative to the practical paper can be viewed as a "practical theory paper". It assesses intellectual skills without physical manipulation of equipment, except perhaps the use of a ruler to measure something on a diagram or to assist with the drawing of diagrams. Intellectual skills such as extracting and interpreting data from diagrams (e.g. temperature or volume readings), drawing diagrams to scale, drawing and interpreting graphs are assessed. It requires students to have experience in practical work, which means student must be constantly exposed to practical activities during lessons. Experience from past performance of students in the two assessment approaches has indicated that students achieve better in the hands-on practical test than in the alternative paper.

1.6 RESEARCH PROBLEM

Research in context-based teaching and learning informed to a large extent the choice of a context-based science curriculum for Swaziland. As noted earlier studies on the assessment component have not been conclusive so far. The few studies done in Swaziland have been largely based on the Matsapha Lessons and Linking School Science with Industry and Technology (LISSIT) materials. From the use of these curriculum materials, Lubben, Campbell and Dlamini (1995) and Putsoa and Maphalala (1999) observed that utilising real life contexts in learning activities in Science gives students a purpose for engaging in learning of the subject matter at school and encourages them to appreciate its social and economic implications of Science. Lubben *et al.* (1995) further observed that contexts help students to make connections between the science taught and learned formally in classrooms and experiences beyond the classroom. Contexts also illustrate the applications, importance and limitations of Science.

These studies, in addition, show that context-based science teaching and learning approaches develop positive affective outcomes such as motivation, interest and empowerment of learners and increase learner participation in their own learning. This approach was found to be good in diagnosing students' misconceptions. However, the effect of the approach in developing students' understanding of basic scientific concepts was the same as that of students who were taught Science following conventional non-context-based teaching styles (Lubben *et al.* 1995). Similar findings were reported by Ramsden (1997). Context-based teaching achieves relevance in school science curricula by drawing on real life experiences of students. In their investigation Lubben *et al.* (1995) measured the effect of contextualising science teaching on students through the analysis of the relevance of the approach for teaching and learning, as well as observing classroom events (students' participation and interest) and assessment tasks (for students' understanding of concepts). Their study focused on content learning outcomes, classroom processes and learning.

Studies that focused on students' perceptions are exemplified by Dlamini, Lubben and Campbell's (1996) study on liked and disliked learning activities. The study established that students' motivation to learn in Science was enhanced by contexts perceived to have immediate or future personal links. A similar study by Dlamini (2003) compared students' views about context-based and non-context-based lesson activities through the use of worksheets. These students had been taught in a non-contextualised way. Students were found to prefer non-contextualised lesson activities. These students were comfortable with non-contextualised lesson activities because there were clear links between the lessons and the examinations, with less possibility of students' misunderstanding. The students recognised, however, that opportunities for cognitive development and application of scientific principles in their lives were enhanced by contextualised teaching.

A study that focused on assessment was conducted by Putsoa *et al.* (2003) who compared the performance of students in contextualised items and in non-contextualised equivalent items. The students had been taught in a non-contextualised way. The findings of this study were inconclusive. There was statistically insignificant improvement in the students' performance in the contextualised items.

The studies reported above showed that students who had been taught in a contextualised way performed in a similar way to those taught in a non-contextualised way in non-context-based content tests. They also show that students taught in a non-contextualised way but assessed using context-based questions and equivalent non-context-based questions perform in similar ways in the two sets of questions. Meanwhile, the other studies show that students appreciated context-based lesson activities or even liked activities that had personal relevance, but still preferred non-context-based activities. No study was found that related context-based teaching and learning to context-based assessment in terms of assessment of content knowledge or in ascertaining students' views about context-based assessment tasks.

Studies on the nature of students' responses to contextualised questions have shown that students react to such assessment questions in different, individual and unpredictable ways (Ahmed and Pollitt 2000). The contextualised questions invoke everyday images that help or hinder students' attempts at answering the questions (Wistedt 1994). When answering such questions students may ignore the context and use their subject matter knowledge in an acceptable way. Students may also take into consideration relevant aspects of the context before answering the questions and experience difficulties selecting the relevant content. This often causes them to under perform. As to which calibre of students take which approach as suggested by Wistedt (*ibid*) is not clear, but Ahmed and Pollitt (2000) observed that context-based questions favour high performing students.

The studies by Lubben *et al.* (1995), Putsoa and Maphalala (1999) and Putsoa *et al*. (2003) outlined above indicate that students' performance in non-contextualised tests is independent of whether the teaching was contextualised or not. They indicate that non-context-based teaching has the same effect on performance in non-context-based tests and context-based tests. The Matsapha Lessons and Linking School Science with Industry and Technology (LISSIT) materials used in the first two studies above employed a technological approach to lessons that emphasised practical work. Students carried out numerous practical activities. It was surprising that despite these activities the teaching approach still did not have a significant impact on students' understanding of scientific concepts. The apparent lack of impact of context-based teaching on students' achievement was of interest to this study. After all the positive effects context-based teaching seemed to have, what was preventing the students from demonstrating better understanding of science content?

Students' achievement in assessment is affected by extrinsic factors such as teaching approaches, assessment models, learning conditions and subject matter, as well as intrinsic factors like students' preferred learning and assessment styles, motivation, attitudes, interest and perceptions, to name a few. It has been illustrated in Section 1.4 above and Section 2.3 below that perceptions of assessment influence the effort students and teachers invest in learning, teaching and assessment activities. Perceptions influence how students react, process, and respond to questions. A search within the students and the teachers on how they viewed context-based assessment tasks appeared to be a reasonable starting point towards understanding the lack of impact of a context-based teaching approach on students' understanding of the content in Science.

Exploring the understanding of students' and teachers' perceptions of assessment requires a context within which students and teachers can experience and communicate their views about assessment. The context for this study involved context-based assessment and performance assessment models for a context-based teaching approach. There seems to be a dearth of information regarding performance assessment or views about it in Swaziland. Furthermore, most of the studies concentrated on the subject matter and how students' achievement was linked to the teaching approach or assessment approach. As noted in Section 1.5.4, the use of practical assessment in science subjects has been limited to Biology, and only at the senior secondary school level. Opportunities for assessing practical skills and procedures are therefore limited.

Performance assessment in the form of hands-on practical activities seems to have promise to support the practical nature of the contextualised teaching approach encouraged in Swaziland. The performance assessment model seems open to a complementary use with context-based assessment by integrating context-based questions to assess complex abilities of students. It seems to provide an alternative assessment model that emphasises students' demonstration of procedural knowledge and skills that were not easily assessed by the forms of assessment used in the studies referred to in this section.

This study therefore, set out to establish what educators could learn from the views and experiences of students and teachers regarding the use of these alternative assessment models in a contextualised science teaching setting in Swaziland. The assessment models were used to provide experiences from which participants could formulate and describe their perceptions.

1.7 RESEARCH QUESTIONS

In pursuing the aim of the thesis, namely to explore and describe the perceptions and experiences of science students and teachers regarding the use of performance assessment and context-based assessment models (see Section 1.2 above), this study sought answers to the following questions:

1. What perceptions and experiences do students and teachers have about the use of performance assessment as alternatives strategies for determining attainment of learning outcomes in Science?

2. How do students and teachers view the use of context-based assessment in assessing learning in Science?

1.8 SIGNIFICANCE OF THE STUDY

Through the National Science Teaching Panel (NSTP) the Ministry of Education in Swaziland mandated the development of a context-based curriculum. Research on context-based teaching and learning is still lacking, and studies on assessment are even rarer. There is a need for further exploration into the potential of context-based science teaching and learning for promoting understanding of basic scientific concepts and other higher order learning abilities and skills. This study provides additional information from the students' and teachers' perspectives on the use of alternative assessment models. It presents students' and teachers' perceived effectiveness and limitations of performance assessment, as well as context-based assessment for context-based teaching and school-based assessment as encouraged by the aims of science education (see The Government of Swaziland 1997) and the assessment framework (see The Government of Swaziland 2000a).

1.9 CONCEPTUAL FRAMEWORK

Instruction, learning and assessment are intricately linked classroom processes (Gipps 1994). This relationship is briefly represented in the top unboxed part of Figure 1.1 below. Figure 1.1 also shows that there is an interactive relationship between assessment, learning, teaching and the perceptions students and teachers hold about these classroom processes (Aschbacher 1993).

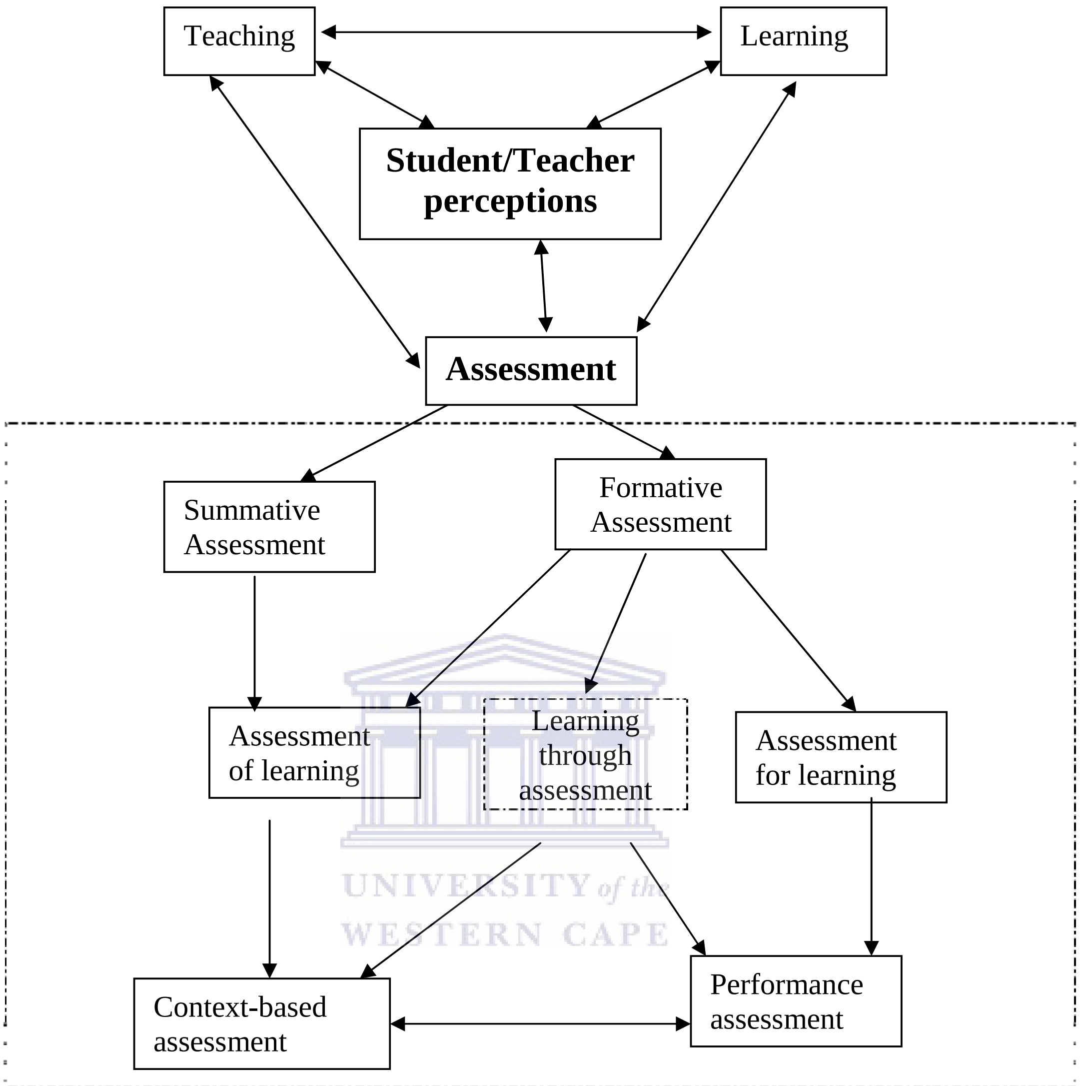


Figure 1.1 Relationships between different components of educational assessment, classroom processes and students' and teachers' perceptions

The "boxed" part of Figure 1.1 shows two main uses of assessment: formative assessment and summative assessment, in relation to students' learning and the assessment models used for the context of this study. Formative assessment is indicated as comprising assessment of learning and new views of assessment: assessment for learning (Gipps 1994) and learning through assessment. The functions of summative assessment and formative assessment overlap in the assessment of learning. The different ways of viewing formative assessment are discussed in sections 1.9.1 to 1.9.6 below.

The unboxed part of Figure 1.1 is expanded and developed into the conceptual framework of this study, shown in Figure 1.2 below. This framework was adapted from a conceptual framework proposed by Shepard (2000) for new views of assessment that are derived from an emergent re-conceptualisation of assessment that is based on cognitive and constructivist learning theories and a reformed vision of a curriculum.

Adaptation of the framework (shown in *italics* in Figure 1.2) involved redefining the reformed curriculum into a context-based curriculum and matching the principles of classroom assessment to performance assessment and context-based assessment models. Thus the framework in Figure 1.2 shows the interrelatedness of:

- a context-based curriculum (A) that emphasises relevance of science, authentic problem-solving and interest in Science;
- constructivist learning theories (B) that emphasise the importance of social contexts and prior knowledge, self-monitoring of thought processes and learning, in the construction of knowledge for deep understanding; and
- classroom assessment for a new paradigm of assessment for learning (Gipps 1994) (C), through assessment models (performance and context-based assessment) that are integrated with instruction to support learning and demand higher order thinking from students.

It also shows how the components represented in cells A, B and C are influenced by perceptions and beliefs held by teachers and students (D).

For each cell A, B C and D, the principles of the components of the framework are summarised and are further highlighted in the sections below. The principles of the components of the framework complement the alternative assessment approach.

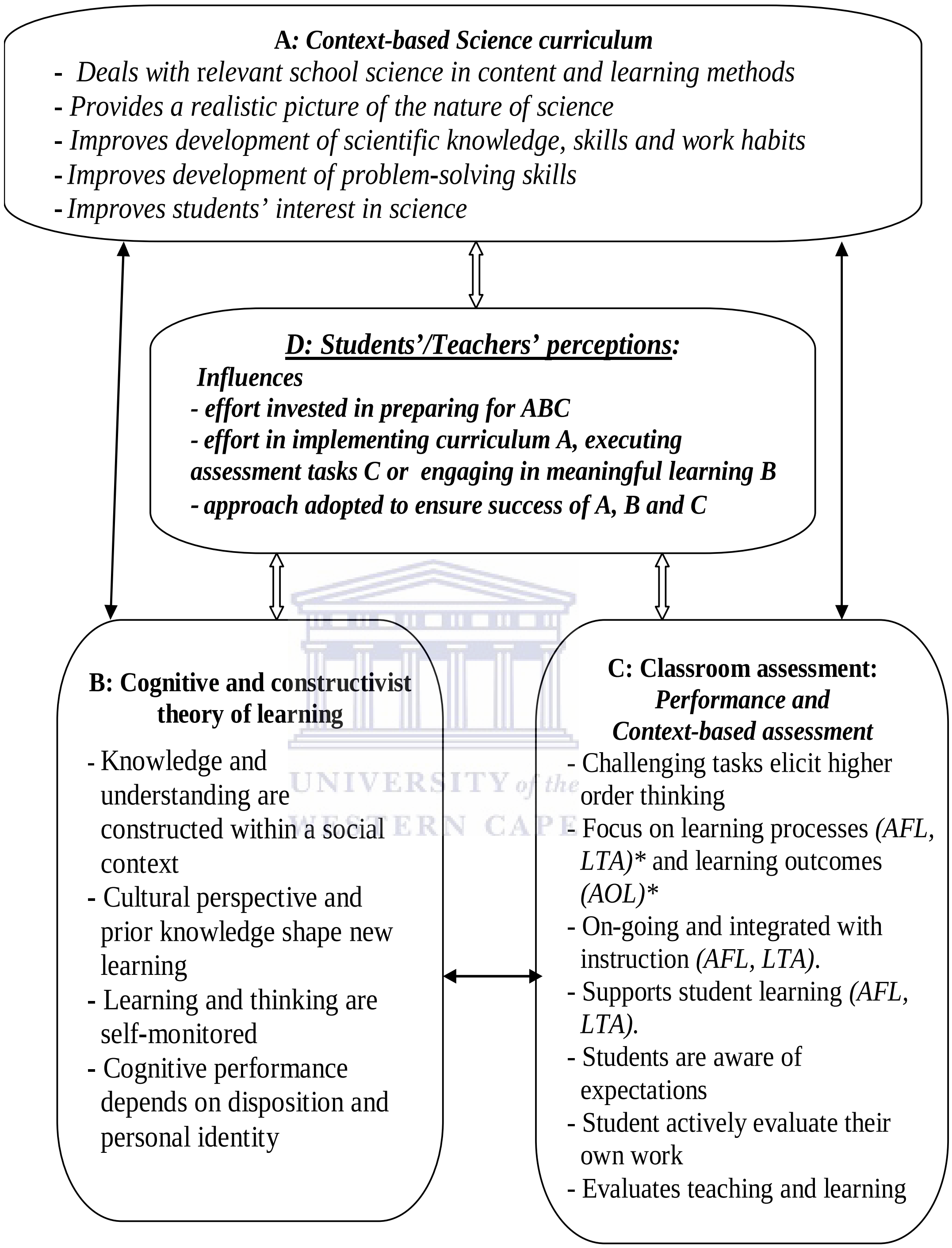


Figure 1.2 Conceptual framework for a new assessment paradigm (* AFL = assessment for learning; AOF=assessment of learning and LTA = learning through assessment)

The above framework guided the design of this study and its components are elaborated on below. It was adapted from Shepard (2000:1074) to include relationships between a constructivist curriculum, learning and classroom assessment to perceptions.

1.9.1 Learning through a context-based curriculum

Demonstrating further the relationship between assessment, teaching and learning is the view that methods used for assessment are determined by educators' views of learning (Sharikzadeh 2003). At the same time assessment practices drive and shape teaching and learning in powerful ways (Ramsden 1997 cited in Maclellan 2001). New views on assessment are based on constructivist theories of learning that emphasises learning as a continuous and active constructive and collaborative process. Students construct meanings rather than simply memorise somebody else's meanings to regurgitate them as answers during assessment.

Context-based teaching facilitates connections between the different components of scientific knowledge by students and fosters a new understanding of this knowledge within a social context. It also complements learning through the use of social issues or social contexts that are familiar to students or apply in real world situations and enables students to construct relevant and meaningful knowledge (Shepard 2000; Sharikzadeh 2003). Through context-based teaching students are encouraged to link school science to contextual variables like culture and life experiences through the use of events that are relevant to students, and in which scientific concepts are embedded (George 1999; Putsoa 1999; Lubben *et al.* 1995). Social contexts are excellent starting points for concept development and make it easy for students to transcend the boundaries between science as experienced in school and science as experienced in the real world context (Shepard 2000).

1.9.2 Assessment for learning

Assessment for learning (AfL) emerged from a paradigm shift that is taking place in assessment practices (Gipps 1994; Sharikzadeh 2003). Others like Neesom (2000) and Black and Wiliam (1998a) equate it to formative assessment. In assessment for learning there is no real distinction between assessment and instruction. Assessment is a regular part of the teaching and learning processes and information from the assessment is used to shape these processes. AfL places less emphasis on scores of the number of facts and procedures that students can reproduce, but places more emphasis on descriptions of students' abilities (Shepard 2000, Gipps 1994). It takes place all the time enabling students to self-assess their progress continuously and the teacher to adjust their teaching in response to students' needs (Neesom 2000). According to Gipps (1994) AfL encourages students to engage in interactive assessment tasks and have access to resources for intellectual work. The feasibility of such a view may be affected by the level of training teachers have in conducting interactive and aided assessment.

1.9.3 Learning through assessment

Learning through assessment (LTA) complements assessment for learning (AFL). According to Goldstein (1994) assessment is an interactive and dynamic activity in which students engage in learning procedures while being tested. Thus, students may undergo conceptual changes during a test taking experience. Their ability to respond to questions changes, even though this change may depend on the confidence students develop as a result of the successful completion of preceding tasks. This study assumed that the dynamic nature of assessment is more pronounced in more interactive assessment models such as performance assessment and context-based assessment models. These two models of assessment can be used as measurement activities to produce grades and as learning activities that could result in a change of students' cognitive states (Gipps and Stobart 2003; Black and Wiliam 1998a).

Assessment for learning and learning through assessment are facilitated by assessment models that allow the integration of teaching, learning and assessment. Performance assessment and context-based assessment models allow such integration. They can be designed to be challenging using tasks that elicit higher order thinking; involve students actively in evaluating their own work; evaluate teaching, as well as learning; or to be integrated with instruction in an ongoing way (Gipps and Stobart 2003; Veronesi 1997).

1.9.4 Learning through social interaction

Learning is best achieved in a social setting where two or more individuals engage in a discourse about a topic. Social interaction promotes learning and improved cognitive structures by exposing students to different and contradicting ideas. Different ideas motivate students to reflect on and re-examine their ideas and maybe restructure and modify them. The need to communicate ideas to others in a group compels students to articulate their ideas more clearly and realise new links that lead to a better differentiation and organisation of their cognitive structures (Good and Brophy 1995). This view of learning implies that assessment for learning can be achieved by making it part of the learning process, as well as a collaboration between teachers and students. In this case peer collaboration, coaching and mentoring become useful strategies for both instruction and assessment for learning. Peer collaboration and coaching facilitate learning by modelling effective ways of thinking and scaffolding complicated performance. It also allows for constructive mutual feedback and the valuing of critical thought (Dietel *et al.* 1991).

1.9.5 Perceptions of students and teachers

Teachers and students hold ideas, beliefs and opinions about teaching, learning and assessment that influence the ways in which they react to these classroom processes. Teacher knowledge, beliefs and perceptions affect the way they react to and implement curriculum innovations, as well as and their cognition and behaviour in teaching to facilitate or inhibit learning (Aschbacher 1993; Yung 2001). When teachers' theories and beliefs vary from the philosophy of curriculum innovations, teachers either re-structure their beliefs or adjust the implemented curriculum or assessment approach according to their belief system. Formative assessment requires teachers to adjust their roles of teacher and assessor to become mentor and judge simultaneously. New theories of learning require a radical shift in the philosophy of assessment to accommodate its role and relationship to learning. The translation of the new views of learning and assessment into pedagogical form depends on personal beliefs teachers hold about assessment and their willingness to change (Yung 2001). Successful implementation of an assessment approach will depend on teacher perceptions of the approach in addition to other contextual variables such as time and class size (James, Griffin and France 2005). According to these researchers assessment innovations are likely to be received and reacted to differently by different teachers. The success of implementing these innovations depends on teachers' beliefs about the function of assessment and the assessment model itself. The understanding of teachers' reactions to new assessment models is crucial in an endeavour to improve assessment and instruction in schools (Aschbacher 1993).

Assessment practices impact on students' lives, motivation and other affective aspects such as self-esteem, confidence, as well as how they perceive their school subjects (Schäffner, Burry-Stock, Cho, Boney and Hamilton 2000). Students also have strong views about different assessment formats. Perceptions of students about assessment influence how students approach their learning as much as students' approaches to learning affect the ways in which students perceive assessment (Struyven, Dochy and Janssens 2003). Students' perceptions of a given assessment activity, their experiences and perceived self-efficacy or perceived importance and value of assessment tasks, determine the effort students invest in the task or their learning for the task (Maclellan 2001; Brookhart and Bronowicz 2003.

1.10 LIMITATIONS OF THE STUDY

The following limitations have been identified from this study:

1. Findings from the study are meant to reflect the type and nature of perceptions that students and teachers participating in this study held about performance and context-based assessment models. The use of these findings for classroom situations beyond those involved in this study should only be as indicators of possible perceptions that may be held by students and teachers in those classrooms. This limitation should not, however, undermine the contribution of the study to information on the perceptions of students and teachers about alternative assessment strategies and their possible use in the science classroom.

2. Context-based paper-and-pencil test items require students to have a particular reading ability and an ability to analyse and understand the scenarios before they can respond. The problem of English as a second language may contribute to problems of interpretation and understanding of the tasks, as well as poorly

expressed responses from the students. Language problems of interpreting questions were minimised by the use of pictures to represent contexts to avoid lengthy descriptions and the inclusion of local labels for some concepts. Some short answer questions were also used. Language related problems and poor expressions of answers may have affected the perceptions investigated.

3. The use of context-based teaching and alternative assessment tasks were a new experience for the participants. Since encounters with practical performance assessment may be limited, a long period of fieldwork was used as a means of promoting familiarity. There is a possibility that more time was needed for teachers and students to form elaborate perceptions. It is noted as a practical point that change is not an event but a process that often needs a long time to take effect (Hord 1987). Hord further warns that possible and permissible change may be effected only to a limited extent because of demands placed on individuals dealing with change and the pressure to produce results from innovations. There is sometimes a danger of innovations being rejected due to premature conclusions.

4. This study would have benefited from a comparison between the perceptions and achievement/performance concerning the tasks or the nature and quality of students' answers. However, this was beyond the scope of the investigation.

1.11 OUTLINE OF THE THESIS

This thesis is organised into eight chapters. The first three chapters are the introduction, the review of related and relevant literature and the research design and methodology. The fourth and sixth chapters are dedicated to the data and the presentation of the results. The fifth and seventh chapters focus on the discussion of the results. The eighth chapter presents the conclusions and recommendations from the investigation. Below are the synopses for each chapter.

Chapter 1: Introduction

This chapter gives the rationale for this study and describes the role and importance of assessment in education and the need for alternative models of assessment that can provide more information about students' competencies than do traditional standard assessment models. The chapter also presents the education context in Swaziland by focusing on the organisation of school level education, science curriculum and reforms, as well as assessment procedures used in schools in the country. The aim of the study, namely to explore and describe perceptions and experiences of science students and teachers of performance assessment and context-based assessment models is presented, as well as the research questions. The conceptual framework is also described.

Chapter 2: Review of related literature

Chapter 2 discusses the literature that was found to be relevant to this study. It reports on the functions of assessment and its role in school education. Problems of traditional assessment models are highlighted to show and justify the call for alternative models of assessment in a re-conceptualisation of assessment of students' learning. Performance assessment, its scoring and implementation, as well as context-based assessment practices and group assessment are discussed and perceptions about performance assessment and context-based assessment are alluded to. The chapter ends with a discussion of the methodological implications of the literature review for this study.

Chapter 3: Research design and methodology

This chapter focuses on the research design and methodology for this study, which embraces the case study approach. It describes the sampling procedures used to select the participants, the different instruments that were developed and used for different purposes in the study, the administration of the data collection process and the data analysis. The chapter ends with noting the methodological challenges experienced in the course of the study.

Chapter 4: Results I: Perceptions of performance assessment

Chapter 4 documents the results obtained on the perceptions of performance assessment tasks by the student and the teacher participants.

Chapter 5: Discussion of results I

This chapter comprises a presentation and discussion of the results reported on in Chapter 4.

Chapter 6: Results II: Perceptions of context-based assessment

Chapter 6 presents the data and results on the perceptions of the students and teachers regarding context-based unit tests and questions.

Chapter 7: Discussion of results II

Results from Chapter 6 on perceptions of context-based tests and questions are discussed in this chapter.

Chapter 8: Conclusions and recommendations

This final chapter provides a summary of the main findings and answers to the research questions. It also gives a brief discussion of the findings, the conclusions drawn from the results, the recommendations based on the findings and suggestions for further research.

Appendices

The appendices comprise the syllabus objectives and the relevant content from the topics used in this study, as well as the learning outcomes, which were based on the content of the Electricity and Air and Living Things units and the syllabus objectives. Presented in the appendices are also letters of request to the schools and the Ministry of Education, and a letter granting the permission to conduct the study and the data collection instruments (interview schedules and questionnaires).

Data sources in the form of interview transcripts and student questionnaire responses are, however, not provided in the appendices. These documents are omitted from the appendices because information from the mentioned sources are extensively quoted and referred to, to support and authenticate assertions made in the results sections reported in Chapter Four and Chapter Six.

The next chapter reflects on recent and relevant literature concerning the main theme of the investigation, namely the perceptions and experiences of students and teachers about performance assessment and context-based assessment models and their role in the learning and assessment of junior level Science.

Figure 1.3 Below presents a time-line for the main activities of the study. The arrows represent some of the sequence steps involved in the development of the instruments, their validation process and implementation. It was not possible to show links for all the related activities without congesting the diagram.

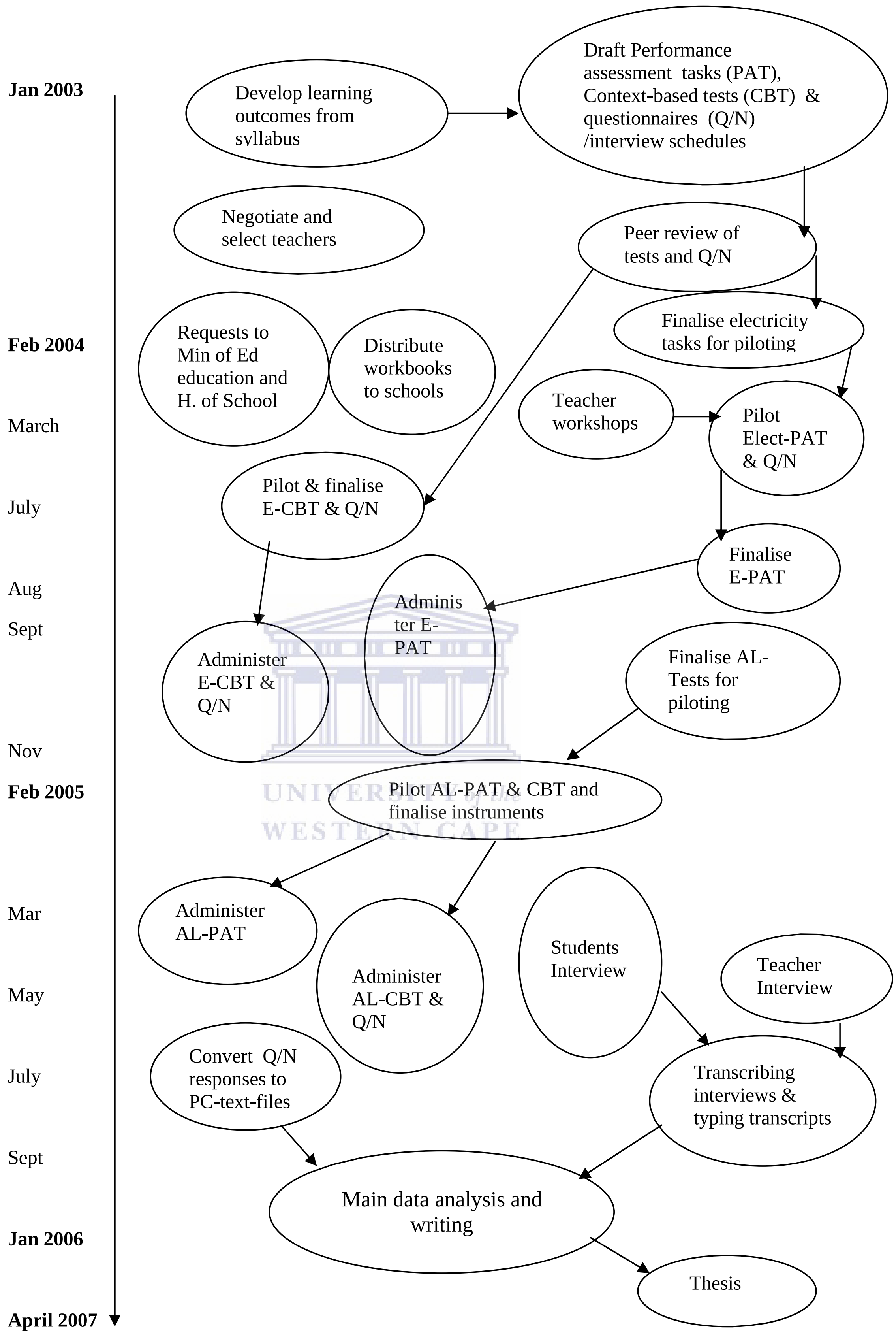


Figure 1.3 Thesis activities and time-line

2. CHAPTER 2

REVIEW OF RELATED LITERATURE

2.1 INTRODUCTION

This chapter presents a reflective overview of studies, research reports, reviews of literature and commentary notes on research work that have relevance to this thesis. The review is organised into four parts:

- Background information about assessment; its role in education and problems that are motivational for a paradigm shift in assessment.
- Alternative assessment models focusing on performance assessment (hands-on practical activities) and context-based (real world-based or authentic) assessment.
- Students' and teachers' perceptions about assessment issues.
- Group assessment.

The chapter ends with a summary on methodological implications of the reviewed literature for this study

2.2 ASSESSMENT: ITS ROLE IN EDUCATION

For many it is impossible to visualise formal education that is not punctuated by assessment events designed to check students learning (Broadfoot and Black 2004). Assessment is an endeavour by classroom teachers to ascertain the status of students' knowledge (cognitive understandings and abilities), as well as skills and attitudes (psychomotor and affective factors) as variables of educational interest (Popham 1999). Every teacher engages in some form of assessment of students' learning through formal or informal means and by using different assessment approaches. Broadfoot (1996) notes that through assessment strengths and weaknesses of learning by individuals, as well as the functioning of institutions and education systems can be ascertained. Assessment is thus an essential element in teaching and learning processes (Gipps and Stobart 2003; Shepard 2000).

Assessment achieves recognition of its importance from the information it provides and its uses. Information from assessment can be used for summative and formative purposes in which judgments and decisions are made regarding students' learning, teaching methods and education policies (Biehler and Snowman 1990). Figure 1.1 above (Section 1.9) shows the relationship between summative assessment and formative assessment and how the latter relates to a new model for using classroom assessment as assessment for learning.

2.2.1 Summative assessment

Biehler and Snowman (1990) point out that summative use of assessment information involves production of a summary of the students' overall performance. This summary forms the basis for judging how well students achieved curriculum goals. The summary is also useful for comparing each student's achievement with that of others students (norm-referenced assessment) and in preparing students' reports for parents, administrators and inspectors or other interested agencies. Furthermore, crucial decisions affecting students' lives are made, such as their progression to the next class, further studies or getting awards. Summative assessment data impacts on learning after the fact. A student may realise what is at stake and decide what to do after seeing the result of their performance.

Summative assessment data does not always have impact on classroom instruction and learning although influences decisions which may have profound educational and personal consequences for students (Sadler 1989).

2.2.2 Formative assessment

Black and Wiliam (1998a; 1998b) state that there is no tight definition or internationally agreed upon and widely held meaning for formative assessment. Different scholars describe formative assessment in different ways. One way noted by the two mentioned authors is that formative assessment can refer to all activities undertaken by teachers and/or their students that provide information to be used as feedback to modify teaching and learning activities in their classrooms and to meet students' needs.

According to Biehler and Snowman (1990) and Elliot, Kratochwill, Cook and Travers (2000), for example, teachers use assessment information to:

- judge students' level of attaining curriculum goals;
- judge the effectiveness of their instruction to help students achieve instructional objectives; and
- decide how to improve instruction and promote the attainment of curriculum goals by their students.

The different ways in which formative assessment information is used as indicated above, link assessment to classroom instruction and learning in ways that are encouraged by new views of assessment.

Another link of formative assessment to classroom instruction is through feedback from teachers to students on learning progress. Feedback received by students regarding their learning comes in the form of marks or grades, comments and endorsements on their work. According to Elliot *et al.* (2000) such feedback enables students to:

- become aware of and correct their mistakes;
- become aware of the expected or targeted learning outcomes;
- become aware of the kind of performance that they need in order to succeed and where they need to spend more time and effort; and
- actively assess learning at the personal level and set goals and academic expectations for themselves.

However, students hold the responsibility to act on feedback from assessment activities to improve their understanding of subject matter and performance (Stepanek 2002).

Formative assessment supports learning through feedback and students' positive reaction to such feedback. Gipps (1994) extends the uses of assessment from support for learning through feedback to using assessment to enhance students' competencies. Having acknowledged that assessment is undergoing a paradigm shift from a psychometric and measurement perspective to an assessment view, she advocates for an alternative way of viewing classroom assessment. In this view students are active participants in the assessment process. Assessment takes the role of helping students to learn and teachers to improve instruction. In this view the functions of assessment extend beyond assessment of learning to assessment for learning and skills development (Gipps 1994). Assessment of learning focuses on what students have learned while the activities used in assessment for learning are designed to facilitate the development and consolidation of students' knowledge and skills (Gipps and Stobart 2003; Shepard 2000; Wiggins 1996/1997). Such a view requires assessment to be an on-going process so that it continuously informs the management of learning by both students and teachers (Taber 2003).

The use of assessment instruments (e.g. tests and examinations papers) for summative and formative assessment may overlap. Wiliam and Black (1996) note that an assessment instrument originally designed for summative functions may be used formatively to help students prepare for further summative tasks. This can be observed when teachers use papers from past examinations for students to practice for future national examinations. In some instances the results from tasks and exercises intended for formative assessment may also be used to judge students achievement in summative ways. These uses present formative and summative assessment as taking up two ends of a continuum along which different formats and uses of assessment tasks and information can take place.

Formative uses of assessment were extensions from formative curriculum evaluation introduced by Michael Scriven in 1967. According to Bloom, Hastings and Madaus (1971) and Wiliam and Black (1996) the term formative evaluation was introduced by Scriven in 1967 in connection with curriculum improvement. Formative evaluation was born from a search for forms of evaluation that would not have the same effect of anxiety and defensiveness on teachers and students as did summative evaluation (Bloom *et al.* 1971). Bloom *et al.* (1971) were the first to extend the use of formative evaluation to its use as practiced today in regard to the testing of students' learning (Wiliam and Black 1996). They defined formative evaluations as

> involving the collection of appropriate evidence during the construction and trying out of the curriculum in such a way that revisions of the curriculum can be based on this evidence (Bloom *et al.* 1971:117).

Below is a brief presentation of the historical background of assessment.

2.2.3 Assessment: some historical perspectives

The formal use of tests for selection purposes was first practiced in the Chinese civil service examination system around 210 BC. According to Madaus (1990 cited in Madaus and Kellaghan 1993) and Min and Xiuwen (2001) the Han dynasty was faced with the need to select candidates who could serve as government officers. Successive dynasties developed a series of formal, written, competitive tests along with standardised procedures for administering and scoring them for the purpose of selecting appropriate candidates (Madaus 1990 cited in Madaus and Kellaghan 1993; Black 1998). According to Black (1998) the tests were intended to ensure that test attainment was the route to success, as well as opening up opportunities for selection on merit rather than parentage or patronage. These tests focused on memory and recall of information and as such their success in selecting the right candidates was questionable. The system was abolished in the early 1900 when educational reform took place in China (Black 1998; Wilbrink 1997; Min and Xiuwen 2001).

The Chinese testing model set in motion developments in assessment in Western Europe that resulted in the dominant presence of examinations observed today (Wilbrink 1997). According to Madaus and Kellaghan (1993) and Stray (2001) the testing ideas used in China filtered into Europe from the 16th Century, where written examinations were introduced to replace oral examinations. Prior to the introduction of the written examinations, oral examinations introduced in the 14th Century were used for promoting students to the next level and the ranking of students on merit (Wilbrink 1997). Formal testing began to be used to define what was expected from students and to make students show periodically how their learning was progressing. Testing also became a way of ensuring knowledge transfer from teachers to students. Policymakers began to realise the potential of examinations as a mechanism for exercising power over students, teachers and schools (Foucault 1977 cited in Madaus and Kellaghan 1993). Madaus and Kellaghan (1993) also cite Madaus and Kellaghan (1992) as suggesting that the linking of results of tests and examinations to rewards or sanctions was responsible for motivating individuals, groups and institutions to perform well, rather than the influence from the tests or performance in those tests. Examiners who managed examinations were able to control the actions of teachers and students through the use of assessment instruments and information.

Further development in the use of tests led to their use to rank candidates from high to low according to their performance in all written examinations. Later on the quantification of performance was introduced. Quantitative marks were assigned to individual test questions. This was the onset of the psychometric use of test results, a shift from "socio-moral assessments of members of status groups towards purely cognitive assessments of individuals" during the 18th Century (Stray 2001:47). According to Madaus and Kellaghan (1993) the use of quantitative marks made it easy for students' performance to be manipulated by aggregating them, organising, averaging and ranking or classifying the marks. In this form marks could be used in a variety of ways such as comparing individuals, groups, schools or position these in statistical distributions. Testing ideas were introduced to the United Sates from Europe in the middle of the 19th Century.

Outcomes of the developments in assessment issues taking place in western countries also filtered into countries colonised by the western countries. Swaziland for example, adopted the British education system and even with the recent curriculum developments she is still largely influenced by that system.

Madaus (1985) notes that the use of test results expanded from simple descriptions of students' performance and comparison, to high stakes uses such as serving as an administrative mechanism to implement policies. Through test results crucial decisions about promotion and retention, placement of students for remedial programmes, certification of successful completion of studies, as well as resource allocation to schools could be made (also see Wilbrink 1997).

The psychometric functions of assessment as observed today date back to the 18th Century although some developments took place along the way. These uses of test results have also evolved over the years from their summative functions noted in the paragraphs above to formative functions.

Bloom *et al.* (1971) note that Scriven introduced formative evaluation as a reaction to resistance by curriculum developers to change finalised curricula even when there was evidence of the need for such changes. Scriven is said to have viewed formative evaluation as involving collecting appropriate evidence during the construction and piloting of a new curriculum on which to base revisions of the curriculum, rather than wait till the curriculum was finalised. Bloom *et al*. (1971) came across formative evaluation while searching for a term for an assessment form that caused less anxiety and defensiveness in students, teachers and curriculum developers than did summative evaluation. Summative evaluation was, and still is, used in reference to tests that were administered at the end of course units or programme of study. Tests were designed to assist in judging the degree to which outcomes of the course were achieved by students, as well as to grade students for purposes of reporting to parents and others and to certify students for progression to the next level of their studies. Formative evaluation helped students and teachers improve what they wished to do. Wiliam and Black (1996) note that from the earliest use of summative and formative evaluation, it was in relation to the functions served by the assessment activity rather than the tests themselves. Thus from its introduction formative evaluation was intended to improve teaching, learning and curriculum development.

It is evident from this very brief historical perspective of assessment that the summative uses of assessment results have been around for a while. It took China more than a thousand years to change an assessment system designed for selection purposes in a reform process of the education system (Black 1998). Formative use of assessment information is a more recent development. Some of the original principles of summative and formative evaluation still apply to assessment practices today. There are calls for the expansion of the uses of assessment information to include more direct support for learning. Below is a further illustration of the need to seek other models of assessment to complement the use of these traditional assessment models to improve learning and instruction.

2.3 PROBLEMS OF TRADITIONAL TESTING MODELS

For a long time oral examinations were used for the evaluation of learning until they their shortfalls were identified and the written examinations were a better alternative. One shortfall was

> the oral exam's inefficiency in the face of growing numbers and lack of standardization, which precluded the comparison and ranking of students on the same set of questions taken under the same conditions. (Madaus and Kellaghan 1993).

Written examinations and standardisation procedures evolved and improved over the years to what has become known as 'traditional assessment formats" in the form of objective type questions such as multiple choice questions, structured and short answer questions. Uses of assessment information also expanded from comparing, ranking and classifying students to informing important policy decisions regarding education systems.

Again, for a long time educational assessment has utilised traditional testing models; standardised and non-standardised tests that worked well with the summative use of assessment information. Standardised tests take the form of multiple choice and/or closed questions that are administered, scored and interpreted in a standard manner (Popham 1999). Their preparation often follows a rigorous process of validation by content and assessment experts; reliability testing; and other standardisation procedures and criticism before they are used in assessing students (Sanders and Horn 1995; Popham 1999). Standardised tests are thus more appropriate for use in high stakes assessment where their outcomes are useful when comparing students, generalising about their achievement, indicating levels of attainment based on set standards and informing education policy decisions. Commercial testing agencies have acceded to the challenge of producing standardised tests intended for widespread use as they are too costly to be developed for use by a single teacher in a school (Popham 1999).

Non-standardised assessment on the other hand is generally used by teachers in their classroom for formal or informal purposes. They take the form of teacher constructed tests used in the evaluation of student responses (Sanders and Horn 1995), although these may comprise items taken from standardized assessment instruments.

Methods of assessment that are based on traditional testing models (structured or short answer and multiple choice based assessment) whether standardised or not have been under intense criticism for their ineffectiveness in measuring higher order thinking skills, and intellectual and manipulative processes, as well as their emphasis on rote learning and recall (Sanders and Horn 1995). They have also been criticised for the effects they have had on instruction and the curriculum (Gipps and Stobart 2003). Traditional testing models convey the idea that there is one correct answer to every question. They also encourage students to place greater value on learning by rote memorisation and recall of memorised information in tests (Moorcroft *et al.* 2000). Real life experiences of raising questions about observations, constructing responses, identifying problems and finding solutions to those problems are generally omitted from these assessment tasks. Teachers have been observed to focus their teaching on what is usually asked in tests and examinations and preparing students for high achievement in such tests (Gipps and Stobart 2003; Maclellan 2001; Moorcroft *et al.* 2000). There are, however, a few reasons suggested by Madaus and Kellaghan (1993) that encourage the practice of teaching to the test. These are that teachers will teach to the test because:

- important decisions that affect students and schools are based on results of tests or examinations;
- a tradition of past examination questions develops a tradition that defines what is to be taught, particularly in settings where high stakes examinations operate; and
- society tends to treat examination results as the goal of schooling because they are the basis for future choices in life.

Teaching to the test has had some positive outcomes for teachers and students in terms of high student achievements in external examinations. However, teachers are also under pressure to deal with as many topics as possible each year and to prepare students for achievement in these examinations. In the process they lose opportunities of exploring the subject fully with their students (Moorcroft *et al.* 2000). As a result of these unfavourable effects of assessment on educational practice many educators in formal settings are exploring the effectiveness of using alternative assessment models to appraise the attainment of learning outcomes by students (Popham 1999).

2.4 ALTERNATIVE ASSESSMENT MODELS

Assessment is undergoing a paradigm shift from a measurement model to a standards model and from a testing culture to an assessment culture (Gipps 1994; Dochy and McDowell 1997). According to Maclellan (2001) the measurement model is based on the traditional learning theory where learning is a process of knowledge reproduction, context independent and transferable to other contexts. This corresponds to the testing culture where teachers are seen as carriers of knowledge to be transferred to students' heads (Dochy and McDowell 1997). The assessment culture and standards model of assessment view learning as a process of knowledge construction that is situated in a particular context, and the knowledge constructed is not necessarily transferable to other contexts. The teacher is not a transferor of knowledge but a mentor who provides opportunities for students to use the knowledge and skills they already possess in order to understand new topics. The teacher is therefore expected to provide interesting and challenging tasks for students (Dochy and McDowell 1997).

Assessment in the standards model demands representation of meaningful, significant and worthwhile endeavour to accomplish the assessment task. In this way assessment reflects the ways in which knowledge and skills are used in real world contexts (Newmann and Archbald 1992 cited in Maclellan 2001). The standards model is concerned with the level at which knowledge is embedded in deep (and new) understanding and can be demonstrated in authentic tasks. It also emphasises the value of education as a means of promoting the development of individual students rather than emphasise individual differences. Alternative assessment models seem to complement the new views of assessment. The assessment tasks are believed to be more likely to give a more comprehensive picture of students' capabilities. According to Sambell *et al.* (1997) alternative assessment:

- favours the integration of assessment, teaching and learning;
- involves students as active and informed participants;
- utilises authentic, meaningful and engaging assessment tasks that mirror realistic contexts in contrast with the artificial time constraints and limited access to support found in conventional exams;
- focus on both the process and products of learning; and

- moves away from single test scores towards a descriptive assessment based on a range of abilities and outcomes.

Alternative assessment models encourage critical thinking and decision making among students and take into consideration students' thinking processes and products during scoring (Herman, Klein and Wakai 1997).

According to Downing (not dated) and Linn and Gronlund (2000) alternative assessment models take different formats described by terms like authentic assessment, portfolio assessment and performance assessment. The distinction between authentic assessment and the other two formats is only slight; with the label "authentic" also applying to both performance and portfolio assessment. One distinction, however, is that performance-based assessment determines how a student performs on a given task, while authentic assessment utilises activities that are in context of real life situations, that is, it emphasises the practical application of the tasks in real world settings. Some educators prefer to use the different terms interchangeably to avoid confusion.

Alternative assessment techniques extend beyond paper-and-pencil measurements, assessing students on a number of abilities based on what they can do in terms of conducting actual tasks and products or outcome of the task (Silbermann (not dated); Gipps and Stobart 2003). The tasks used for assessment of learning can also be used as exercises through which students can further explore their understanding and application of knowledge in a topic of study. Since they focus on both assessment of achievement and developing of understanding, they help students learn (Gipps and Stobart 2003; Moorcroft *et al.* 2000). The tasks may range in complexity from simple extended multiple choice questions – restricted response format, to an in-depth scientific investigation – extended response format (Linn and Gronlund 2000; Moorcroft *et al.* 2000) as is the case in performance assessment.

2.4.1 Performance assessment as alternative assessment

Assessing knowledge of content is easier than assessing skills and values when using traditional assessment models. Alternative assessment models such as performance assessment have been developed to provide a complementary way of assessing a broad range of student abilities.

Gipps quotes Stiggins and Bridgeford's (1992) succinct description of performance assessment as

> … [a] systematic way to measure learners' ability to use previous acquired information/knowledge in solving problems or completing a specific task; real life situations or simulated assessment exercises are used to elicit original responses which are directly observed and rated by a qualified judge (1994:99).

Different educators use the term performance assessment to refer to different kinds of assessment approaches. Some equate performance assessment to any form of constructed response assessment. Others contend that genuine performance assessment must possess at least three features, namely (i) multiple evaluative criteria - more than one criterion is used to judge a student's performance; (ii) pre-specified quality standards to be used in judging the performance and (iii) judgment appraisal where human judgment determines acceptance of demonstrated performance (Popham 1999). Performance assessment relies on teacher observation and professional judgment to draw inferences about student achievement. Performance assessment is associated with higher order skills of planning, measurement and observation, and thus requires students carrying out tasks to be familiar with the skills and techniques necessary to investigate ideas around scientific concepts (Gott and Duggan 2002).

Performance assessment requires students individually or in small groups to demonstrate their skills by actually performing a given task and using their knowledge and skill in context, not merely by completing a task when cued. They therefore can provide teachers with information about how a student understands and uses knowledge. The performance expected from students can be in the form of manipulations of equipment and the generation, recording, as well as computation of data and drawing of conclusions from the data. Performance assessment tasks measure students' ability to perform tasks that correspond to important instructional objectives including those in the psychomotor and affective domain (Linn and Gronlund 2000). They are primarily used to measure learning outcomes that cannot be easily measured by objective test items or other paper-and-pencil tests. They are thus recommended for use with less structured problems and tasks such as problem- identification, collection, organisation, integration and evaluation of information and for learning outcomes that involve the creation of a product. Hands-on activities provide ample opportunities for the assessment of an array of skills (Aschbacher 1991; Shavelson, Baxter and Pine 1991).

From the above discussion of performance assessment this thesis takes a view of performance assessment as an approach that integrates science investigations such as hands-on practical tasks to measure and evaluate students' content and procedural knowledge and their ability to use the knowledge in reasoning and solving problems in social contexts. Students execute context-embedded tasks that are meaningful and engaging to demonstrate their knowledge, skill and work habits through:

- manipulating and operating scientific instruments and equipment to generate relevant data;
- recording, analysing and interpreting data;
- drawing relevant conclusions from data and using the data in a real world context; and
- communicating the product of their investigation orally and in written reports.

2.4.2 Scoring in performance assessment

In performance assessment grading usually makes use of teacher observation of students while they perform the assessment task and the professional judgment of such performance to draw inferences about student achievement. The teacher is able to see beyond the written response into students' thought processes, whether students learned for the test or they really understand the concepts and are able to apply them (Roeber 1996).

Performance assessment tasks do not have clear-cut right or wrong answers. Teachers need to construct special scoring guides or rubrics that specify criteria by which students' work is judged. The criteria define scoring in terms of degrees of successful implementation of the assessment task and meeting the criteria (Aschbacher 1993; Brualdi 1998). Performance assessment is ideal where teachers (trained scorers) directly observe students while manipulating or operating laboratory equipment and material to perform the task (Shavelson and Baxter 1992).

Borich and Tombari (1997), Brualdi (1998) and Downing (not dated) provide the following formats for presenting rubrics:

- Rating scales (where the criteria of proficiency levels are specified in a continuum from an excellent to an unacceptable product and may be represented by numerical figures).
- Checklists (performances are presented as specific actions that are scored as either present or absent and are appropriate for performances that can be presented as specific actions).
- Holistic scoring (overall quality of performance is estimated and scored as a numerical value).
- Narratives/anecdotes (observer writes narrative reports of students' performances and determines the degree of acceptance of performance against set standards).

The choice of rubrics depends on the capability to be measured. They do, however, have to be constructed carefully and the raters trained if the rubrics are to be used successfully and effectively in scoring students' performances (Borich and Tombari 1997). This study used rating scales as they were similar to a familiar way of scoring of students' work for the teachers.

2.4.3 Assessment and learning

A good assessment is one that is built on current theories of learning and cognition and grounded in views, skills and capacities that students will need for their future experiences (Herman 1992). This statement calls for links between assessment formats and the curriculum where learning outcomes are outlined, particularly when new programmes with different pedagogy and learning objectives are developed.

Whilst no single form of assessment model can appraise in totality a student's school learning experience, it is possible to develop measures that can indicate learning that is closely related to the curriculum (Sanders and Horn. 1995). However, the call for including alternative assessment as part of the assessment of school learning is not simply the use of alternative forms of assessment; it also involves the alternative use of assessment as part of the learning process (Gipps and Stobart 2003). Such a use would make assessment part of the learning process, an integral part perhaps, and more relevant to the curriculum. The call for alternative assessment is motivated by constructivist reforms in science curricula and cognitive research (Shavelson *et al.* 1991).

Aschbacher (1993) explored facilitators and barriers of implementing performance assessment (portfolio, journals, open ended questions and essays). Factors like commitment to the assessment model and instruction, support of the administrators, technical assistance and working with colleagues facilitated execution of these alternative assessment approaches. Factors inhibiting the implementation process revolved around:

- assessment anxiety;
- lack of time to learn, plan, practice, use and reflect on the performance assessment approach;
- need for training and ongoing support; and
- resistance to change by teachers and administrators.

Teachers in Aschbacher's (1993) study perceived a change in their attitude about their teaching, assessment practices and use of alternative assessment. Teachers were more positive about the value of working with alternative assessment models. They changed their view of the role of students in their learning. Teachers began to expect and encourage more student talk, change their focus from correct answer in content of students' expressions to the nature of students' explanations. Teachers also felt that through the alternative approaches they could diagnose skills students had never expressed before, and got better insight into students' thought processes. Similar observations were made by Nuttall (1992).

Aschbacher (1993) also reports that teachers perceived some student and instructional benefits from the assessment models used. There was a perceived increase in students understanding which was demonstrated by different ways in which students could express their knowledge. Teachers reflected more about their teaching practices in order to align instruction and assessment. Aschbacher (1993), however, observed that the teachers exhibited some reluctance to developing learning outcomes, and the standards and criteria for assessing those outcomes. The use of performance assessment needed greater input from teachers. Thus, teachers needed to engage with students in ways they had not used before, such as monitoring in small groups, conferencing with students and coaching performances, as well as assume greater evaluation responsibility than in traditional assessment models. This could constitute an increase in the workloads of both students and teachers. For teachers to implement new curricula or assessment reforms to the benefit of learning and teaching, teachers may need the support of keen and committed professionals in teacher education and researchers (Aschbacher1993).

Integrating new assessment forms into classroom instruction demand a dynamic conceptual shift for teachers and administrators and a strong background and understanding of current theories of teaching, learning, assessment and curriculum development.

The call for alternative assessment models is also a response to the implicit or absent link between assessment, the curriculum, learning theories and cognitive research. Curriculum reforms that subscribe to constructivist ideas show promise for promoting the links between assessment, learning and the curriculum. Shavelson *et al.* say that current curriculum reforms that are grounded on constructivist learning theories view students as

> … active agents in the teaching-learning process, constructing personal and shared meanings in the subject matter. The subject matter is well contextualised in a culture of learning and problem-solving, one that encourages group work, as well as individual work. Hands-on activities and long-term projects are the rule rather than the exception (1991:348).

Current theories of learning and cognition that emphasise skills and capacities for future use are the basis of good assessment (Herman 1992). Performance assessment models apply approaches that are suggested by modern learning theories, such as constructivism. Students are viewed as active participants in the construction of meaningful knowledge, not “as recipients of discrete bits of knowledge” (Linn and Gronlund 2000:267). When students are engaged in performance assessment tasks they use their background knowledge in interpreting the tasks and in constructing meaning of the tasks and their demands.

Performance assessment tasks are more complex than multiple choice tests as they measure multiple reasoning and knowledge (declarative and schematic dimensions of knowledge). Good performance assessment tasks are essential if they are to positively influence teaching. However, changing the assessment format is no guarantee of a change in teaching styles. So, using performance assessment with teachers who teach to the test will not necessarily improve their teaching (Shavelson and Baxter 1992; Aschbacher 1993).

2.4.4 Implementing performance assessment in Science

The implementation of new curricula and performance assessment models is believed to require that teachers and students take up new roles and change their mindset from a focus on content coverage to outcomes based on standards achieved (Aschbacher 1991; Aschbacher 1993).

Studies on the use of performance assessment models have shown that specific scoring criteria and examples showing expected competencies are essential for consistent evaluation through such models. Indicating to students the expected performances regarding the tasks motivates them to improve their performance (Slater (not dated)). That is, the main assessment criteria may be shared with the students prior to their attempting the task so that they know in advance what is expected from them (Gipps and Stobart 2003). With such knowledge students can monitor their performance as they work on the task, thus meeting a requirement of performance assessment on students noted by Shavelson *et al.* (1991). Slater (not dated) further notes that students' performance on the tasks improves if students interact among themselves and with the teacher while engaged in the task. This characteristic of performance assessment models shows that important human support can be availed to the students during the assessment process, while such support is not permitted in traditional assessment models. Peer and teacher support make performance assessment tasks good teaching activities, which complies with Shavelson *et al.'s* (1991) view of the interchangeable use of good instructional activities and good assessment tasks.

Experiences from the Assessment of Performance Unit in the United Kingdom noted by Nuttall (1992) show some of the strengths of performance assessment activities.

Nuttall notes that the students enjoyed the assessment tasks and experienced reduced test anxiety when the assessment tasks were similar to those used as classroom activities. The performance assessment tasks also revealed student attainments that were commonly not recognised or awarded credit before. Furthermore, students received higher scores in their performance assessment. Nuttall advises that with all these positive aspects of performance assessment comes the demand for a significant amount of teacher time and energy to prepare, administer and grade the tasks. Furthermore, teachers may find it difficult to administer the tasks and monitor the entire class at the same time. Making use of assistance from other teachers is a possibility with a potential of disrupting other school activities unless carefully planned (Nuttall 1992).

2.5 CONTEXT-BASED LEARNING AND ASSESSMENT

As part of the paradigm shift in assessment noted earlier in this chapter, another development taking place is the inclusion of students' everyday life and potential experiences or other relevant world contexts in their teaching and assessment. Standardized multiple choice and open-ended questions in examinations are being replaced by structured questions that are embedded in contexts (Taber 2003; Ahmed and Pollitt 2001). Much research effort has therefore been put into the use of context-based activities of teaching, learning and assessment models in response to calls for teaching and assessment in Science in ways that are perceived to be more relevant.

As noted in Chapter One there is evidence that the use of contextualised tasks as learning activities allow students to integrate their knowledge, skills and attitudes when executing tasks. They also learn to coordinate skills required for performing complex tasks and maybe even help them to transfer what they have learned in science classrooms to their daily life activities. However, these tasks are associated with a high cognitive load that may constitute learning difficulties for some students (van Merriënboer, Kirschner and Kester 2003).

Boaler (1994) notes from literature three reasons for using context-based learning, namely to:

- provide students with a concrete and familiar experience to make learning more accessible;
- motivate students and increase their interest by providing examples which enliven lessons; and
- promote the transfer of mathematical learning by showing examples of links between this school subject and real world problems and phenomena.

Context-based approaches to teaching Science have been around for more than twenty years (Bennett 2003). Recent instructional theories focus on the use of authentic, real life tasks as the driving force for learning (van Merriënboer *et al.* 2003). A number of curriculum developments now emphasise teaching Science through social contexts (Koosimile 2004). As such, a wide range of materials have been developed in response to a desire to address the observed limited relevance of science material used for teaching and the dropping of student interest in Science evident in the low numbers of students pursuing Science or the students' perceptions that Science is a difficult subject (Bennett 2003). Teachers are faced with a challenge: to make science lessons interesting and meaningful for their students.

Bennett (2003) notes another concern of having to provide courses for students who did not specialise in Science at secondary school level and for whom the conventional academic science courses were fairly inaccessible. Different approaches have been followed in constructing context-based teaching materials. Bennett describes four such approaches with varying amounts of context-based teaching. There are those comprising complete courses where contexts form the framework in which scientific knowledge and understanding form a coherent system of scientific concepts. Another approach comprises courses where scientific contexts and applications are the focus of instruction. The context drives instruction rather than the coverage of science content. These courses were developed for tertiary level non-Science majors in the United Kingdom. The third group of teaching materials comprises particular contexts and applications strategically used to replace some conventional topics in Science. The fourth group utilises even less context-based teaching by integrating only a few units within a conventional topic.

The outlined different formats and approaches to context-based teaching imply the need for caution in the design and use of context-based assessment items or tasks so as to match these to the curriculum formats in use. Some of the concerns raised by researchers regarding contextualised examination questions (discussed below) could be explained through these different formats of contextualising science teaching. This is with regard to the focus of the assessment task, whether the focus is on assessing students' scientific knowledge or their abilities to recognise and identify scientific principles used in their environment or apply scientific knowledge to their environment.

In Swaziland context-based curriculum materials that follow a technological approach to science teaching and learning have been used. These materials were developed to serve as alternative teaching material for Electricity, and Air and Living Things in the junior secondary school science curriculum and were known as the Matsapha Lessons (Lubben *et al.* 1995). Another set of curriculum materials were developed for two units: Forces, Support and Movement and Acids, Alkalis and Gases through the Linking School Science with Industry and Technology (LISSIT) project.

Studies involving the Matsapha Lessons on students' competency in identifying relevant scientific ideas correctly showed that only a slight improvement was produced (Lubben, Campbell and Dlamini 1997). The competence was also independent of concept achievement. These observations were obtained from data generated from two end-of-unit tests. One test consisted of the standard short answer examination type questions focusing on recall and understanding. The second test was also a paper-and-pencil test but practically oriented and contextualised. This test explored students' application of scientific knowledge and investigation abilities. Results showed that overall the control group that had not been taught through the Matsapha Lessons, performed equally well on the standard concept understanding test in circuit electricity. High achieving students in the study became better equipped to describe reasonably valid methods of investigating, even though they had not been taught how to do this. However, some confounding variables were noted regarding the unanticipated lack of improved performance of the experimental group. For example, the sequencing of teaching units, the introduction of a new teaching approach and material, teachers' and students' degree of familiarity with the teaching approach were suspect in the apparent lack of improvement in the experimental group. Proper induction of teachers was identified by Lubben *et al*. (1997) as an important factor in the process of acquainting them with curriculum innovations.

In another report involving the Matsapha Lessons, Dlamini *et al.* (1996) found that students' motivation to learn in Science was enhanced by real life contexts perceived to have an immediate or future personal link to students' experiences. Students' interest also increased with their perceived expertise in the context. Context-based teaching has a potential of shifting pedagogy from the teacher to the students to increase their active participation and promote concept development. Through context-based teaching students speculate about solutions and explanations for everyday problems and occurrences. Student discussions provide opportunities for teachers to assess students' assertions and diagnose their conceptions of Science as taught through the use of social contexts. Improved concept development can be expected from the students' discussions, demonstrations or experimentation activities that specifically target scientific principles in the context or those that counter alternative conceptions diagnosed from students' expressions. However, successful concept development in contextualised teaching requires teachers to accept and utilise student centred teaching styles, alertness and consciousness to opportunities for concept development (Dlamini *et al.* 1996).

The success of context-based teaching has been observed not only in students' increased enthusiasm and motivation but also in the delivery of content in Chemistry (Belt, Leisvik, Hyde and Overton 2005). Belt *et al.* (*ibid*) conducted a case study in the United Kingdom in which students were introduced to thermodynamics and kinetics. Students were required to identify fuel sources for a hypothetical newly established city through interpreting and evaluating a variety of Physical Chemistry data. Students were also expected to develop and use a variety of skills relating to group work, such as communication, organisation, problem-solving and critical thinking. All this work culminated in oral presentations and reports by the students. Students were, in addition, asked to indicate their perceived level of difficulty with regard to the tasks. Participating students appreciated studying Chemistry within an applied context and felt that the approach could lead to the development of their subject knowledge and their perception of its relevance. Some students expressed a perceived increased confidence in approaching problem-solving in the future. A number of students found the calculations in tasks challenging until they realised a familiar method to use in the calculations. Further difficulties reported by students were: working out how to approach the task, knowing where to start with the task, understanding what was asked, as well as deciphering each task to know which bits of information to use (but were fine once that was done).

Context-based teaching may cause some difficulties for teachers who do not have similar experiences as those of their students or who are not familiar with the environment of their students. As a result of the differences in experiences of teachers and students some teachers may not recognise students' contributions as relevant and therefore reject those contributions (Koosimile 2004). Koosimile observed that teachers readily supplied or accepted scientific versions and explanations of reality, but not those supplied by students. Similar observations were reported by Lubben *et al.* (1997) where teachers accepted only correct responses and rejected or ignored other responses without exploring their limitations. In these studies the conflict between teachers' and students' ways of explaining or interpreting phenomena tended to remain unresolved, resulting in experiences that were not in support of contextualised learning. In such instances students may develop negative perceptions such as that their ways of describing reality are of lower value or inaccurate. Students who are convinced of their ideas (having experienced occurrences and reactions first hand) could also doubt teachers' knowledge and interpretation of natural phenomena (Koosimile 2004). These observations could clearly be unfavourable for contextualised teaching and learning.

Despite the unfavourable observations noted above, context-based teaching and learning still has potential for increasing students' motivation and interest. Their learning, retention and recall also could improve through the presentation of real life examples that are likely to engage their interest (Barker and Miller 2000; Boaler 1993). Context-based teaching helps students see links between school science and its applications in everyday life, and is as effective as conventional teaching approaches in helping students learn science content (Bennett 2003; Ramsden 1997; Lubben *et al.* 1997; Dlamini *et al.* 1996). Barker and Miller (2000) explored the level of understanding of basic chemical ideas in thermodynamics and chemical bonding of Advanced (A) level students and the effects of using the Salters Advanced Chemistry course in teaching these concepts. The Salters Advanced Chemistry programme makes use of a context-based teaching approach, which was used in the UK for the teaching of basic Chemistry concepts in thermodynamics and chemical bonding. Barker and Miller (2000) found that the programme had a significant positive effect on a number of the participating students. Students demonstrated an understanding of certain difficult concepts cited positive experiences of the approach and were able to clearly recall specific contexts and chemical ideas. Some misunderstandings still remained, which the authors believed could be attended to by a change in the teaching approaches used in lower classes prior to the post-16 and A level stages. Some areas of difficulty were that students were not readily making the anticipated links between different conceptual areas. This was seen as a possible indication of students' lack of transfer of knowledge from the thermodynamics topic to other chemical reactions.

Transferring knowledge and skills across different topics in one single knowledge domain (subject) is a difficult process for students. Barker and Miller (2000) demonstrated in their study that students had difficulty using thermodynamics concepts to explain other concepts in chemical reactions. Transfer of scientific knowledge to life contexts, which may contain the use of concepts (relevant or irrelevant) from different subjects, as well as personal and social dimensions, may be more difficult. Social issues bring in personal knowledge and value systems that add more complexity to the knowledge transfer process (Yang 2004).

The problems of transfer have also been reported in studies carried out in Mathematics. Wistedt (1994) observed the intense debates that have ensued following unsatisfactory results in mathematics in Sweden, results that have shown that students have difficulties understanding mathematics. Teaching Mathematics by linking it to students' informal knowledge and everyday settings is often advocated. In studying development programmes that required teachers to build on students' experiences and conceptual understanding, Wistedt video recorded lessons; audio recorded group sessions and interviewed students. From the findings of the study, she advises that contextualised tasks that are to be used for teaching must be meaningful to the students. Students make sense of and understand tasks that are grounded in practical experiences although their additional knowledge of a context does not always assist them in reaching the anticipated response to a task. Tasks may invoke everyday images that may help or hinder students solving given tasks. Wistedt (1994) further notes that students respond to a task in two fundamental ways. They may simply ignore the details of the task and solve the problem using their mathematical knowledge. They may also interact with the task and take into consideration all relevant aspects of the task before attempting to respond to it.

Some contexts in mathematics may lead students to some confused state when examples used encourage them to think in different ways than they would when tackling a real world problem. When working on tasks students may use procedures that are based on the real world situation or context presented rather than mathematical procedures (Boaler 1993). This behaviour suggests a non-transfer of mathematical procedures and skills to the context and that students do not connect the mathematics presented in each context with mathematical concepts learned in class. If transfer is to be enhanced through the use of contexts, then it must be recognised that students interpret the same mathematical situation differently. Boaler (1993) warns that although contexts are usually intended to provide meaningful situations from which students can learn, the intended meaning may be misinterpreted or ignored by students, and whatever learning that occurs, tends to be associated with the context and forgotten when the task is over.

Wistedt (1994) recognises that students working on context-based tasks have a challenge of reflecting on their own understanding and constructing links between everyday references and mathematical references. Students are expected to formulate the task as a mathematically relevant problem before they can begin to execute it. Yang (2004) further calls for research on skills transfer from a scientific domain to relevant social domains.

Boaler (1994) also notes that when students learn mathematics using contexts they attempt to integrate real world variables with the mathematics of the task if encouraged to do so. Boaler explored whether students from two schools following different teaching approaches could transfer their mathematical knowledge and understanding to different task contexts when taught following a process-based teaching strategy or taught in a content-based learning environment. Six questions of different sets of contextual information were used for each of the two schools. Of the six questions two were abstract, three were contextual without much engagement of students with the context and one question was contextual and required greater involvement of students in working on the question. This question required students to allocate jobs of specific hour durations to individual employees in a fabric workshop. Boaler (1994) found that in the school following the process-based learning environment, the performance of boys and girls in all questions was equivalent whilst in the school where students were in a content-based learning environment a higher proportion of girls underperformed compared to boys, and an even higher proportion of girls underperformed in the fabric workshop than the boys. However, the students generally underachieved in the fashion workshop question compared to abstract equivalents of the question. Underachievement was attributed to the students taking into account real world variables when working on the task, rather than focusing on its mathematical demands. In the other school following a content-based learning environment, students were not encouraged to consider real world details of context in tasks when working on these tasks. Boys appeared more able to focus on the mathematics (for example number of hours available for different activities of the fashion task) in the task. The girls were found to use common sense and mathematical knowledge when responding to the task, which did not earn them any credit. They seemed to place more value in the situations and conditions of the task and thus experienced greater difficulty of abstracting the mathematics in the task.

Mixing mathematical knowledge and commonsense knowledge can be an unfortunate experience when performing tasks in a way that is sensible in real world encounters but of no use in an assessment task. The contextual situations and the conditions presented in the contexts seemed to interfere with students processing of questions, inhibiting them from abstracting the mathematical concepts from the contexts. In the case of the fashion workshop task, the girls engaged more with the context of the task and its aspects because they knew a lot about the non-mathematics related activities involved in the task context. Sometimes a context-based task may appear to require a student to approach it from a real world situation (such as the fashion workshop). Approaching a task from a real world situation tends to require several variables to be considered (such as the reality of depending on others for doing their jobs and meeting time deadlines). Sometimes the real world approach to a task may be unusually demanding to work on so that students may experience a cognitive overload situation (van Merriënboer *et al.* 2003). It is apparent from Boaler's (1994) study that familiarity with a task may encourage students to engage too deeply with a task and lead them to underachieve.

From concerns that

> These pseudo-real contexts, far from enabling students to see links between the mathematics learned in school and problems encountered in the real world, encourage students to see school mathematics as a strange and mysterious language which is of no use to them in the real world (Boaler 1994:554),

Boaler recommended that if the use of contexts causes students to under-perform, then contexts that involve real world variables should only be used if students are required to use those variables to do the task. Real life situations in contexts are more favourable that fictitious contexts. Contexts need to encourage students to analyse them both mathematically and in terms of real world variables and to think and understand them rather than recall things about the contexts.

In order for students to work on and generate acceptable answers for context-based mathematical tasks they have to suspend reality as perceived from the real life situation and ignore their common sense knowledge that stems from experience.

Teaching, learning and assessment are three salient components of the educational life of a student. The move to integrate learning and assessment and promote assessment for learning seems to match the use of context-based teaching and assessment. Not only has contextualising science been advocated for teaching and learning in Science and Mathematics, some developments have also taken place in assessment in Science. Studies such as those reported by Boaler (1994) and Wistedt (1994) above have looked at issues of teaching and assessing students' learning in Mathematics in a contextualised way.

Contextualising science teaching is believed to make the content more relevant for the students (George 1999; Putsoa 1999; Lubben *et al.* 1995). It also allows teachers and examiners to assess students' abilities to use their knowledge in different situations. Presenting tasks used for learning or assessment in context makes these tasks more concrete, less abstract and less demanding for students (Ahmed and Pollitt 2000; Ahmed and Pollitt 2001).

Students perceive, interpret, react and respond to different tasks in different ways, as noted above. They construct their own meaning in different situations of learning and of assessment. Some contexts may invoke emotions in individual students, while other tasks might be more familiar and more relevant to some students than to others. Students who are very familiar with a context will know more about it, may engage more with it and face the challenge of selecting aspects of the context that are pertinent to answering the questions (Ahmed and Pollitt 2001; Wistedt 1994; Boaler 1993). This may also affect their performance in the task (Boaler 1994). The degree of familiarity with a particular context determines the extent of student engagement with the task, and their convictions and commitment to support such convictions (Dlamini *et al.* 1996). Ahmed and Pollitt (2000) further note that students who may not be so familiar with the context, may think they have not covered the necessary content of the task and therefore cannot answer it. These observations raise questions about whether contexts used in assessment items do in fact reduce task complexity as they are believed to do in contextualised teaching activities.

Contextualising questions does reduce their abstractness but there appears to be more to contextualising an assessment task than reducing the abstractness of concepts. More student competencies seem to be assessed whenever contextualised questions are used. Ahmed and Pollitt (2000) note a few problematic episodes student encounter. For example, students may be assessed on their abilities to identify the actual task they are to execute from the context in which it is set, that is, de-contextualise the question. Students may have to reconstruct the tasks to a scientific representation in addition to answering the question. If they do not do this they often fail to complete the task as intended. Thus, according to Ahmed and Pollitt (2000) the use of contexts in examinations should be based on clear assessment outcomes and should link as closely as possible to the subject matter being tested. Furthermore, they feel that if the intention is to assess students' ability to do science and not whether they can identify science in context, there is no need for contexts in assessment.

Regarding the use of contextualised assessment questions in examinations Ahmed and Pollitt (2000) conducted a study on how students solved contextualised questions. They used questions selected from examination papers from Science and modified for problem-solving. They asked students to talk through their thought processes while working on the questions. Students were video recorded while working on the questions to help researchers' analysis of students' activities. Students were found to react in individual and unpredictable ways to different contexts in questions. In some instances students experienced interference between schemata provoked by context and those provoked by the science content of questions, which led to misunderstanding and errors. More able students seemed able to ignore irrelevant information in the task while less able students tended to answer the questions in terms of context rather than the problem posed. Other drawbacks of poor selection of contexts used in phrasing questions were that contexts could:

- obscure the science ideas that questions are about - particularly so if students assume that information in a given context is relevant to answering the question;
- distract students from accessing their subject knowledge;
- stimulate an incorrect schema by personifying non-person objects (an unrealistic situation); and
- be fictional and lead students to thinking that they cannot answer the question because they have not seen it before or spend time wondering if they missed it when it was taught. In a way students can confuse context and content if fictional contexts are used.

Students were also found to use their everyday language and knowledge to answer questions instead of using scientific concepts, which resulted in a loss of credit, also observed by Boaler (1994). This experience may cause uncertainty among students on how to answer questions that are based on real world contexts which they know a lot about if their everyday knowledge leads them to lose marks. Students use their everyday language and interpretations to answer assessment problems because it makes the best sense to them. For assessors to discredit students for what they may believe to be true can cause extreme concerns for the students if it causes them to fail their subject (Boaler 1994), possibly in the same way as suggested by Koosimile (2004) regarding variations in teacher-student experiences in context-based teaching.

Despite the numerous negative views and experiences in the use of contextualised assessment questions there are still some good reasons that support contextualised assessment. Tasks become more relevant and familiar to some students who may not have developed an interest in the formal aspects of science. The complexity of some tasks may be reduced through the use of concrete objects and terms that enable students to show their understanding of the principles in real cases. An assessment benefit is the examination of students' application of knowledge, a higher order skill less dependent on recall and memorization. Students analyse the questions and decipher the science content in the task in order to engage in a process that demands understanding of both the context and the content embedded in it. Real life contexts, therefore, allow the assessment of a student's ability to select relevant knowledge and ignore what is not relevant, as well as to use scientific knowledge in various situations. A context that is familiar to students is more likely to help them answer a question meaningfully than an unfamiliar one (Pollitt and Ahmed 2000; Ahmed and Pollitt 2001).

Modern theories of learning view students as active participants in the learning environment (Linn and Gronlund 2000, Shavelson *et al.* 1991). Students construct new knowledge from classroom events (and real world contexts) by anchoring new concepts to, and modification of, existing schemata. Students construct individual schemata in idiosyncratic ways. In the same way that different students construct different mental representations of questions; they will construct different mental representations of real life contexts included in questions. Ahmed and Pollitt (2001) view the question answering process through a five stage model beginning with pre-examination learning that is followed by an in-examination reading of tasks. Task reading invokes several schemata in students and stimulates mental representations of what the task demands. Each student forms his or her own representations through the activation of relevant and irrelevant concepts in their minds, which may affect the student's interpretation of the task. Students need to search through the activated concepts for those that match the answer to the task. They then generate and write down the answer in words that best represent their understanding of the demands of the task. Thus, the task reading stage and the concepts invoked by the process represent the most crucial part of the context used in questions. To reduce complexity of context-based question Ahmed and Pollitt propose the use "focused context" (2001:5) in questions. This entails examiners identifying the focus of the context they wish to use in assessment tasks and ensure that all parts of the question lead to answering it by avoiding any side issues that might distract students from the intended focus of the question. The

> … context must be a way of analysing reality and applying science, not merely a setting within which interesting scientific phenomena may be explored (Pollitt and Ahmed 2000:15).

They advise that the context must be an integral part of the question.

Ahmed and Pollitt (2001) also suggest that students' interpretations of tasks are less likely to be different from those of the examiner if the context used in questions is focused. A focused context is believed to activate relevant schemata that help students understand the task and direct students' thinking towards responding in the right way. They found from their investigation that expressing questions in a focused and natural way made questions more valid and more understandable to the students. This observation implies that questions that fit naturally to a real world context will reduce stimulation of irrelevant schemata in students and are more likely to be performed better than unfocused questions. Highly focused contextual questions are less ambiguous in terms of language use. These researchers acknowledge that focusing the context to the most salient aspects that correspond to main issues discussed in the context, does not necessarily reduce task difficulty for students who do not know the scientific concept(s).

Context-based learning supports understanding of school science concepts and procedures and provides the motivation for students to engage in school science but does not automatically help students to access their understanding of school science to deal with everyday situations (Campbell, Lubben and Dlamini 2000). These researchers studied students' level of awareness of social and economic implications of Science, ability to design valid experiments to solve a given dilemma, ability to apply scientific concepts to solve problems and their perceived source of the knowledge used in each of the cases. Paper-and-pencil probes were used with Form II (Grade 9) students in Swaziland who had been taught through a context-based approach that linked school science with industry and technology (Linking School Science with Industry and Technology (LISSIT) materials). Each probe presented a context-based instance of an everyday situation and students were to demonstrate their awareness of social and economic implications of Science; their ability to design experiments, as well as their ability to apply scientific concepts. Students were to indicate the source of the information (books, television, radio, school science, other) they used to respond to the probes. Analysis was by means of a coding scheme from clusters of similar responses generated from a preliminary analysis of a sample of responses.

Campbell *et al*. (2000) found that the use of everyday situations in science teaching does not readily enable students to use school science in dealing with everyday life situations. While students drew on procedural knowledge of experimentation gained from school science, they did not readily utilize learning about social and economic aspects of Science or apply school science knowledge in solving everyday problem situations (Campbell *et al*. 2000). Contextualised learning helps students develop understanding of conventional science content as expected in examinations and to apply scientific knowledge in their everyday life, but this requires the construction of strong links between classroom science and the science applied at work in the community. A two-way flow of knowledge between school and everyday life experiences is necessary for a meaningful impact of school science on students' lives (Campbell *et al*. 2000).

The current use of contexts for teaching, learning and assessment in Swaziland has been encouraged by the outcomes observed by researchers who have spent time and effort exploring the use and implications of contexts in science teaching, learning and assessment. Educators who are in the first place interested in the use of a content-based curriculum and are concerned by traditional assessment models that tend to be devoid of relevant science to students' lives, are trying to find alternative ways of improving this sought after relevance. The developments towards greater inclusion of contextualised science activities in learning and assessment are not suggesting a replacement of currently used assessment models but a search for more efficient ways of using assessment to promote and measure learning.

Strong arguments cautioning contextualising science teaching and assessment have been advanced, such as that:

> [i]f students are taught in context and assessed in context, they may never develop an understanding of the abstract concepts beneath what they have learned. This could result in students seeing every context in which they are taught as a new piece of science content. If they then come across unfamiliar contexts in an examination they may not realise that they have done this before and simply need to transfer their knowledge from a different context (Ahmed and Pollitt 2000:3).

This view seems to be portraying a scenario where students are taught the contexts not that context is used to teach Science. Contexts are meant to serve a purpose for doing school science, but not to replace the science taught. Contexts supplement the teaching of scientific concepts, principles, theories and facts, and provide a useful base for assessment. Teachers need to understand the role of contexts in teaching Science.

The literature above has shown that contexts in assessment can be misleading if poorly chosen and presented. Contexts need to be closely aligned to the content, be realistic and meaningful for students to find executing contextualised tasks worthwhile. In addition it should be acknowledged that students have had years of training and practice in tackling non-contextualised questions. The students and teachers in the studies presented above studies have had very little time and exposure of working with contextualised assessment tasks. Hord (1987) warns that change takes time and more time may have been needed for students to adjust to the use of contextualised questions before positive effects could be observed. The use of the Salters Advanced Chemistry course reported by Barker and Miller (2000), gives some hope for contextualised teaching and assessment in Science.

2.5.1 Validity in context-based questions

Taber (2003) notes with interest the shift towards more contextualised questions results in more complex questions in which students must discriminate between relevant and irrelevant information. In some instances the context may even lead students away from the targeted scientific knowledge, making the validity of the interpretations made about scores awarded to students' responses to context-based questions doubtful. The schemata invoked by the thought processes students engage in while reading and interpreting the rich information provided in contextualised questions can be unpredictable (Ahmed and Pollitt 2001). It has been mentioned above that different students have different experiences and will have different interpretations of contexts presented with questions. The interpretations are influenced by the language competence of students and their familiarity with the context. In spite of these concerns the use of real world context in assessment continues in external examinations (Ahmed and Pollitt 2001) for a number of reasons. Much work has been done in context-based teaching and the advantages documented. However, much research work is still necessary to contribute to the debate on the use of context-based questions in assessment. Suggestions for improving validity when constructing context-based questions are presented by Ahmed and Pollitt (2001) as follows:

- focusing contexts in assessment questions to improve understanding of questions;
- making questions fit naturally to real world contexts;
- including issues that are central to the context, central to the question; and
- selecting contexts that are essential to the questions rather than simply presenting interesting contexts.

Ahmed and Pollitt (2001) found that shaping questions in a realistic manner was more likely to stimulate similar schemata in students as intended by the questions. Students experienced less interference from the context in the question de-contextualising process. Ensuring that context-based questions are free from comprehension problems is a challenge for assessment considering that students' prior experiences influence the operations and constructions they engage in while learning or executing assessment tasks. An additional challenge observed by Lubben *et al.* is that by their "very nature, testing for the ability to identify relevant science ideas to solve a given problem requires large amounts of information" (1997:39). This observation implies that it may not be that easy to cut back on contextual descriptions in constructing context-based tasks.

2.6 PERCEPTIONS OF ASSESSMENT

Students' and teachers' perceptions of assessment affect their reactions towards assessment tasks, interpretations of tasks and scores, as well as the amount of effort, or lack of, that students and teachers invest in preparing for and executing assessment activities (Maclellan 2001; Brookhart and Bronowicz 2003; Aschbacher 1993; Yung 2001).) Madaus and Kellaghan note from historical developments of assessment that the

> … power of tests and examinations to affect individuals, institutions, curriculum or instruction is a perceptual phenomenon: if students, teachers or administrators believe that the results of an examination are important, it matters very little whether this is really true or false - the effects are brought about by what individuals perceive to be the case (1993:4).

Ideas from the literature on the nature and effect of students' and teachers' perceptions of assessment on their behaviour and actions regarding classrooms practice are presented below.

2.6.1 Teachers' perceptions of assessment

Teacher knowledge, beliefs and perceptions affect the way they react to and implement curriculum innovations (Yung 2001). Yung observes growing evidence that teachers have theories and belief systems that play an important part in their cognition and behaviour in teaching. Understanding the reactions of teachers to new assessment models is thus crucial in an endeavour to improve assessment and instruction in schools Aschbacher (1993).

When, according to Yung (2001), teachers' theories and beliefs about a curriculum and its implementation vary from the philosophy of a curriculum innovation they either re-structure their beliefs or adjust the implemented curriculum according to their belief system. As mentioned above, new theories of learning require a radical shift in the philosophy of assessment, its role and relationship to learning to a broader model of educational assessment. Formative assessment requires teachers to adjust their roles of teacher and assessor to become a mentor and judge simultaneously. However, Yung (2001) notes that the effects of assessment reforms on classroom teaching are context-specific and inconsistent. Thus, whether the new views about the curriculum and assessment can be translated into pedagogical form depends on personal beliefs of teachers and their willingness to change.

Yung (2001) demonstrated how teacher belief systems, particularly their conceptualization of the role of assessment can facilitate or inhibit learning. Three teachers implemented a new assessment scheme. The teachers' treatment of the assessment scheme reflected their beliefs about fairness in relation to specific goals and general goals. These were their beliefs about assessment in relation to specific goals of learning, teaching and the development of students' abilities in relation to general goals of maintaining the procedures of public examinations, giving students an opportunity to learn during assessment and ensuring a balanced all round education for students. The teachers dealt with the goals (specific or general) differently. One teacher treated the assessment in relation to ways of achieving the general goals. This teacher maintained fairness through differentiating students' abilities by refraining from giving clues to help students solve the tasks at hand or allowing any discussion among students. Another teacher structured the assessment to intentionally allow conferencing between students and the teacher through the use of questions to get students to think and learn the subject while being assessed. The third teacher maintained fairness in assessment by extending completion of tasks as homework so as not to deprive students a chance to do other school activities. Each teacher extended the laboratory hours beyond the time allocated. Yung (2001) suggests that new models of assessment should not make demands on students and teachers that are incompatible with the learning context.

From the study Yung (2001) observed some confusion among the teachers regarding the use of the same set of assessment activities for formative and summative results. Some teachers tend use assessment results interchangeably and do not distinguish formative and summative purposes of assessment (Bachor and Anderson (1994). Wiliam and Black acknowledge that "significant tensions are created when the same assessments are required to serve both formative and summative functions" (1996: no page number). Shavelson (not dated) also makes similar observations regarding the requirement for teachers to be responsible for both summative and formative assessment. He notes that teachers experience conflict in the use of assessment information to help students narrow the gap between what they can do and what they need to be able to do, while simultaneously judging students' performances for purposes of grading and certification. Yung (2001) advises that formative assessment and summative assessment purposes need to be separated to allow teachers more formative use of school-based assessment to help students learn. He also suggests that teachers need to be trained on how to use assessment formatively. On a similar note Wiliam and Black (1996) are of the view that the tension is caused by interpreting the same evidence differently or taking different actions based on the same interpretation. They propose some approaches to separate the uses of the assessment evidence.

According to Yung (2001) teachers need to be taken on board the new assessment paradigm by discussing the philosophy and intentions of new assessment models during teacher education programmes and meetings. Otherwise teachers may bring their own interpretations that are embedded in traditional testing paradigms that may interfere with the good intentions of school based assessment. Yung supports Gipps (1996 cited in Sebatane 1998) that teachers need to be trained to acquire assessment skills to complement views for assessment for learning. Indeed if teachers are expected to change what and how they teach in fundamental ways they need sustained support to try out new practices and theories (Shepard 1995).

Teachers in another project perceived the grading and ranking of students to be the primary purpose of assessment, although the developmental function of motivating students, diagnosing learning and evaluating teaching was also perceived important. However, the importance of the developmental function depended on how students used the feedback to strengthen their knowledge, develop their thinking and improve their presentation of information. These were observations made by Maclellan (2001) from a study on teachers' and students' perceptions of assessment experiences and practices. Maclellan also observed some inconsistencies between teacher aspirations to formative assessment and their assessment practice, which was influenced by the traditional measurement model of assessment. These observations were seen to be indicative of aspirations towards new assessment models but which were still to be realised in practice. These findings corroborate Yung's (2001) assertion regarding the influence of teacher beliefs in implementing newer views of assessment.

McNair, Bhargava, Adams, Edgerton and Kypros (2003) also observed that teachers' classroom practices do not always correspond to their (declared) philosophies or what they studied in their teacher education programmes. They attributed the mismatch to contextual variables. Teachers continue to feel the pressure to provide students with information that is more directly assessed through standardized tests and provide them with test-taking practices. In such an approach assessment is conducted mainly for accountability and reporting. MacNair *et al.* drew these conclusions from an exploration of the types and frequency of teacher use of paper-and-pencil tests, observation notes and their use, as well as the utility of assessment techniques by primary school teachers. Teachers in MacNair *et al.*'s study had difficulties with the systemic integration of other assessment approaches for the purpose of formative assessment to inform instruction on an ongoing basis. Teachers viewed observations as an assessment strategy that provided information needed to individualise instruction. However, their use of observations was predominantly for behavioural rather than academic issues. That is, the teachers did not fully utilise observations as a source of information on students' academic qualities of skills, knowledge and abilities, how students approach learning, as well as the difficulties they experience. This characteristic makes observation an essential tool in formative assessment. Teachers also experienced some difficulties implementing several assessment approaches simultaneously. They demonstrated a lack of clear understanding of the use of assessment to support instruction. MacNair *et al.* (2003) identified a need to help teachers develop skills of implementing different assessment approaches right from initial training if they were to be adequately prepared for implementing multiple assessment approaches.

Observed inconsistency between teacher aspirations and practice of assessment could result from, for example, teacher lack of sufficient knowledge of the demands of alternative assessment models. Harmon (1995) was quoted by Roberts and Wilson as saying that alternative assessment models can be successfully implemented

> … only if teachers understand their use and depth of the content they demand, are empowered to make instructional decisions, and are supported by school districts which encourage teacher change (1998:2).

The need for taking teachers on board concerning the debates involving the use of alternative assessment models in classrooms is clear, if they are to take these ideas and use them in their classroom. Teachers also need to be part of educational reforms, as well as receive support and assurance that the reforms are in the best interest for their students (Roberts and Wilson 1998).

Roberts and Wilson (*ibid*) conducted a study that evaluated the impact of an integrated assessment system on teachers' assessment perceptions and practices. Teachers' perceptions were explored at three levels: usefulness of alternative assessment strategies for developing students' understanding in Science; usefulness in grading students' work; and their usefulness in classrooms. Teachers took part in a year-long assessment project that was designed to apply new theories of learning and methodologies in assessment to the practice of teacher managed classroom-based assessment. Teachers and students followed a curriculum that was designed to engage students in an issues-oriented, hands-on approach to thinking about scientific issues of relevance to their lives. Student understanding was assessed in an ongoing manner using a specially designed matching assessment system. Findings showed that the teachers appreciated the support received from other teachers in their working groups. They also showed that while teachers perceived alternative assessment to be useful, the reality of their use to guide and grade large numbers of students' papers became a challenge. Teachers continued to use traditional assessment approaches even though they felt that alternative assessment were good, as Yung (2001) confirmed. According to Roberts and Wilson (1998) several authors contend that the reality of assessment reform is fraught with problems and difficulties that may slow the teacher down. Thus, teachers' perceptions tend to be altered by the reality of the implementation of assessment reforms. Other factors identified to affect implementation were quality of leadership, support for teacher professional development, as well as teacher collaboration.

Kampfer, Horvath, Kleinert and Kearns (2001) note that teachers have extensive knowledge and experience from working with students that would be a useful guide for assessment innovations. In a study involving special education teachers in the United States of America, Kampfer *et al.* (2001) used a survey to explore the amount of time and effort teachers required for an Alternate Portfolio Assessment process used in their state. The teachers were also asked to make comments about their use of the assessment. They found that teachers perceived the assessment programme to be time consuming. Components found to be time consuming were creating portfolio entries and completing portfolio components. This experience was attributed to possible inadequate understanding of how to develop and implement entries. The time requirements could be reduced by training the teachers. Time consuming aspects associated with the implementation of portfolios were related to variables of social relationships, progress documentation, and providing support to students. Other issues noted to be affecting implementation of the portfolios were inadequacy of training in the use of portfolios and their scoring, the short timeframe between training and implementation of portfolios and teacher motivation to use the assessment model. Initial training was, however, valuable in illustrating portfolio development and incorporation into regular instruction.

Validity and reliability are the main weakness of alternative assessment. Teachers in Kampfer *et al.*'s (2001) study felt that the rigour involved in the portfolios was more a test of their ability to implement the assessment model than a process to evaluate the effectiveness of the assessment programme. Scores received by high performing students on their portfolio work were perceived to be an inaccurate measure of students' abilities. High scores were perceived to be associated with the degree of student involvement in portfolios, regularity of portfolios in instruction and benefit to students.

James *et al.* (2005) observed that assessment that is aligned to standards measures the degree to which students can demonstrate their understanding and performance relative to the identified standards of learning in a given context. The success of a given assessment model depends on the requirements of the assessment model and other contextual factors such as time, class sizes, proper training to implement that assessment, as well as teachers' perceptions about the assessment approach. Teachers may experience some difficulties in implementing certain assessment models, feel uncomfortable and uncertain about administering them and therefore abandon their use.

Summary on teachers' perceptions

It is shown in the above discussion that teachers do hold beliefs, conceptions or perceptions about assessment practices that influence their implementation of assessment tasks. The need for training of teachers to acquire assessment skills is evident so that teachers can adequately understand the value of many of the principles and benefits that can be derived from new assessment formats. It is also evident that assessment innovations bring in new demands on teachers and sometimes instructional time and school resources may necessitate support of the school administration and other teachers.

2.6.2 Students' perceptions of assessment practices

"What do kids think about the assessment practices that are used to evaluate them?" This is a question asked by Schäffner *et al.* (2000:3) as an issue that researchers of assessment matters seem to have omitted over the years. Assessment practices impact on students' lives, motivation and other affective aspects such as self-esteem, confidence, as well as how they perceive their school subjects (Schäffner *et al.* 2000). This section of the review discusses research work and findings on students' perspectives regarding their assessment.

As Schäffner *et al.* (2000) above note, rarely do discussions on classroom assessment focus on what students think about the assessment work assigned to them or how students perceive the use of such assessment and the grades awarded for assessed work. Most of the discussions are on what teachers do with students, how they interpret and use or should use assessment information (Brookhart and Bronowicz 2003). Studies on students' perceptions about assessment have shown that such perceptions have considerable influence on how students approach their learning. Students have also been found to have strong views about different formats of assessment (Struyven *et al.* 2003).

Students' perceptions of the assessment tasks they are working on, their experiences, perceived self-efficacy and perceived importance and value of assessment tasks determine the extent to which assessment shapes and drives student efforts on tasks or their learning (Maclellan 2001; Brookhart and Bronowicz 2003). Students perceive task importance and value in relation to their goal orientation (perceived need to execute a given task) (Brookhart and Bronowicz 2003). Students who perceive a need to understand the material embedded in an assessment task will successfully negotiate a solution to the task and in the process engage in deep learning or consolidate previously learned material. Tasks that students perceive to be on rote learned material are less likely to encourage students to invest effort in higher order objectives or deep learning. In short what is important to be assessed strongly determines what is considered important to be learned (Maclellan 2001, Wiliam 2001). Student perceptions of assessment were therefore important in directing students' learning, establishing learning outcomes and the overall goals they had to achieve.

Brookhart and Bronowicz (2003) explored students' perceptions of task characteristics - interest, importance and value, task difficulty, as well as perceived self-efficacy and goal orientation using a multiple-case study. They found that students perceived teacher expectations of studying as stimulating them to prepare for tasks by rehearsing and going over previous work in order to prepare for the tasks. Primary school and high school students exhibited differences in the form of studying they adopted, with the primary school students focusing on quality of outcome of studying and high school students focusing on the studying process. High school students also linked interest and grades to the importance of task, although importance was not necessarily linked to interest. Perceived self-efficacy was linked to the content and mechanics of responding to the task for the primary school students. Perceived task importance and meaningfulness seemed to stimulate high school students to care about how others performed in the task.

Perceptions of students and tutors were explored by Maclellan (2001) through a questionnaire dealing with different aspects of assessment: experiences, purpose, nature and demand level of tasks used, timing, procedures for marking and reporting. Students in the study seemed to have an underdeveloped conception of assessment and its benefits. They were not aware of the role of assessment in learning such that they could not exploit fully the assessment experience to improve their knowledge level.

Maclellan's (2001) findings are supported by Gibbs and Simpson (not dated) that students tackle assignments intended for learning to maximize marks and progress rather than to maximize learning achieved from engaging in those tasks. According to Maclellan (2001) it is important to involve students in the assessment processes if there is to be success in transforming their learning through assessment. To assist students in this endeavour, teachers too need to understand the benefits of involving students in assessment processes so that they can implement changes in assessment practices that necessitate students taking greater responsibility in their own learning and improve their metacognitive capacities.

Segers and Dochy (2001) also agree with Maclellan (2001) by noting that students' perception of their learning environment (teaching, learning and assessment) is important in interpreting their learning outcomes. Segers and Dochy investigated students' perceptions of the learning environment and learning-assessment environment in schools that had adopted a problem-based learning approach. Students worked in tutorial groups to analyse and discuss the assigned problems. Students also generated learning objectives from the tasks, which they used to process the subject matter. Students' ability to handle problems such as to analyse and solve novel and authentic problems were assessed through an overall test – a written final examination designed for problem-based learning. Students had no prior experience of working on the innovative tasks. A questionnaire and semi-structured interview were used to obtain data. Innovative assessment models used were self-assessment and peer assessment. Students seemed to appreciate the new innovative assessment models, identifying them as stimulating their thinking, learning, critical reflection and structuring of the learning process. However, the students perceived the education process rendered by the new assessment models as hindering the reaching of goals. Students were content with memorizing information from textbooks and reproducing it as expected in previous traditional assessment models rather than engage in assessment models designed for enhancing learning such as the "Overall Test" used in the study.

Segers and Dochy (2001) indicated that group assessment needs time to allow for concrete discussions and reflection by students and more time for students to carry out the analyses and explanations required in the novel tasks. Since students had no prior experience of working under the conditions of the innovative tasks, it became evident that teachers would need to improve their instructional techniques in order to align them with new programmes or modes of assessment, as the one investigated

Gijbels, van de Watering and Dochy (2005) report on a study purporting to establish the effects of a written assessment task integrated into a problem-based learning environment on students' perceptions and performance. Six assessment tasks were completed by one group of students on a voluntary basis while the other group did not, but both groups wrote a final examination comprising multiple choice items at the end of the course. Members of staff involved in the course were interviewed and they completed a questionnaire. Integrating the tasks with classroom instruction was found to encourage the students to study more critically and systematically. Gijbels *et al.* (2005), however, advise that the results of the study may be inconclusive on the effects of the assessment tasks on students' performance. Students worked on the tasks on a voluntary basis so that it may be possible that better performing or highly motivated students participated, thus the observed better performance. Another possible contributing factor to the improved performance was that the participating students received more teacher time in the form of feedback on the tasks. On a positive note, these cautions indicate a possible success of problem-based assessment under the close mentorship of students by their teachers.

Students' perception of a task will influence the time and effort they invest in the task. Perceived importance, usefulness and value of engaging in the task are thus motivators for student effort (mental effort and the overt physical act of doing the activity). In a study seeking to test a theory of classroom assessment Brookhart and DeVoge (1999) worked with third grade language arts classes over four classroom assessment events. They sought to describe the level of perceived task characteristics, perceived self-efficacy, amount of invested mental effort, achievement and the relations between these four events. Students displayed both goal oriented (the task had some learning value) and performance oriented (getting good grades) reasons for investing effort in the tasks. The amount of effort invested was linked to students' goal orientation, that is, the reason or motivation for studying. Another motivation for students' efforts was perceived importance, value and future functionality of the task. If students identify some value in a task they are likely to devote more time and effort. For example, the value of the task in terms of learning benefit and value of the task in increasing their score, were two task characteristics students identified with in this study.

Students play an active role in constructing knowledge about their subject through their experiences with assessment tasks and in their interaction with each other (Moni, van Kraayenoord and Baker 2002). Recent shifts to authentic classroom-based assessment models which encourage learning that is integrated with assessment, collaborative assessment achievement and self-assessment imply a more active role by students in their assessment. Students also develop other attributes such as beliefs, attitudes, practices and understanding about assessment. Students' attitude to assessment affects their participation in assessment and the value they place on assessment method and task. Exploring students' perceptions of assessment is one way of increasing their input in assessment (Moni *et al.* 2002).

Moni *et al.* (2002) report of a study on perceptions of Year 8 students regarding literacy assessment in English. Students' open accounts and interview transcripts were openly, transparently and selectively coded in order to identify appropriate themes for reporting. The study revealed that students had varied assessment backgrounds when they entered secondary school level that affected their understanding of assessment practices at this level. Students were more confident in tasks that appeared familiar and perceived as revision particularly at the beginning of the year. Such confidence was maintained as the year progressed. Familiarity in genre and topics had a positive effect on students' confidence while their confidence was reduced in unfamiliar tasks, which were also disliked. Unfamiliarity also generated weariness in students at the beginning of a task, but this soon faded as students continued working on the task. Moni *et al.* suggest that students develop tightly held views about the nature of assessment tasks and about their impact on students' lives from previous experiences with assessment. As students progress through school they may become increasingly negative and concerned about assessment of their learning and may resist assessment tasks. If teachers understand that some students hold such perceptions they can explicitly expose their students to tasks that deliberately channel the resistance to assessment towards developing strategies that benefit students. The designs of tasks need to take into account that students have prior assessment experiences.

Summary on students' perceptions

The studies and reviews described above reflect useful findings regarding students' perceptions of assessment for consideration in classroom use of assessment. Students'

- perceptions;
- experiences;
- perceived self-efficacy; and
- perceived importance and value.

of assessment tasks determine the extent to which assessment shapes and drives student efforts and time invested in tasks or their learning. Students' also perceived that teacher expectations influence the extent of their preparation for the task.

Students may appreciate new innovative assessment models and find them stimulating for their thinking and learning. Students were also encouraged to reflect on and structure their learning process. However, students may have an underdeveloped conception of assessment and its benefits. They may not be aware of the role of assessment in learning, and as a result not fully exploit assessment experiences to improve their knowledge levels.

There were suggestions that in order to assist students to fully benefit from assessment processes teachers need to understand the benefits of involving students in assessment processes. In such instances students can also implement changes in assessment practices that necessitate their taking a greater responsibility in their own learning and hence improve their metacognitive capacities.

2.7 GROUP ASSESSMENT

Much research has shown the advantages of using co-operative learning in classrooms. Through co-operative work students develop social skills, professional working skills, as well as build a sense of community within the classroom. Group work also enhances student achievement and self-esteem (Griffin, Griffin, Warkentin, Quinn, Driscoll and McCown 1995). These authors note that co-operative learning augments the extent to which students actively process content by affording them the opportunity to discuss and negotiate meanings with each other. Discussions encourage students to reason at higher cognitive levels. Students also receive social support from each other which can promote their persistence on challenging tasks, reduce frustration, increase autonomy and contribute to academic and career aspirations. Co-operative learning requires students to learn from each other through explaining their point of view, listening to others, give and receive help from each other, as well as help each other develop deeper understanding of the material being learned. Griffin *et al.*'s (1995) observations of the benefits of using co-operative assessment tasks are corroborated by suggestions like those by the Southwest Educational Development Laboratory (1996) that if students spend a lot of time working cooperatively on learning tasks, they should also be assessed in settings that are similar to those used in their learning settings.

Griffin *et al.* (1995) also observe the increase in research that is extending co-operative learning principles and benefits to co-operative assessment. They also note an increase in the use of various forms of peer, collaborative and co-operative learning in small group activities to achieve learning outcomes. When students work collaboratively with each other they deepen their understanding of subject content and may also take greater responsibility for their learning. The absence of peer collaborative assessment has been noted with concern to be undermining co-operative learning despite its benefits. Thus, collaborative learning has extended to assessment approaches (Boud, Cohen and Sampson 1999). Other co-operative assessment benefits reported in the literature included an increase in comprehension and recall, improved test scores and students' preference for co-operative test taking, increased motivation, as well as increased quality of communication among students and the teacher while the test is being executed.

Like other educational processes, co-operative assessment has drawbacks and these were identified by Bilsky-Torna (1993 cited in Griffin *et al.* 1995) as stifling academically stronger students and weaker students riding on the coattails of stronger students. Griffin *et al*. (1995) also observed that some students tended to view group assessment as a way of getting free marks and decreased their prior preparation for an examination, as well as monitoring of their strategies for studying. Students may thus fail to recognise the importance of group assessment as a tool for learning.

The effects of co-operative assessment on student qualities like individual achievement, depth of processing information and perceived ability were investigated by Griffin *et al.* (1995). Students were given tests individually and then in groups. Students had positive perceptions of group assessment recognising its benefits on grades, and increased understanding and learning of content through discussions. Students' perceptions depended on how group assessment was used. Some perceptions were specific to certain aspects of assessment. For example, making students discuss examination questions encouraged them to be responsible for constructing and acquiring knowledge from the questions with the help of each other. Another example was that co-operative assessment was also a problem-solving task for the students. Students focused on explaining to others why certain responses were correct or incorrect, clearing confusing issues, evaluating each others' reasoning and determining best answers and rationales for answers. Students valued comparing different viewpoints as they discussed different perspectives of the task. They also indicated that they monitored their peers' discussion when there were disagreements regarding answers to the tasks and especially, when there was an even split of group members arguing for different responses. Discussions raised many and diverse answers to tasks allowing students to clarify, re-examine and evaluate their knowledge of the information discussed. However, students restricted their discussion to questions they had responded to differently in order to clarify the differences.

The diverse views and abilities of students in groups may determine the extent of group discussion and productivity. Group productivity has implications for group scores and therefore motivation of the group members. In a study exploring the effects of group ability composition on group processes and outcomes, Grade 7 and Grade 8 science students were assigned hands-on performance assessment tasks individually, and then in groups. The students also participated in a written test containing some items analogous to the hands-on tasks. Webb, Nemer, Chizhik and Sugrue (1997) observed that for medium-low-ability or low-ability students working with more able students improved the quality of their responses and their scores during group problem solving or individual work. Achievement of high-ability students was generally not affected by group composition.

The effect of group composition was observed to be particularly strong for low- and medium-ability students and less so for high-ability students. When low- and medium-ability students worked with high-ability students they got help and performed better in the group test and achieved higher scores in subsequent tests. Groups with high-ability students tended to give more correct answers and quality explanations. Low-ability students learned from the high quality group discussions and participation.

Similar findings were reported by Fawcett and Garton (2005) from a study on the effects of peer collaboration on primary school students' problem-solving abilities. Students worked in pairs of specific abilities based on a classification that was based on pre-test results. High-ability, low-ability and mixed-ability pairs of students worked on tasks in interactive (talk) and non-interactive conditions. In the interactive situations students were allowed to discuss and share ideas and in the "no talk" situations discouraged from any oral exchange of information. Fawcett and Garton (2005) found that students collaborating with the same ability or lower ability peer or working independently showed no significant improvement in the abilities tested in the task between the pre-test and post test.

High-ability students were observed by Fawcett and Garton (2005) to regress from pre-test to post-test activity except when working with low achieving peers in the "talk" situations. Regressing in performance was attributed to students' lack of need to internalise new information and a lack of accepting the challenge to construct new meanings. The improvement of high-ability students in interactive conditions with low-ability students was explained in terms of verbal exchanges between the students, such that the expert students explored variation in their own and their peer's knowledge then restructured their own knowledge and thinking, corrected misconceptions, filled gaps in their understanding and developed other strategies for solving similar problems. These students could also have improved their understanding through consolidation of learned material. Students working in interactive conditions improved more than those in the non-interactive condition. These results demonstrated advantages that would serve just as well in assessment conditions as Webb *at al*. (1997) reported.

Findings from Fawcett and Garton (2005) are in support of an earlier study by Tao (1999). Tao (1999) also found improved performance for students who worked on qualitative problems and tests in pairs over students who worked individually. Improved performance was attributed to students experiencing conflict and co-construction that were conducive to problem-solving. Students' success in problem-solving is not necessarily a matter of their ability, but can be linked to the nature of their interaction with each other in a group and the use of subject concepts and principles in the discussion. Whenever conflicts arise, students are forced to resolve the conflict by reflecting on and reviewing their own understanding and justifying their positions.

The observations and conclusions highlighted above showed that pairing students with other students or more expert students in social interactive conditions provided better learning opportunities. Students shared ideas among themselves before these ideas developed into internal schemata. New constructions or reconstructions of ideas were enabled through the interactions. Peer collaboration provides supporting conditions that allow students to explore their ideas freely and to check them against those of others in the team. Talk and activities on shared tasks play an important role in the development of students' understanding of subject knowledge (Tao 1999).

Pairing students on merit had advantages for some students and disadvantages for others. The findings reported by Fawcett and Garton (2005) support the questioning of the fairness of using different group composition on assessment through collaborative group work posed by Webb *et al.* (1997). As shown in Fawcett and Garton's (2005) study, some group compositions were more advantageous than others, concerning group productivity and performance of the group and individual, and student learning. Webb *et al.* (1997) recognised that a competent member in a group places that group at an advantage but may also encourage dependence or dominance by certain individuals in the group. The competent member may feel exploited by the other group members. These observations show that assessing students through group work brings in some complexity not experienced in individual assessment. There is no guarantee either, that group composition fairness can be achieved by distributing high-ability students evenly among the groups (Webb *et al.* 1997). Neither is there a guarantee that low-ability students will necessarily learn from their high-ability peers (Fawcett and Garton 2005). Training students to take assessment in groups may be necessary to develop skills of working together.

2.7.1 Group assessment and learning

Group work increases student learning through collaboration among students. Such collaborative learning also increases social and emotional outcomes such as social skills, self-esteem and attitudes towards others (Webb *et al.* 1997; Fawcett and Garton 2005).

Within the group students provide each other with information which can be used by individuals to assess their own knowledge and skills. They can even extend this awareness to modify existing ideas in light of the input from the group members. Contributions from individuals within the group are discussed until the group agrees on a particular idea or response to use in the assessment task (Boud *et al.* 1999). The sharing of ideas also leads to improved performance by students working on an assessment task in groups, particularly when they resolve conflicts and establish cognitive equilibrium.

2.7.2 Group assessment and marks

The justification for group assessment is based on the grounds that if collaborative learning and teamwork are valued, they must be reflected in an assessment process that emphasises judging of students' work on the basis of their collective efforts (Boud *et al.* 1999). While collaborative group work is a common occurrence in science practical work, group assessment is new and difficulties can be experienced due to unfamiliarity with the assessment format and other complexities as noted by Webb *et al.* (1997). Boud *et al.* (1999) noted that students who are used to their work being judged individually can be quite resentful of others gaining credit for what they perceive as their own contributions, particularly in instances of high competition among students for class positions and prestigious awards in recognition of their abilities and effort.

The issue of marks and high grades is seen as highly important in accounting for students' achievement in learning. Students place great value on grades for the encouragement they get to engage in learning activities. If there are no grades, particularly in courses where grades strongly feature, then the un-graded tasks are regarded with lower status and less effort is invested in them (Boud *et al.* 1999). These are more than often the effects of performance oriented perceptions.

2.8 IMPLICATIONS OF THE LITERATURE FOR THIS STUDY

The literature survey described above, illustrates the broad nature of this study. The study looks into students' and teachers' perceptions of performance and context-based assessment as alternative ways of assessing students' learning and as ways of aligning assessment to curriculum implementation in response to current theories of learning.

The literature on performance assessment provided information on the characteristics of performance assessment tasks adopted in this study. That is, performance assessment tasks were constructed and used as hands-on science practical tasks that required students to:

- manipulate scientific equipment to generate data;
- record, analyse and interpret the data;
- draw relevant conclusions from the data; and
- use the conclusions to explain real world practices.

It also provided guidelines for constructing reliable and valid performance assessment tasks and scoring rubrics. In addition, the studies provided some methodological ideas for the procedures to use when administering and scoring the tasks. These guidelines and procedures are outlined in Chapter 3. Thus, some of the more relevant benefits from the literature in regard to the methods of administering performance assessment tasks involved sharing the main scoring criteria with the students before they worked on the tasks, as suggested by Slater (not dated) and Gipps and Stobart (2003). This sharing was intended to help students know in advance what was expected when they performed the tasks and therefore improve their achievement. Another suggestion by Slater used in this study was that students' performance improves if they interact among themselves and with the teacher while they perform the tasks, thus group assessment was used and teacher mentoring allowed.

As far as context-based assessment was concerned, two implications of the studies enumerated above can be identified for this thesis. It is important that teachers are aware of the demands of context-based teaching so that they can meaningfully use real life contexts to develop students' conceptual understanding in Science, and be constantly on the lookout for opportunities to do so. Attempts were, therefore, made to involve teachers who had some previous exposure to the Matsapha Lessons and, therefore, familiar with the contextualised teaching approach. Familiarisation workshops were organised for the participating teachers to improve on their effort to assist students to link school science to their lives. Another implication involving assessment was the necessity to ensure that the context in assessment questions was focused on the most salient points of the task in order to reduce the provocation of irrelevant concepts and therefore, improve the validity.

Studies on perceptions about assessment and group assessment provided some guidance on the collection, analysis and interpretation of the data on perceptions of students and teachers. Thus interviews and questionnaires focusing on students' and teachers' considerations of good and not so good aspects of the tasks, as well as the method of administration, were used. Group assessment was adopted for administering the performance tasks based on the recognition that it is an acceptable way of assessment which tends to align assessment to teaching. Group assessment was also used due to equipment shortages in the schools.

The next chapter presents the methodological design of this study and takes into consideration the ideas and views advanced in literature.

3. CHAPTER 3

RESEARCH DESIGN AND METHODOLOGY

3.1 INTRODUCTION

This chapter deals with the research design of the study, the selection of participants, the research instruments

and their development and validation, teacher orientation workshops, ethical considerations, data collection and analysis. It also discusses methodological challenges experienced during the course of the study.

3.2 RESEARCH APPROACHES IN EDUCATION

Educational researchers choose what research design to use from two main research approaches: quantitative and qualitative approaches. These research designs fall in a continuum from quantitative to qualitative approaches.

Quantitative research approaches involve fixed, deductive designs in which what is to be done and the procedures to be followed are specified before the main part of the study (Krathwohl 1998; Robson 2002). Studies involving quantitative research designs describe phenomena using numerical data that is analysed through various statistical procedures. Large samples of participants are used to produce statistically meaningful data and generalisable findings. Quantitative research designs involve stating and testing of hypotheses. They also maintain controlled contextual factors that might interfere with the data or its interpretation (Gay and Airasian 2003).

Qualitative research is flexible and inductive in design and uses verbal narratives to describe data from which explanations about a phenomenon under study emerge (Krathwohl 1998). Experiences (feelings, beliefs, thoughts and actions) of the participants and the events in their environment are interpreted in terms of the meanings participants bring or develop in the setting (Winegardner (not dated)). Interpretations of participants' experiences and meanings are described using their words although numerical data may be used to elaborate on the findings (Schumacher and McMillan 1993).

Through a qualitative research approach the researcher is able to work closely with the participants in their natural environment while their experiences, processes and events in their lives unfold. There is high interaction through talk with the participants and use of different strategies to obtain information from them and their environment (Robson 2002). In qualitative research all observations are important and the test for validity is through triangulation rather than statistical means.

Qualitative research approaches can follow different designs such as case studies, ethnographic studies, historical studies and legal studies (Schumacher and McMillan 1993).

3.3 DESIGN OF THE STUDY

Case study designs allow researchers a high level of flexibility, and are tolerant of low pre-specification of data to be explored. The design of a case study evolves and develops as the research progresses (Robson 2002). In addition, in case study designs the case may be an individual, a set of individuals, a situation or an event a researcher wishes to study (Krathwohl 1998; Robson 2002).

Case studies do not seek objectivity or generalisability of their findings. Instead, they provide detailed descriptions of the phenomena under study to enhance reader understanding of the phenomena. Descriptions allow readers to form their own interpretations and understanding of the phenomena by integrating case study data to their knowledge and allow them to form their own generalisations to populations of their experiences (Winegardner (not dated)).

Case studies are, however, labour intensive in terms of data collection, analysis and reporting. The thick descriptions that give case studies their character tend to result in lengthy documents (Winegardner (not dated)).

Case study design was preferred and used for this study because the demands of the study for data collection and the nature of the data sought made the case study more appropriate. According to Gay and Airasian (2003) and Schumacher and McMillan

(1993) the case study design allows for the investigation of one phenomenon within its real life context so as to develop an in-depth understanding of the phenomenon, regardless of the number of sites or participants involved in the study. If more groups are involved as in this study, the groups are not viewed as statistically comparable or mutually exclusive, but the groups are used to provide information for the focus of the study. Students from the different schools were treated as one group. Teachers too were treated as one group.

In addition, the case study approach served other useful functions for this study. It allowed an exploratory research investigation on the perceptions of students and teachers, as well as the development of theoretical constructs.

3.3.1 Case study in exploratory research

Case studies are appropriate for exploratory and discovery research. They can be used as studies into topics for which very little previous research is available. They are also useful as precursors for further enquiry (Robson 2002; Schumacher and McMillan 1993). In Swaziland there appear to be no reported studies on perceptions about performance assessment. Only a few studies have revealed students' and teachers' views about contextualised science teaching in Swaziland, such as those by Dlamini *et al.* (1996), Campbell *et al.* (2000) and Dlamini (2003). This lack of research on perceptions about performance assessment and context-based assessment models in Swaziland made the case study design appropriate for establishing perceptions about these assessment models. Findings of this study may set the scene for further research in the use of these or other assessment models and perceptions about such use in Swaziland.

3.3.2 Case study in developing theoretical constructs

Schumacher and McMillan (1993) further note that case studies are useful in developing concepts or constructs from data through inductive analysis of observations and data. Data analysis in this study involved constant comparative analysis through which statements from respondents or segments of data from these statements were continuously compared to others in developing theoretical categories and themes that were used to describe perceptions of the participants. The level of detail in data collection for case studies allows one to derive theoretical constructs through constant comparisons.

3.4 PARTICIPANTS AND THEIR SELECTION

Convenience and purposive sampling were used to select participants for the study. According to Cohen and Manion (2000) convenience sampling allows a researcher to choose, from a research population, participants to whom access is easy given contextual constraints, and is an appropriate sampling technique for case study research. Purposive sampling enables the researcher to choose participants based on experience and knowledge of the group to be sampled (Gay and Airasian 2003). Familiarity with teachers and schools gained from a nine year period of working in pre-service and in-service positions made the selection easier. However, Swaziland, being a small country, with a population of just over one million, adhering to ethical requirements was a challenge. Thus, specifying selection criteria that would make it easy for certain individuals to identify the participating schools, such as proximity and accessibility, was avoided. Nonetheless, the following were used as priority criteria for selecting participants for this study.

i) Willingness of the teachers to participate in the study. The willingness of the teachers was established through informal discussions and requests to the individual teachers. This criterion was important because the study was intensive and extended over four school terms depending on teaching progress in schools and thus called for perseverance.

ii) Willingness of the concerned members of the school administration; namely Head Teachers and Heads of the Science Departments to allow the school to participate in the study.

iii) Previous involvement of the teacher in the development of contextualised curriculum materials and/or the use of the materials in their teaching.

iv) In-service education and training (INSET) schools were preferred although the working conditions did not allow for the use of only INSET schools. Thus non-INSET schools were also used. Only two INSET schools were eventually used as one withdrew.

INSET high schools in Swaziland are a selection of eight schools that are used by the Ministry of Education and the University of Swaziland (UNISWA) - Department of In-service Education to host workshops for science teachers. In 1998 each of these schools was supplied with a large consignment of science equipment to enable them to run INSET regional workshops.

Table 3.1 below presents a summary of the classes and number of students in each of the participating schools. The schools were codenamed M1 to M4. A detailed description of the schools and the teachers is provided in Chapter 4, Section I.

Table 3.1 Summary of participating schools

School	Number of classes	2004: Number of Students in Form II	2005: Number of Students in Form III
M1	2	40 Form IIA 39 Form IIB	31 Form IIIA 26 Form IIIB
M2	2	40 Form IIA 43 Form IIB	26 Form IIIA 38 Form IIIB
M3	2	52 Form IIA 52 Form IIB	49 Form IIIA 43 Form IIIB
M4	1	44 Form IIB	36 Form IIIB
Total		**310 Form II**	**249 Form III**

Table 3.1 shows that at the beginning of the study in 2004 there were three hundred and ten students and two hundred and forty-nine students at the end of the study. The table also shows that in three of the schools there were two classes participating. In these schools the participating teachers taught two Form II Science classes and wanted both classes to participate in the project. The teachers felt that it would be unethical and difficult for them to teach two classes at the same level and treat them differently.

3.5 INSTRUMENTS AND THEIR DEVELOPMENT

Data collection did not target the teaching or assessment approaches or their effects on learning, but the perceptions associated with the students' and teachers' experiences of conducting the assessment approaches and what they thought about these thereafter. Data used to describe the participants' perceptions was thus obtained through the following instruments:

1. Open student group questionnaires that focused on feelings and experiences of using performance assessment tasks and context-based questions.

2. Students' group interview schedule focusing on feelings and experiences of using performance assessment, context-based questions and multiple assessment tasks.
3. Student open questionnaire focusing on feelings and experiences of using context-based questions.
4. Teachers' interview schedules focusing on views of performance assessment tasks and context-based tests.
5. Observations of lessons on performance assessment tasks.

Interview schedules for both teachers and students were developed from ideas obtained from studies on perceptions about assessment practices such as Brookhart and Bronowicz (2003), Maclellan (2001) and Yung (2001).

Before perceptions could be ascertained the participants needed to experience the use of performance assessment tasks and the context-based tests. As mentioned in Chapter One the Matsapha Lessons were used for teaching the content in the topics Electricity and Air and Living Things to complement the assessment approaches that were to be used by the participants. Studies conducted using the Matsapha Lessons, such as those by Lubben *et al.* (1995; 997) and Dlamini *et al.* (1996), demonstrated the success of using these lessons in teaching the SWISP syllabus, for which they were designed as alternative (not substitute) teaching material. Teachers identified well with these materials, they allowed the teachers to engage in learner-centred teaching approaches. With these materials, teachers were able to bridge the gap between school science and science embedded in contextual experiences. Thus, these materials proved to be appropriate for the contextualised background within which this study was conducted. The assessment experiences of the participants' were provided through three performance assessment tasks and two paper-and-pencil context-based unit tests. The development of these assessment instruments and the research instruments is described below.

3.5.1 Performance assessment tasks

Performance assessment in this study was used to refer to assessment activities that involve students in "hands-on" practical activities where students manipulate material (scientific instruments and equipment) to generate a product (Shavelson and Baxter 1992). The product illustrates the students' conceptual and procedural understanding and the students' abilities to use that knowledge to reason and suggest solutions to a given problem situation. Products of assessment take different formats, for example textual, pictorial, kinaesthetic (physical action) work or a combination of these (Rhoton and Bowers 1997).

a) Structure

A task development shell proposed by Solano-Flores, Jovanovic, Shavelson and Bachman (1999) was used when developing the performance assessment tasks. Each task indicated clearly what students were expected to investigate in four sequential stages, as the example presented in Figure 3.1 below shows. The tasks are provided in Appendix IIA and Appendix IIB.

Stage 1 of each performance assessment task was on planning and design. Students were required to first draw up a plan outlining a procedure on how they would arrange the equipment to collect and record the data for the given investigation. Stage 2 was the hands-on investigation. Students were required to perform the investigation according to their plans, that is, assemble the equipment in order to make observations and measurements, as well as collect and record appropriate data. Stages 1 and 2 tested students' ability to plan and carry out an investigation. Stage 3 focused on the analysis and interpretation of the data, while in Stage 4 students used the data to respond to given questions. In the last two stages students were assessed on their ability to:

- analyse data (draw patterns, perform calculations of required variables);
- interpret data and draw conclusions; and
- recognise and select data that was relevant for use or explaining a given context.

All the tasks followed a similar format: plan, perform, analyse and use the data to respond to questions on a real world context. Teachers needed to approve the plans before students proceeded to the hands-on stage and the rest of the task. This was to ensure that the plan would produce useful data for the subsequent stages of the task.

Electricity Practical Performance Assessment Task 1

Group Number ……… School ……………………………… Date ………….
Names of Group Members ……………………………………………………………………

Instructions:
You will work on this practical task in groups. You are allowed to discuss among yourselves in your group and you may check with the teacher for some information you may need.

Assessment Task 1

In this activity you will plan and carry out an investigation on the electrical conductivity of a set of material.
You will use an electric circuit to test the materials.
You will then classify the materials into conductors and non-conductors of electricity.
You will also choose the most suitable material to use in a given situation.

Material provided circuit board connecting wires 2 x 1.5V cells, switch lamp and holder various materials labelled P, Q, R, S, T, U.

Procedure
Use the space below to write out your plan of how you will test the given materials and indicate how you will record your results. Your plan should show the necessary diagrams. Show your plan to the teacher before you begin testing the material (*Working space provided*).

1. Connect your electric circuit using the material provided and test that it is working.
2. Use the circuit to test for the electrical conductivity of the given material.

Questions
a) Classify the materials you tested into conductors and non-conductors …(2)
b) Why have you classified them in this way? (Two-line writing space provided).…(2)
c) The picture below shows the top part of power-cable pole.

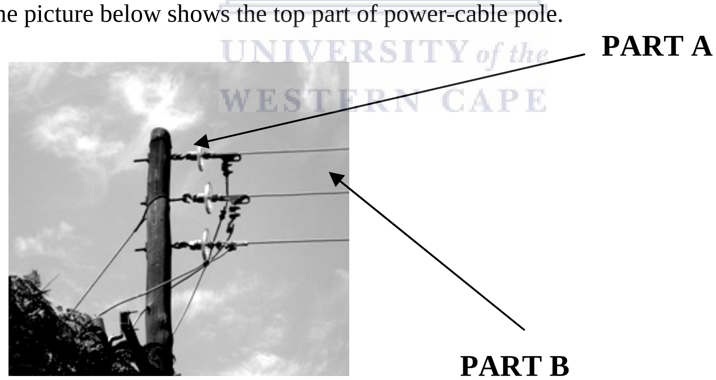


Figure 1 Top end of electric power-cable pole

i) Which of the materials you tested would be most suitable for replacing **Part A,** labelled in the picture?……..………. (1)
ii) Explain your answer. ……… (2)

d) i) Which of the materials you tested would be most suitable for replacing **Part B,** labelled in the picture? …. (1)
ii) Explain your answer.………… (2)

Figure 3.1 Example of a performance assessment task from the Electricity unit

Three performance assessment tasks were planned for each unit in order to provide students with reasonable experience in working on the tasks. However, because of time constraints, only three tasks could be used for the two units. The development of these tasks and their scoring guides is described below.

b) Steps in constructing performance assessment tasks

When constructing the performance assessment tasks and their scoring rubrics, the content validity of the tasks and the reliability of scoring using the rubrics were ensured by following guidelines outlined in Linn and Gronlund (2000) and Borich and Tombari (1997). According to the guidelines, prior to constructing the assessment tasks, learning outcomes for the units must be constructed. Learning outcomes describe specific knowledge, skills and other attributes of individuals that are important for students to display after a learning experience (OECD/PISA 1999:11). Learning outcomes were constructed for the two content units. An analysis of the junior secondary Integrated Science syllabus aims, the objectives and the unit content was conducted before the learning outcomes could be developed. After construction the learning outcomes were checked against the syllabus, as well as the teaching materials to ensure that the outcomes addressed the syllabus requirements and the teaching materials. These were validated as described in section 3.6 below. The syllabus aims, objectives and content, as well as the learning outcomes for the two units, are provided in Appendix IA and Appendix IB. The learning outcomes were used to guide subsequent steps in the construction of the assessment tasks to ensure item and content validity, that is, the representation of the unit content and skills (cognitive, intellectual and physical) (Popham 1999).

The next step was to isolate the learning outcomes that focused on skills that could not be easily assessed by objective type questions. In this study the learning outcomes of interest were mainly those on the psychomotor abilities and procedural knowledge of investigations. Performance assessment tasks were then designed to address the selected outcomes.

c) Performance assessment tasks for Electricity unit

Performance Assessment Task 1 from the Electricity unit (see Figure 3.1 above and Appendix IIA) was intended to be a simple task that also introduced the new assessment format to the students in a non-threatening way. The task was on the electrical conductivity of a set of six objects marked P, Q, R, S, T and U (iron nail, aluminium foil, carbon rod, wooden-rod, PVC-object, and piece of porcelain). Students were required to test the objects (materials) for their electrical conductivity following their plan and record their observations. The follow-up questions needed students first, to classify, with justification, the materials into conductors and non-conductors of electricity. Second, they also had to identify, with reasons, the most suitable materials for use in replacing electrical insulation material or conducting materials as presented in a picture of an electric pole.

In Performance Assessment Task 2 students were given a set of four different wires of one metre-length each: thick nichrome wire (0.50mm), thin nichrome wire (0.03mm), thin constantan wire (0.02mm) and thin copper wire (0.05mm)). In carrying out the task, students were required to measure the electric current passing through, as well as the voltage across each of the wires. They then had to calculate the electrical resistance of each wire. Again they planned how they would connect the circuit components and the meters to be able to collect the data needed. Real world context questions were on the suitability of the wires for a given function. Electric cells were to be given on approval of the circuits assembled in the groups, to ensure proper connection and protection of the meters. Performance Assessment Task 2 and its rubrics are provided in Appendix IIA.

d) Performance assessment task for the Air and Living Things unit

The performance task given for Air and Living Things (see Appendix IIB for task and rubrics) assessed students' knowledge of properties of oxygen and carbon dioxide gases and their skills in testing for these gases. Their ability to make observations and record data were also tested. Students were given three colourless gases labelled Gas A, Gas B and Gas C, to determine the identity of the gases. The names of the gases were given to the students to reduce the number of tests students might want to carry out. Students also decided how they would test for each of the gases and what materials they would use to test the gases.

e) Rubrics

Performance assessment tasks do not necessarily have clear-cut right or wrong answers. They elicit a number of acceptable responses from respondents and thus may have multiple solutions. Their scoring is defined in terms of degrees of successful implementation of a given performance task. Students' success in a performance assessment task is normally scored by using rubrics - specially prepared scoring guides (Brualdi 1998).

Rubrics can be in the form of rating scales, checklist of criteria presented as specific actions, holistic overall performance of students or anecdotal reports on student performance, as discussed in Section 2.4.2. Performances on tasks are scored by giving an estimated numerical value to the performances that can range from excellent to unacceptable and presence or absence of specific actions (Borich and Tombari 1997; Brualdi 1998; Downing (not dated)). The ideal way of scoring performance tasks is to directly observe and score students' actions while they carry out the tasks (Shavelson and Baxter 1992).

The suggestions noted above were observed when constructing rubrics for scoring each performance assessment task. Thus the following guidelines were followed when developing and implementing rubrics for the performance tasks. They have been adapted from Airasian (1991 cited in Brualdi 1998) and Linn and Gronlund (2000).

1. Identify the overall performance task and perform or imagine doing it.
2. List the important aspects of performance or product that would result from performing the task.
3. Express the performance criteria in terms of observable student behaviours avoiding ambiguous words.
4. Limit the number of performance criteria and levels of proficiency to a number that can be observed during the performance of the task.
5. Arrange the performance criteria in the order in which the behaviours are likely to be observed.
6. Get peers to review your stated learning outcomes, performance tasks and rubrics to see if they are feasible and that they match, to produce a more valid and possible reliable scoring guide.

7. When implementing the rubrics, rate all students on the same task before going on to the next task, to improve consistency in scoring and minimise variations.

Constructing rubrics took place at the same time as the construction of each task. After construction each performance assessment task was tried out to ensure that the instructions were clear and that the tasks would lead to meaningful and usable data. Actions and results of those actions were identified and used to create the scoring criteria that matched each of the stages of the assessment task. Criteria that were considered important were selected, modified and assigned a numerical score each. The criteria were then sequenced in the order in which they were anticipated to occur according to the task instructions. This was done to aid scoring. The sequenced criteria and their scores comprised the rubrics for scoring each performance assessment task. An example of a rubric is shown in Figure 3.2 below. Scores for questions dealing with the use of data from Stage 2 of the task were indicated alongside the questions themselves, as shown in Figure 3.1 above.

Thus, each rubric comprised sections and scores corresponding to the different stages of the performance assessment tasks; namely planning, investigating (manipulating equipment and capturing of data), analysis and interpretation of data.   An example of rubrics for scoring a performance assessment task is shown in Figure 3.2 below.

i) Rubrics for scoring Electricity performance tasks

The rubrics given in Figure 3.2 below reflect criteria and the ratings against which performances in the tasks were scored. This rubric was used in scoring Performance Assessment Task 1.The criteria also show the degree to which students correctly planned and used the plans to assemble the circuit, correctly test the given objects and record their observations. The criteria for scoring the recording of the data included a table or other form of recording that would show, for example the material tested, the expected observations, in terms of whether the light bulb produces light and indicate the brightness of the light. The scoring of students' analysis, interpretation and use of the data were guided by questions given in the procedure and report sheet.

Electricity Performance Assessment Task 1 Scoring Guide
School ………………………………. **Date……………….**
Rating scales used to guide grading of performances. They will be used to guide the grading of each group of students working on a given task; each member gets the same mark. Students discuss their ideas and can consult the teacher should they need to refer to a diagram.
NB: Please outline to them what you assess.
Activity 1: Please rate the performance of the students / group according to marks given in **bold.**

Assessment Aspects	**Assessed ability**	**Mark / Ratings**	
		Group No. 1	**Group No. 6**
A: Planning of testing of materials	Correctness of diagram of circuit to be used to test materials **[0-4]**		
Plans show	Data recording table with appropriate headings e.g. (material (P,Q….); light bulb, brightness of light, conduction … **[0-4]**		
B: Manipulation of equipment	Connects circuit that matches diagram **[0-2]**		
	Tests materials correctly **[0-2]**		
C: Capturing data	Makes accurate observations **[0-2]**		
	Records data correctly **[0-2]**		
D: Analysis and Interpretation of data	Question in paper **[12 marks]**		
E: Co-operativeness	Group works co-operatively with one another (no one student dominates) **[0-2]**		
Total	**30 marks**		

Comments:

Figure 3.2 Example of rubrics for scoring a performance assessment task

The scoring guide for the second performance assessment task on the electrical resistance of pieces of wires indicated degrees of correctness of circuit (both drawn and assembled), presence and positions of the circuit components, an indication of variables to be recorded as data, measurements and calculations (see Appendix IIA).

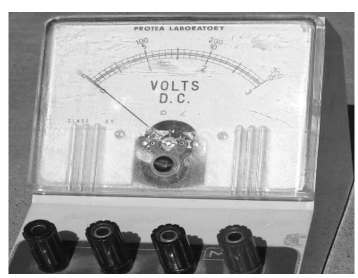
ii) Rubrics for scoring Air and Living Things performance task

Scoring criteria were presented in four steps with the scores for rating the performance of the students ranging from a minimum of 0 to a maximum of 5. In the plan for testing gases steps to follow when testing the gases and the outcome of the test are shown. Criteria focusing on performance in the gas testing stage of the task dealt with the correctness of procedure for testing gases, as well as the recording of observations and correct identification of the gases. Identifying gases used in given situations was part of the question section (see Appendix IIB).

3.5.2 Context-based unit tests

Context based tests were constructed for each of the units on Electricity and Air and Living Things. Each test item comprised real world contexts as descriptions and/or pictures, as well as questions based on the scientific concepts relevant to the contexts. The tests are provided in Appendix IIIA and Appendix IIIB respectively. The structuring of the context-based questions was influenced by suggestions from Ahmed and Pollitt (2001). One such suggestion was to focus the context in the question. They define “focus” as the “extent to which the most salient aspects of the context correspond to the main issues addressed in the question” (Ahmed and Pollitt 2001:5). In questions that are designed to fit naturally into a real world context, the scientific concept that is central in the question is also central in the context. They also advise against the use of real world contexts that may be interesting, but have no significant role in the answering of the question. In observing these suggestions all contexts used for the questions had direct links to the scientific concepts sought. Figure 3.3 below shows an example of a context-based question taken from the Electricity unit test.

2. Londiwe went to the shop to buy a PM 9 battery (Figure 1 below) for her radio. She asked the shopkeeper to check whether the battery was “full” or not.
   Draw on Figure 1 below how the shopkeeper would test the battery. (2)



3V 15V 300V +

PM 9 Battery Voltmeter

Figure 1: A voltmeter and a PM 9 battery

Figure 3.3 Example of a context-based question

3.5.3 Student questionnaires and interview schedules

In order to ascertain students' views on the alternative assessment approaches used in this study, interview schedules and questionnaires were developed. For the context-based tests the questionnaire items sought information about students' likes, dislikes, and feelings about the use of contextualised items as test questions. This questionnaire was attached to each topic test as shown in Appendices IIIA and IIIB.

Two sets of semi-structured student interview schedules were developed to solicit views of participating students on the use of performance assessment tasks. The first interview schedule was administered to students of School M4 and, because of problems discussed in Section 3.11.2 below, it was modified into a questionnaire to be administered to individual students. The second interview schedule was also used as a questionnaire. These instruments sought information about what students liked or disliked, or thought was good or not good about performance assessment tasks, working on tasks in groups, and the use of different models of assessment in one topic. The second interview schedule/questionnaire also included questions requiring the students to say which stage of performance assessment task students liked and to establish aspects of their readiness to take performance assessment tasks. The interview schedules/questionnaires are included in Appendix VIIA.

3.5.4 Teacher interview schedules

A semi-structured interview schedule for a teachers' focused group interview was developed. The interview was intended to draw out teachers' ideas about perceived good and poor aspects of the two assessment models used, their perceived effects on students' learning and on teaching, requirements by teachers and schools to enable their use of these assessment approaches, as well as their views regarding the future use of the assessment models.

Focused group interviews were used to provide greater understanding of issues surrounding the use of performance and context-based models of assessment from the teachers. Schumacher and McMillan (1993) encourage the use of group interview by stating that a social environment is created by a focused group interview where group members are stimulated by perceptions and ideas from other group members.

On review of the transcript from the focus group interview it became clear that views on context-based assessment were not adequately discussed. Another interview schedule was then developed which focused on teachers' views on good and poor aspects of context-based questions, teacher observed students' problems and strategies they used in dealing with such problems and students' answers when marking, challenges teachers experienced regarding context-based tests, future use and main concerns about context-based assessment. These instruments are provided in Appendix VIIB.

3.6 VALIDITY AND RELIABILITY

According to Gay and Airasian (2003) validity is concerned with the appropriateness of interpretations and inferences researchers make regarding test scores or responses in questionnaires, interviews, or observations records made during the course of a study. Reliability is concerned with the consistency with which an instrument measures whatever it is measuring. In terms of tests this would refer to the consistency of scores awarded to an individual in that test by different scores or at different times (Gay and Airasian 2003). In qualitative research this would be the extent to which independent researchers discover the same phenomena, and to which researcher and participants agree on the descriptions of the phenomena (Schumacher and McMillan 1993).

Instruments used in this study were of two types, namely assessment instruments and data collection instruments. To improve the degree to which these instruments allowed for appropriate interpretations about the respondents perceptions, the procedures described below were used. The procedures for validating the assessment instruments, namely performance assessment tasks and the context-based tests are described first.

3.6.1 Peer review

The process of constructing the assessment instruments was followed by a review that was conducted by four colleagues at the University of Swaziland (UNISWA). The two electricity performance tasks were also reviewed by a scholar from the University of York who has vast experience in research involving context-based teaching and learning. He was also involved in the development of the Matsapha Lessons. Three of the five reviewers were very familiar with the context-based teaching approach having been involved in the development of contextualised teaching materials and research surrounding their use. Each reviewer was given the abstract of the study, the syllabus objectives and syllabus content for each unit, the learning outcomes, the assessment instruments, as well as the questionnaires. They were asked to check for the appropriateness of the learning outcomes, the content validity of the tasks in measuring the specified learning outcomes, the suitability of the contexts, as well as the wording of the items. Amendments were made to the assessment tasks as discussed and suggested by the reviewers. Modifications were mainly on simplifying the language used in the phrasing of questions. Some examples of major changes were:

- making contexts more realistic;
- contextualising some questions; and
- adding pictures or diagrams to make questions or context clearer (crops growing under trees in Question 2 in Air and Living Things test; Question 3 in Electricity test; picture of electric power pole in Performance Assessment Task 1).

Validation of performance assessment tasks and context-based paper-and-pencil unit tests was viewed in terms content validity. Content validity demands that the test items represent a degree of measurement of the intended content area (item validity) and adequate coverage of the intended content area (sampling validity) **(**Gay and Airasian 2003). The syllabus information and the learning outcomes provided the target content coverage against which the assessment instruments were to be validated. Content validity of the tests was considered collectively for the performance task(s) and test in each unit. That is, content coverage was considered for all tests conducted for each unit such that they complemented each other in providing a wider coverage of content for each unit. Performance assessment tasks focused on only a small section of a unit or topic, which necessitated a collective judgment of content validity for the assessment forms used.

Validity in this study was considered in terms of credibility and translation fidelity as suggested by Krathwohl (1998), and are discussed below.

3.6.2 Credibility

According to Krathwohl (1998) credibility entails judging the integrity of the researcher in carrying out the study. It involves both the steps taken to ensure the data is credible and being forthright about the strengths and weaknesses of the study and its findings. In ensuring credibility of the study Krathwohl advises that the study should build credibility with readers by utilizing studies that are already known and accepted and avoid the weaknesses of those studies. In this study suggestions from the literature were used when constructing the tasks. For example the construction of performance tasks benefited from Solano-Flores *et al.* (1999) who propose a framework for constructing and administering performance assessment tasks, as well as Linn and Gronlund (2000) and Borich and Tombari (1997) who give guidelines on maximizing the validity and reliability while constructing performance tasks and their rubrics (shown below). The construction of context-based tests benefited from the experiences and suggestions by Ahmed and Pollitt (2000, 2001).

Credibility in the study was also enhanced by triangulation, a cross-validation procedure of comparing inferences made on participant consistency in responses to different questionnaire items and interview, as advised by Gay and Airasian 2003).

3.6.3 Translation fidelity

According to Krathwohl (1998) translation fidelity involves judging whether the meanings of the explanation, rationale or theory presented as findings are congruent with the forms of behaviour represented by the exploration. It judges whether the concepts and constructs used as findings are a good match for the excerpts of data selected as supporting evidence for the findings. In this study perceptions represented by themes and categories derived by an inductive process from respondents' statements are supported by samples of excerpts from which the perceptions were derived. These excerpts enable readers to make their own judgment regarding congruence between the generated concepts and constructs and the forms of behaviour they are explaining.

3.6.4 Reliability

Reliability also concerns the accuracy with which a test measures the skills or attainment it is designed to measure. This is the consistency with which students are scored in tests (Gipps 1994). The tests used in this study comprised both closed and open ended items. For some questions the answers were specific and predetermined. In other questions the students used their own experiences and ways of expressing their answers. To improve the reliability of the scores given to the students in the performance assessment tasks the guidelines described in Section 3.5.1(e) above were followed.

Performance assessment tasks and the context-based tests were intended to provide students with experiences of being tested through such assessment procedures. The tasks were used to familiarise participating students and teachers with the two approaches to assessing students' learning so that their views about these models of assessment could be explored through other research instruments: the interviews and questionnaires. As there were no readily available research instruments to explore students' views and perceptions about assessment, other ideas were sought from studies that explored students' and teachers' perceptions, to inform the construction of the instruments. These studies are discussed in Chapter Two Section 2.6.

3.7 PILOTING OF INSTRUMENTS

As mentioned above, the performance tasks and the context-based topic tests and questionnaires were reviewed by peers at UNISWA. After incorporating comments and suggestions from the reviewers the instruments were piloted.

3.7.1 Piloting the Electricity tasks and tests

Performance assessment tasks from the Electricity unit were piloted with one Form III (Grade10) class from School M2 in March 2004 with the help of three participating teachers. The scoring rubrics were also piloted during this time. Discussions were held with the teachers about the structure of the performance tasks and how they were to use the rubrics to rate the performances of students on the tasks. Two stations were set up for each task and the students worked on the tasks in six groups of five to six students. On completion of one task the students disassembled their circuits and moved on to another station. On completion of the performance assessment tasks by students, observations made by teachers on the use of the tasks and scoring rubrics were discussed. Suggestions from this discussion were incorporated and amendments were made to the instruments to get them ready for implementation. The structure of the scoring rubrics was modified to provide space for indicating actual scores for each group rather than tick the score from a list of scores of 0 to 5. Criteria were restructured to show their targeted actions clearly. Other changes are mentioned below.

Modifications that were made on the performance assessment tasks after piloting were as follows:

i) Provide space for circuit diagram for the plan just below procedure 1 (Task 1 and Task 2).

ii) Highlight (bold) requirement for showing how results would be recorded.

iii) In Task 1 a question requiring students to justify their classification of electrical conductivity of the materials tested. This was intended to get them to think deeper about the classification and therefore also improve their answers to the rest of the questions. Another additional question was given (uses of conducting material) to provide a contrast on the uses of (non-conducting) materials given in the situation.

Teachers also received training in administering and scoring performance assessment tasks during this piloting exercise.

Piloting of the unit topic test was conducted with one Form III class in a non-participating school. The piloting group had not learned the unit in a contextualised way as this approach was not normally used in schools and no school that used the contextualised science teaching approach and Matsapha Lessons could be found. Students also completed a questionnaire at the end of the test. Reviewing the responses to the questions indicated that some modifications had to be made to include some less demanding questions or leading questions. The following changes were made to this test:

1. A new Question 2 on showing how to connect battery to voltmeter was added.

2. Question 2 became Question 3 and modifications to this question included a question on sources of electrical energy and the bulbs were reduced from four to two in Question 3(e).

3. Questions on energy changes and representation of the units V and W in electricity were added in Question 4.

3.7.2 Piloting Air and Living Things tasks and tests

Two performance assessment tasks (although only one was eventually used due to time constraints) were piloted with two Form III classes in February 2005, one task in each class. Again students worked on the task in groups of six. Only two teachers were able to attend this piloting session. The scoring rubrics for the tasks were also piloted with the performance tasks. At the end of the administration of the tasks observations by teachers were discussed again. The following modifications were made on the task.

1. A third test tube of gas (air) to be tested was made available.
2. The gas testing materials were not listed and students were required to ask for them.
3. Students were asked to predict and explain which of three plants grown in each of the three gases would produce most starch instead of describing a test to confirm their identification of the gases.
4. A fourth question on rusting was also added.

The topic test on Air and Living Things and its questionnaire were piloted with a Form III class, also in February 2005. Modifications were made in light of the students' responses to the test items and questionnaires, and were as follows:

1. Question 2(b) was added to extend the knowledge assessed to practical activities.

2. The wording in Question 3 was modified to be clearer. For example from a phrase like "how does the body benefit from the changes…" to "how do the changes … help her ...?"

3. Two questions on other changes in the human body during exercise and on the possible effect of exercise on people with heart problems were removed. These

were replaced by a question on the advantages of an increased heart rate during exercise.

4. On the question about the brazier one sub-question on what the siblings wanted to do with the brazier was deleted. One question was modified to demand two reasons rather than one because the response produced from the piloting of the test was linked closer to informal knowledge about the possible dangers of a brazier. It was not clear whether the omission of the scientific answer was due to question demands or other factors. This change was intended to get students to go beyond informal answers and to provide more "scientific" answers.
5. Question 5 on painting roofs was split into two parts so that students would be guided towards responding correctly to the question as they had omitted some crucial information in the piloting of the test.
6. The last two questions were exchanged.

3.8 ORIENTATION OF TEACHERS

Teachers attended two orientation workshops of two days each as part of their familiarisation with the context-based teaching approach.

Teachers were formally invited to the workshops through hand delivered invitation letters. During the delivery of the letter each participating teacher was given separate copies of the Electricity and the Air and Living Things Teachers' Guide to read and familiarise themselves with the content, structure and teaching approach followed in these curriculum materials. They were also given the workshop programme indicating what the focus of the workshop would be (see Appendix VI). The first workshop was held in March 2004 and the second workshop was held during the first week of the school holidays in April 2004. Day one of the first workshop and the two days of the second workshop were held at the Department of In-service Education Laboratories at the University of Swaziland. The second day of the first workshop was held at School M2 where the Electricity performance assessment tasks were piloted, as mentioned above.

3.8.1 Workshop 1

The first day of Workshop 1 was dedicated to familiarising the one new teacher with the contextualised teaching approach to be followed for the study, and to refresh the memory of the three other teachers who were familiar with the approach. This was to make sure the participants had a similar understanding of how to conduct the lessons. Electricity was the first unit to be discussed since it was the next topic in the teaching sequence of the Swaziland Integrated Science programme (SWISP) curriculum. During the workshop teachers compared the activities in the Electricity unit with those of the same unit in SWISP to ascertain the similarities and differences between the two teaching approaches. According to Dlamini (2000) comparing old and new teaching approaches is an effective way of familiarising teachers with the teaching material. Comparisons were in terms of the nature and degree of guidance given to students, the nature of the learning activities and teaching approaches. The comparison was made to show that the materials used for the study also followed the same syllabus as that used for SWISP. The teachers also discussed a few lessons that were representative of the context-based teaching approach and discussed the different sections of the lessons (the lesson planner, teacher notes, students' activities) and what was expected of the teachers when implementing the lessons. The teachers then tried out the students' activities planned for in the lessons and discussed their experiences in the process.

The second day was dedicated to discussing and administering the pilot test for three performance assessment tasks on electricity and using the scoring rubrics to observe and assess students' practical skills in electricity as described above. After the discussion of the tasks and the rubrics the setting up of equipment for the tasks was carried out. The teacher from the host school then invited the students and divided them into the six groups according to the stations set. After the introduction of the visitors to the students by their teacher the tasks were introduced to the students. It was also explained to the students what was expected from them when completing the tasks. Thereafter the students started working on the tasks in groups ranging from five to six students per group. While the students worked on the tasks the teachers checked their work. The teachers also assisted and redirected students whenever they needed help. Throughout the exercise of using the scoring guide consultations were held with the teachers to clarify procedures in scoring and to ensure that the teachers understood the tasks and the rubrics and how to use them. Things that needed attention in both the tasks and their rubrics were noted. Students' input through the questions they asked during their interaction with the teachers was noted and used to modify and improve the tasks and scoring guides, as stated above.

3.8.2 Workshop 2

The unit on Air and Living Things was discussed. The workshop was run in April 2004. This was long before the units were to be taught, from mid-second term (June to August). This was done on the recommendation from the teachers. Teachers wanted the workshops earlier so that there would be no further disruptions of their teaching caused by attending workshops. They also wanted greater exposure to the teaching approach before they actually started using it. During the two days of Workshop 2 the teachers briefly recapped on the characteristics of the materials, but spent most of the time conducting the students' activities and discussing their instructional requirements.

3.9 ETHICAL CONSIDERATIONS

Ethics of research deal with protecting participants from exploitation and deception by a researcher, as well as from physical, mental and emotional discomfort form the research work (Schumacher and McMillan 1993). It also considers issues of consent, confidentiality and respect for participants (Robson 2002). Ethical issues considered in this study are described below.

This study was explained to all the participating teachers and the administration authorities in the schools. At the time of seeking permission from schools to conduct the study, the purpose and aims of the project were discussed with head teachers and heads of departments in each school. The nature of the data, the data collection process and how the data would be used were explained. Schools were assured of confidentiality when reporting the findings. Heads of departments also inspected the teaching materials and were satisfied that their use would not interfere with the science content in SWISP. Participating students received the same science content treatment as others following SWISP. Only the teaching and assessment approaches were different.

Participating students were aware of their involvement in the study. Although they would not be able to withdraw individually from being taught in the contextualised approach they could withdraw from the tests (performance assessment tasks and unit tests) if they so wished. Students were informed that participation in the interviews and responding to the questionnaires was voluntary. The withdrawal of the entire school was also possible as it happened with the fifth teacher who left. This point is discussed further under Section 3.13 below on methodological challenges.

Participants' contributions were treated confidentially during the analysis and reporting of these contributions. Further confidentiality was observed by not indicating personal details of the group of students or the school in which each teacher taught. Characteristics of the schools that can easily lead to the identification of the school were also omitted from the description of the schools. Codenames were used for schools and pseudonyms for the teachers.

3.10 EXPERIENCING ALTERNATIVE ASSESSMENT MODELS

As mentioned in the aim of this study, the focus was to establish and describe views of participating students and teachers regarding the use of performance assessment tasks and context-based tests. At the time of this study the contextualised approach in teaching Science, as well as the performance assessment models were not commonly practiced in secondary schools in Swaziland, although context-based questions were included in tests and examinations on many occasions. This study, as noted above in Section 3.5 needed to first provide the students and the teachers with the necessary experience of the assessment models prior to exploring their views. Below follows a description of assessment activities that were used to generate students' and teachers' experiences of using the alternative assessment models.

At the beginning of the study three hundred and ten Form II students participated. Each student was supplied with a students' workbook that followed a contextualised teaching approach. The teachers were provided with a corresponding Teachers Guide. Each workbook focused on four units: Electricity; Air and Living Things; Force, Support and Movement; and Acids, Alkalis and Gases. The last two units were not used during this study due to time pressures in schools as students were already working towards getting ready for the external examinations at the end of Form III. The final number of participating students in Form III was two hundred and forty-nine. According to the class registers from the teachers three of the fifty-one students who did not proceed to Form III left school, implying that forty-eight students repeated Form II. This is an average of seven students per class.

Providing opportunities for participants to acquire experience of the assessment models and the data collection exercise took place over a period of about twenty months (March 2004 to October 2005). Data collection was adjusted according to the operating programmes in the schools to keep disruption of school routines to a minimum.

Winegardner (not dated) says that it is important to keep at minimum interruptions to the operating system in a situation under study. The Ministry of Education in Swaziland normally, and correctly so, demands that.

3.10.1 Administering performance assessment tasks

Each teacher was provided with all performance assessment tasks and rubrics. They were also assisted when administering the tasks to the students. All the equipment and material required for the tasks were provided to ensure that all that was needed was available. It was convenient to provide the material to ensure that there was a suitable level of awareness and control of their working condition and availability for all the classes. Equipment and materials that were used for conducting the performance assessment tasks were borrowed from the Department of In-service Education of the University of Swaziland (UNISWA).

On the day that each school arranged for the performance assessment tasks, the task, its implementation, as well as the use of the rubrics were discussed with each teacher. Then the equipment for the task was set up to allow students to work in groups. The provided equipment proved insufficient to allow students to work individually. There was also insufficient time to have had the students to perform the tasks in shifts so as to work individually.

After setting up a particular task the teacher invited the students into the laboratory. Once the students were settled the teacher introduced the task and then informed the students of what was expected, what would be scored and that they could ask for help if they needed to. Thus, students were helped regarding problems of the equipment, as well as in respect of the subject content. Assistance relating to subject content was provided by the teacher in order to avoid any variation in the nature and extent of content related assistance that might interfere with student performances. Technical assistance was also provided when the equipment malfunctioned. At the end of each performance assessment task the teacher's observations and experiences of his or her implementation of the task were discussed. These discussions were audio-recorded.

Teachers observed and scored the performances and display of specific skills by students on each task. Scoring was by means of ratings presented as numerical values in the rubrics (see Figure 3.2 and Appendix IIA). The teachers made their own judgments on how to rate the students' performances on the assessment tasks.

Students from School M4 were the first to carry out Performance Assessment Task 2 from the Electricity unit and they experienced problems. They took a long time to come up with the appropriate circuit. They needed a lot of assistance with the circuits and to get them to work. They also had problems reading the instruments (voltmeter and ammeter) to obtain the readings needed to calculate the resistance of the wires. Therefore, they needed more time to complete the task. From that experience, the procedure of administering the task was adjusted by following Solano-Flores *et al.*'s (1999) approach of giving students ready to use data after Stage 2 of the performance assessment task. Thus, data on the electric current and corresponding voltage for each of the wires was produced with the help of the teacher from School M4. This data was given out to students when the task was administered in the other schools. Students were given this set of data after they had obtained two or three readings, but they were not informed of such data prior to the time the data were issued. Providing students with the data was a way of facilitating completion of the task and ensuring that students used meaningful data to complete the task. In addition, the view that assessment should support learning implies that data students obtain and use through performance assessment tasks should be scientifically acceptable.

3.10.2 Administering context-based unit tests

Both the Electricity and Air and Living Things unit tests were administered and scored by each participating teacher after the unit had been taught.

Students also completed a questionnaire at the end of the test. All test scripts and questionnaires were collected from the teachers.

3.11 DATA COLLECTION

Data collection involved participant observations, student questionnaires and interviews, teacher interviews and discussions with teachers.

3.11.1 Participant observation

During the administration of the performance assessment tasks participant observations were made as students executed the tasks and how the teacher interacted with them. Field notes of occurrences in the groups such as students' efforts in planning, drawing and assembling of electric circuits, procedures they used, as well as the nature of answers they gave for task sub-question were noted. Student behaviour in their groups such as participation and their interactions with each other were also noted. These observations were carried out while students were being assisted with testing circuits, changing connecting wires, replenishing escaped gases, issuing material for testing gases and any other technical assistance they required. As a result some field notes were written out as reflection notes after the activity.

3.11.2 Administering student questionnaires and interviews

Students' views were obtained by means of student questionnaires and group interviews. These data collection procedures were conducted at different periods of time as dictated by the progress of teaching and the implementation of the performance assessment tasks, as well as each school's programme. Each teacher made all the time arrangements with the students for the administration of the interviews and questionnaires.

The first session of the interviews involved students in School M4 who were interviewed in the groups set out for the performance tasks. All six groups were interviewed over two days in August 2004. Interviews were conducted in the last week of Term II when the students were not attending regular classes. The first day of the interview coincided with a movie show so only two groups were interviewed. The other four groups were interviewed on the second day. Although the interview lasted less than 20 minutes with each group, it seemed the interviewing process took a long time for the students to wait for their turn. Some group members did not turn up for the interview. Finding suitable time slots to conduct similar interviews with other classes was difficult. Thus, after the initial interview session with students from M4 the interview schedule was modified into an open ended student questionnaire that was administered by the participating teachers in the remaining six classes. This questionnaire is found in Appendix VIIA. Submissions to this questionnaire and interviews were not used in reporting the results. The reason was that it was difficult to control the administration of the instruments in the different schools as all groups in School M4 were interviewed but not in the other schools. Experiences from the first set of data collection helped modify the organisation and implementation of the group interviews and individual questionnaires.

During the administering of the second set of the interviews and questionnaires, students were arranged into groups in which they would discuss and respond to the questionnaire. After the students had settled down in their groups the questionnaire was read to them. The questions were explained to the students to help them understand questions and the response each item required from them. The introduction was followed by the selection of one group for the interview with the help of the teacher. The selected group was isolated from the rest of the groups and interviewed through a focus group interview. Each group discussed the questionnaire and responded to it as a group. While students in groups were completing the questionnaire the students in the other group were interviewed. The groups working on the questionnaires were supervised by their teachers. All interviews were audio recorded as voice files through a digital device. A total of forty-one group questionnaires and eight interviews were used in the data analysis and reporting of the findings.

Students in each interviewed group were allowed to discuss their responses to the items before the interview and audio recording commenced. During this time the progress of the other groups was checked.

These interview sessions went as follows:

- M3: two days in March 2005. Four groups from two classes were interviewed during class time.
- M2: March 2005. Two groups from two classes were interviewed during class time.
- M4: July 2005. One group from one class was interviewed during study time.
- M1: August 2005. One group was interviewed during negotiated non- lessons.

The delay in conducting interviews with M1 and M4 was due to the difficulties of finding suitable time for administering questionnaires and conducting the interviews. In the case of M4 a teaching period for English was combined with a study period and used for the interview. In the case of M1 the interview was held in the last week of the term after the students had written their mid-year internal examinations.

The purpose of the interviews was to get students to express their ideas about the assessment approaches used to appraise their learning. Sometimes they would take time to respond or seem to have difficulty expressing themselves in English. So they were allowed to express their views in the local language, siSwati, so that they could be free to say what they wanted to say. SiSwati was also used to clarify some questions or encourage students to respond. Translation of sections of interviews that were in siSwati was done during the transcribing process and incorporated into the interview transcripts.

3.11.3 Administering teacher interviews

Both formal and informal discussions were held with the participating teachers. Informal discussion occurred during field visits to check on progress, or soon after the administration of performance assessment tasks. Questions for initiating discussion with teachers were open ended to allow the teachers to talk about their experiences and observations when they taught and assessed their students through the tasks. The teachers also described how they went about the scoring of the students while they worked on the performance tasks. These discussions were also useful in explaining aspects of the tasks or their scoring that were not clear to the teachers. Discussions were audio recorded as voice files.

In formal discussions a semi-structured interview schedules was used. A focus group interview was held in May 2005 with three teachers. Teachers were asked to prepare for this interview through an outline of questions that were to be discussed during the interview (see Appendix VIIB). The fourth teacher was asked to write out her responses to the questions in the interview guide after two attempts to get an interview failed. After looking through the responses, the need to interview the teacher was evident and she agreed to the interview. Another interview that focused on views on context-based tests and questions was held with each teacher. The interviews were audio recorded digitally as voice files. The focus group interview was also video recorded.

3.12 ANSWERING OF RESEARCH QUESTIONS

This study pursued two research questions.

Question 1 was:

> What perceptions and experiences do students and teachers have about the use of performance assessment as alternative strategies for determining attainment of learning outcomes in Science?

Information to answer this question was obtained from student group questionnaire data and student interview transcripts, as well as from teacher interview transcripts. Specific questions were used to get students and teachers to describe their feelings and experiences of good and unfavourable aspects of performance assessment tasks and their mode of administration. Field notes of observations during the administration of performance assessment tasks were also used for triangulation.

Question 2 was:

> How do students and teachers view the use of context-based assessment in assessing learning in Science?

Students' views about context-based tests were obtained through questionnaires. These questionnaires were attached to the test. Students completed the questionnaire immediately after they had done the test. Transcripts of students' interviews were also used to provide additional data. Teacher interview transcripts provided information on their views regarding the use of context-based questions and tests.

3.13 DATA ANALYSIS

Data for the study was obtained from student questionnaire responses, student interview transcripts, teacher interview transcripts, as well as observation notes. These responses were converted into computer text format in preparation for coding and analysis. The constant comparative method of data analysis was used (see Glaser and Strauss 1967; Merriam 1998).

3.13.1 Coding and analysis

Organisation and management of responses from questionnaires and interviews and their coding were facilitated by the ATLAS.ti 4.1 computer programme for qualitative data analysis. The programme allowed coding of responses and the manipulation of data (codes) in various ways, grouping the data and generating frequency tables, all of which enabled the interpretation of the data. The coding process involved several steps as described below.

Through the programme meaningful parts of response statements or data segments were selected and coded using codes derived from respondents' voices or words. Initial codes were brief notations (words or short phrases) that captured the essence of the idea presented in the selected segment of a student's response. Units of responses within the same text file were compared for similarities with those already coded so that similar responses were given the same code. This procedure was followed for each questionnaire question and each class. Thus responses for each questionnaire question and each class were coded and saved as a computer file referred to as a hermeneutic file.

Following the initial coding stage, all the hermeneutic files (coded responses files) for each question from the four schools were merged into a single hermeneutic file. Codes and student phrases from the different schools were compared and checked so that codes dealing with similar ideas were merged into one code or assigned a new code.

The process of merging codes also involved revisiting of data segments and further refining and redefining of codes. A second merging of hermeneutic files was conducted for certain questions in order to obtain corresponding or conflicting entries. For example Question 1 on what students thought was good about performance tasks was merged with Question 3 on what students thought was not good about performance tasks. Codes for similar ideas were grouped together into categories or "code families" to use ATLAS.ti language. The merging of hermeneutic files for different questions allowed for the identification of similar ideas that were expressed in different questions. Themes were sought for the categories from literature and from the categories themselves. The themes and categories were used to construct a code book that was used for an inter-coder reliability check. These themes and categories were used in framing, reporting, as well as discussing the findings of this study and they are outlined below:

1. Theme: Affective disposition
   - Affective disposition motivation extrinsic (ADME)
   - Affective disposition motivation intrinsic (ADMI)
   - Affective disposition self-efficacy (ADSE)
2. Theme: Metacognitive disposition
   - Metacognition (MET)
3. Theme: Social relations
   - Peer collaboration and support (PCS)
   - Group assessment problems (GAPRO)
4. Theme: Psychomotor Disposition
   - Handling apparatus and procedures (PSYCHAN)
5. Theme: Recommendations (RECOM)
6. Theme: Resources
   - Resources for performance assessment tasks (PATRES)
7. Theme: Stages liked/disliked (ST)
8. Theme: Degree of readiness for task (DRT)
9. Theme Task characteristics:
   - Task complexity (TCC)
   - Task importance/value (TCIV)
   - Task requirements (TCREQ)

Task format and presentation (TCFP)

Although the themes are listed alphabetically their discussion does not follow this sequence.

3.13.2 Inter-coder reliability

Inter-coder reliability was determined through the help of three colleagues from the University of Swaziland. Each coder was provided with a codebook and the students' responses to the questionnaires. For the group questionnaire a sample of responses from eight groups (about 20%) comprising one or two groups per class was provided to each coder. After the coding exercise codes were checked. Where errors were suspected in the codes a discussion of those codes was held with the coders and the responses were returned to the respective coders for a further check of codes. An example of an instance where a coding error was suspected was the coding of a statement on the non-participation of students as "PCS (peer collaboration and support)" instead of "GAPRO (group assessment problems)". All codes for individual coders were compiled. Inter-coder reliability was calculated in relation to each of the three coders per question and in relation to all the coders per question, as well as for the entire questionnaire. A sample of how inter-coder reliability was calculated is provided in Appendix VIII Inter-coder reliability ranged from 58% for Coder 3 to 78% for Coder 2 with an overall inter-coder reliability of 70% for views about performance assessment.

For each of the questionnaires from the two context-based tests, responses from three to six students per class, depending on the size of class, were randomly selected for an inter-coder reliability check. Each coder received responses from 60 students. The 60 responses comprised 34 students' questionnaire responses from the Electricity test and responses from 26 students from the Air and Living Things tests. This constituted about 10% for each test questionnaire. Agreeing and disagreeing codes were compared against each coder. Inter-coder reliability ranged from 63% for Coder 3 to 71% for Coder 1 with an overall inter-coder reliability of 66% for the students views about context-based tests.

3.14 METHODOLOGICAL CHALLENGES

Working with participants in their everyday environmental setting and adhering to ethical considerations had its challenges. Some of the challenges that might have had a negative effect on the outcomes of the study are described below.

3.14.1 Attendance of induction workshops and piloting of tests

One teacher did not attend the second induction workshop for context-based teaching and familiarization with the teaching material. The first piloting of performance assessment tasks and scoring rubrics were also missed by one teacher while two teachers missed the second piloting. Teachers who missed any of the piloting sessions also missed out on valuable training for implementing the tasks and rubrics which could have influenced their perceptions. Nevertheless, although Inna missed both sessions of piloting the performance assessment tasks, she managed to use the rubrics quite well, even making valuable suggestions about allocating scores for some performances.

3.14.2 Communication

Problems of communication with participating teachers were experienced. Such problems were anticipated, so that at the beginning of the study each teacher received an allowance to purchase mobile phone airtime (they all had mobile phones) to facilitate their communication regarding their needs or progress. The effects of low communication caused delays in the administration of the performance assessment tasks. The ideal time to administer these tasks was soon after the lessons dealing with the content and the skills assessed. This time was ideal because students could easily remember the information and also be able to consolidate the learning of those concepts and skills covered in the task, as part of formative assessment. It was, therefore, not always easy to tell when it was time to administer the performance tasks. As a result the administration of some assessment tasks were delayed in some schools.

3.14.3 Availability of time

Finding a convenient period of time to administer tests and to interview participants was a major challenge for this study. The prolonged period of data collection required several visits to the schools and perseverance in negotiating for time slots. The difficulty of securing time for activities in this study delayed implementation of data collection in some instances.

3.14.4 Competing activities

A number of extra-curricular activities took place in the schools. These activities included athletics, music, drum majorettes, Dream for Africa gardening projects and other sporting activities like soccer and netball. Extra time was needed for these activities. It was common practice for schools to take some time off teaching (reducing periods by 5 minutes) for these extra-curricular activities, particularly when the school teams had won the first round of competitions and were preparing for the second. The competition for time for the various school activities was experienced by this study in one school where part of the class was taken for music practice while they were conducting a performance assessment task.

3.14.5 Work load

Teachers teaching two classes felt that they preferred to involve both classes in the study, even though the intention of this study was to use one class per school. They said that students in the other classes would feel 'cheated' if they were not part of the study. These teachers were quite aware of the ethical implications of treating students at the same class level differently, particularly when the same teacher taught the same subject to the different classes. Students' enthusiasm to participate was observed in School M3 where a third Form II class that was taught by a different teacher. Even though they were not taught in the contextualised way, they wanted to take the performance assessment tasks. So they were allowed to take the two performance tasks in Electricity although their scripts were not used for the study. In this way, while the study would not use information from students not taught by the participating teachers, it allowed students of the same class level to take part in some of the activities of the participating students. The study observed another ethical issue of free and voluntary participation of non-research-participants.

3.14.6 Attrition

At the beginning of the study five teachers (three female and two male) and the students in the Form II classes they taught were selected to participate in the study. One of the teachers withdrew from the teaching profession and from the study in the early stages of the study. He did not communicate his intention to withdraw from the study and departure from the profession. There was no replacement for the teacher at the school where he taught. No replacement teacher was found for this study because classes in other schools were using the original Swaziland Integrated Science Programme materials, and they were at an advanced stage in their teaching and learning. It would have been impractical and unethical to request them to change to the contextualised teaching approach. In addition replacing this teacher would also have meant that the new teacher would have had to go through the entire induction process. That school would also be likely to participate only in the Air and Living Things unit and not in the Electricity unit. Ethically it would have been improper to make request at very short notice and disrupt their already set programme.

3.15 SUMMARY OF CHAPTER

This chapter describes methodological issues of this study. The research design chosen for the study and its relevance for the selection of participants, data collection, analysis and reporting of results are presented. This chapter has shown that the study was in two parts. One part focused on familiarising the students and the teachers with the alternative assessment approaches selected for this study, that is, performance assessment and context-based assessment models. The familiarisation exercise was also intended to provide the participants with the experience of carrying out these assessment models. The other and main part of the study was exploring perceptions of the participants regarding the use of the alternative assessment models through questionnaires and interviews. Thus the development of the instruments comprised two parts: the assessment instruments, which had to meet a certain level of quality to be used for the familiarisation process, as well as the research instruments. Each set of instruments was taken through the expected validation procedures.

The data and results obtained from the activities described in this chapter are reported in Chapter Four and Chapter Six. Chapter Four reports on the perceptions of students and teachers in relation the hands-on performance assessment model. These results are discussed in Chapter Five. Chapter Six presents the perceptions of students and teachers about the context-based assessment model and these results are discussed in Chapter Seven.

4. CHAPTER 4

RESULTS I: PERCEPTIONS OF PERFORMANCE ASSESSMENT

4.1 INTRODUCTION

This study sought to establish the perceptions and experiences of students and teachers concerning the use of performance assessment and context-based assessment models in Science at the junior secondary school level. The characteristics of each assessment model were discussed in Chapter Two and Chapter Three. The context within which the experiences occurred was context-based science teaching, learning and assessment, as well as hands-on performance based assessment. Four teachers in four schools participated with their students in seven junior secondary school Integrated Science classes. The characteristics of the participating schools and teachers are described below using the codenames M1, M2, M3 and M4 for the schools and pseudonyms for the participating teachers.

This chapter presents results of the investigation into the perceptions and experiences of the participating students and teachers regarding the use of a performance assessment model in assessing learning in Science. The presentation of the results follows a detailed description of the perceptions according to themes and categories generated from students' statements to questionnaire items. Excerpts from the questionnaires, interviews and field notes are used to illustrate the perceptions described.

The chapter is divided into three major sections. Section I deals with the context of this study. It describes the setting and the participants; namely the schools, the classes and the teachers. Section II deals with the students' perceptions about performance assessment and Section III focuses on the teachers' perceptions about performance assessment.

SECTION I

4.2 THE SETTING AND PARTICIPANTS

The setting in which this study took place was urban schools. All schools were located in towns. Participating students and teachers may have had urban or rural backgrounds.

4.2.1 School characteristics

The four participating schools were named School M1, School M2, School M3 and School M4 as reflected in Table 3.1 above, where a brief summary of the schools and participating classes was presented. Characteristics of the schools are described below to give the contextual conditions for this study. Information that has ethical implications is omitted. Such information includes location of the school and distances from each other. Teachers are described separately from the schools for ethical reasons, such as anonymity, respect for privacy and confidentiality.

School M1

School M1 is a co-educational school that runs double stream classes from Form I to Form V. Science teaching facilities are minimal - only one science laboratory is in use and is inadequately equipped. Another laboratory was under construction. Classes were considered by the teacher to be large although they matched the recommended number of 40 students per classroom. There were 40 students in Form 2A and 39 in Form 2B in 2004. The numbers were reduced to 31 in Form 3A and 26 in Form 3B in 2005.

School M2

This school is a co-educational school with good teaching facilities generally, although it has only one laboratory that is used to accommodate ten classes, “which makes it not so easy to conduct classes using the contextualized method” according to the teacher. The school has a full-time laboratory assistant. The numbers of students per class were 40 in Form 2A and 43 in Form 2B in 2004. These numbers were reduced to 27 in Form 3A and 38 in Form 3B in 2005.

School M3

School M3 is a co-educational school. There are two science laboratories, but with inadequate science teaching material. The school runs three streams for each class at the junior secondary school level and a double stream at senior secondary school level. The classes for this study were fairly large. There were 52 students in Form 2A and 52 in Form 2B in the participating classes in 2004. In 2005 the number of students was 49 in Form 3A and 43 in Form 3B. The school makes use of an inexperienced laboratory assistant (that is, a school leaver). This assistant was very helpful during this study.

School M4

This is a girls' school. The school has good science teaching facilities and runs a double stream of classes from Form I to Form V. The number of students in the participating class in this school was 44 in 2004 when they were in Form II and 36 in Form III in 2005.

4.2.2 Teachers and their teaching qualifications

The characteristics of each of the participating teachers and their working conditions are described below. The school characteristics are described separately for ethical reasons.

Inna

Inna is a female in her fifties with twenty-eight years of teaching Science. Her experience further extends to the writing of the contextualised science teaching materials (Matsapha Lessons) used in this study. She has also been involved in the writing of the recent context-based teaching materials undergoing piloting in Forms I-III between 2004 and 2006. She has, in addition, been a marker for Science external examinations at the junior secondary school level for over fifteen years. She taught Integrated Science (SWISP) to students in Form I to Form III with a teaching load of 24 (35 minute) periods (14 hours per week). She holds a Secondary Teacher's Certificate and a Diploma in Education.

Jabulane

Jabulane is a male in his thirties. He holds a Bachelor of Science degree (Biology/Geography) and a Post Graduate Certificate in Education (PGCE). He has eight years of experience in teaching Science but none in the use of context-based teaching materials. During the study period he taught SWISP and Mathematics in Forms I, II and III (Grades 8, 9 and 10) and in Form V (Grade 12), he taught both Science and Mathematics. He also took the students for some computer awareness activities. His teaching load was 34 (40 minute) periods per week (23 hours per week). His additional responsibilities were assisting in the school administrative duties such as producing the school teaching time-table.

Josephine

Josephine is a female in her forties. She holds a Bachelor of Science (Education) degree with 17 years science teaching experience. She also has some experience in the context-based teaching approach as she participated in the development of the Matsapha Lessons but has not used this material regularly. She taught Form I to Form IV classes, teaching SWISP at junior secondary school level, Biology and Combined Science at senior secondary school level with a total of 20 (35 minute) period per week (12 hours per week). Her other school related tasks included membership to the school catering committee, students' science projects and she is house mistress for inter-house athletics.

Lorraine

Lorraine is a female in her forties with 26 years teaching experience. She holds a Secondary Teachers' Diploma and a Diploma in Education. She also has some experience in teaching Science through context-based teaching material. She has been a Science external examinations marker for SWISP for over ten years. Her teaching load was 12 (50 minute) and 6 (1 hour) periods per week (16 hours per week). She taught Form I Mathematics, Form III SWISP, Form IV Combined Science and at Form V she taught the Chemistry component of Science (Physics/Chemistry see Section 1.5.1 c) above). She was also a class teacher and a member of the school's disciplinary committee.

The teachers participating in this study had varied teaching and professional experiences. These experiences were, however, not used in interpreting their perceptions.

SECTION II

This section of the chapter reports on students' perceptions and experiences about the use of performance assessment tasks in assessing science learning. Data were obtained through open ended questionnaires and semi-structured group-interviews, as well as field notes from observations conducted when the performance assessment tasks were administered to the students. Students' perceptions are described and excerpts from questionnaire responses are used to illustrate the sources of the perceptions. Interviews and observation field notes were also used. Students' views are presented according to categories rather than questionnaire items to avoid repetition that was evident in responses to different questions.

4.3 STUDENTS' PERCEPTIONS OF PERFORMANCE ASSESSMENT

The format of the performance assessment tasks is described in Chapter Three. However, it is briefly noted here that each task comprised four stages: the first stage required students to design a plan to use to solve the given problem. The second stage required them to test their plan. In the third stage the students were required to analyse and interpret the results obtained from stage two and then use the information to respond to questions involving the use of scientific information in real life contexts.

Aspects of the questions used in reporting students' perceptions are summarised as follows:

1. good or liked aspects about the performance assessment tasks;
2. learning benefits from using performance assessment tasks;
3. not good or disliked aspects of the performance assessment tasks;
4. kind of preparation students thought they needed for performing the tasks;
5. use of multiple assessment tasks for a single topic;
6. use of groups for performance assessment tasks;
7. recommendations about the use of performance assessment tasks;
8. other feelings about tests used in this study. (See Appendix VIIA for students' questionnaire).

The ATLAS.ti 4.1 programme was used to organise the responses and analyse the data from the questionnaires and interviews and to generate categories as described in Section 3.13.1. These categories were used to frame and report the findings of the study in this chapter.

As can be noted from the summary of the aspects of the questionnaire items, the instrument allowed students to present a wide spectrum of views and experiences from which perceptions could be derived. From the analysis and coding process five main themes emerged from the coded data. The themes were:

1. Cognitive dispositions
2. Affective dispositions
3. Psychomotor disposition
4. Social disposition
5. Recommendations for implementing performance assessment tasks

The themes, categories, sub-categories and perceptions are summarised in Table 4.1 below to show the perceptions aligned to each sub-category. The theme focusing on recommendations deals with students' perceived recommendations for effective implementation of the practical performance assessment model.

Table 4.1 Overview of themes, categories, sub-categories and students' perceptions aligned to sub-categories

Themes and categories	Sub-categories	Perceptions	
Cognitive dispositions:	Task complexity (TCC)	Ease/difficulty of understanding of task	
Task characteristics	Task importance and value (TCIV)	Improvement of learning and understanding	
		Knowledge acquisition	
		Improvement of memory and recall	
		Relevance of knowledge	
		Assessment value	
	Task requirements (TCREQ)	Instructional preparation (TCREQ1)	Availability of time to prepare for task
			Teaching-learning pace
			Quality of readiness
			Prior instruction
		In task readiness (TCREQ2	Students Expectations of task
			Understanding topic
			Task familiarity
		Task resources (PATRES	Equipment requirement
			Time
Social disposition/ Group assessment	Peer collaboration and support (PCS)	Promoting understanding	
		Quality response and success	
	Perceived group assessment problems (GAPRO).	Non-participation and sharing of marks	
		Dependence on others	
		Dominance	
		Conflicting discussions	
Metacognition	Metacognition (MET)	Monitoring own learning and knowledge level	
Affective dispositions	Extrinsic (ADME)	Higher pass rates	
		Shared marks	
		Recognition by teacher	
	Intrinsic (ADMI).	Interest	
		Laboratory safety	
Psychomotor disposition	Handling apparatus (PYSCHAN)	Practical, procedural, process and manipulative skills	
Recommendations for implementing performance assessment tasks	Task complexity	Task instruction and content focus	
	Task importance and value	Retention and recall	
		Improving understanding	
		Skills for future	
	Peer collaboration and support		
	Group assessment problems		
	Motivation	Feedback, teacher support during tasks	
	Resources	Time (teaching and assessment task)	
		Equipment (sufficient functioning)	

The categories and sub-categories are explained in the relevant sections where perceptions are presented. In reporting the data frequency tables are provided to indicate the distribution of the responses in schools and with regard to questionnaire items in each category.

Perceptions associated with any of the categories are illustrated by excerpts from questionnaires, interviews or field notes thus:

> *It is* easy to understand through the experiments (1 M3B G5).

In the above excerpts the notations are interpreted as illustrated below:

1	represents Question 1 of the questionnaire
M3B	represent the School M3 and class B (Form IIB or Form IIIB)
G5	represents the student group assigned the number 5.

Italics are used to add text for clarity or siSwati verbalisations

4.3.1 Task characteristics

Brookhart (1997) describes task characteristics as properties of tasks that deal with the nature of the tasks whether they comprise essay or multiple choice or free response questions or practical tasks and their demand from the students. They comprise task complexity, task importance and value, as well as task requirements (Brookhart and DeVoge 1999). Brookhart and DeVoge (1999) further note that students perceive assessment and lesson tasks differently. Some students may perceive a task to be simple while others perceive it to be difficult. Some students may perceive it to be of importance in their lives while others may not see any value in doing the task. Students' perceptions in the category of task characteristics are discussed below.

a) Task complexity

Task complexity (TCC) is used in this study to denote the perceived level of difficulty of a task experienced by students in interpreting and understanding the performance assessment tasks used in this study. It also involves identifying and selecting appropriate information for use in carrying out the task and in comparing recalled information when selecting appropriate or relevant information for responding to the task (Buckland and Florian 1991). According to Brookhart and DeVoge (1999) task complexity is influenced by students' self-efficacy, that is, their perceived ability to successfully complete a task. Students judge their self-efficacy by weighing task difficulty, amount of effort required for doing the task and the amount of assistance available, against their perceptions of past performances and experience.

Responses that indicated conceptual accessibility to the task, level of difficulty of interpreting or responding to the task and locating or generating answers, were assigned to this sub-category.

In this study students' perceptions about the level of difficulty of the tasks were reflected in statements responding to Questions 1, 2, 3, 6 and 8 of the questionnaire. The distribution of responses dealing with views that related to task complexity are shown in Table 4.2 below. Statements on task complexity made in response to Question 7 are discussed under recommendations in section 4.3.7.

Table: 4.2 Frequency of citations related to task complexity by question and school

Category	**M1**	**M2**	**M3**	**M4**	**Total**
TCC(Q1good/liked)	4	1	16	8	**29**
TCC (Q3 disliked)	1	0	2	2	**5**
TCC (Q6group assessment)	1	0	0	0	**1**
TCC (Q8) open	1	0	1	0	**2**

The data in Table 4.2 show that there were twenty-nine citations that indicated perceived good aspects of performance assessment tasks that were associated with task complexity, while there were five citations that indicated poor aspects of the tasks.

Students who perceived the tasks to be easy, related the low difficulty to interpretation, understanding and responding to the task. They also associated perceived ease with which they understood the task to the presentation of tasks in a practical format and the use of group assessment. Group assessment discussed in section 4.3.4 below.

i) Practical format and easy understanding of task

Perceived low difficulty of tasks was associated with the practical nature of performance assessment tasks. Some students felt that tasks were easier to understand than theory if they had encountered similar practical activities in lessons. It was easier for them to remember those activities and link to them to the task at hand. They also felt the tasks required less studying.

Perceptions that tasks were easy to understand and which were associated with the practical nature of the task, were presented in responses such as the ones captured in the excerpts below.

> Practical tasks are good because they are easy to understand (1 M3B G6).
>
> It's easy to understand practical than theory because in practicals it is not easy to forget what you did practically (1 M3B G3).

Ease of understanding the tasks was also perceived from the perspective of the different stages of performing the tasks. Table 4.3 below illustrates the proportions of responses indicating the stages that students liked or disliked.

Table 4.3 Frequencies of liked and/or disliked stages of task

Stages of task	Liked	Disliked
All-P,T, AQ	**2**	**0**
Answering questions	**3**	**4**
Planning	**7**	**4**
Planning and testing	**1**	**0**
Testing	**21**	**0**
Testing and answering question	**1**	**0**

P=planning; T= testing, AQ=answering questions

Data from the table show that all stages were liked with twenty-one citations in favour of the testing stage and no citation indicating that it was disliked. Seven citations indicated that planning was liked while three citations were on liking answering questions. There were four citations each indicating a dislike for planning and answering questions. There were also entries for liked combinations of stages. Students seemed to have experienced varied levels of difficulty from the performance tasks as a whole and from each of the stages. These experiences are described below.

Students perceived the testing stage to be more meaningful and afforded them opportunities to obtain data required for answering questions correctly.

> In testing you see the process really happening and you can be able to draw conclude (*sic*) (1 M1 G1).
>
> Testing, because when testing we finally get an accurate answer (1 M3A G8).

Thus the testing stage was perceived to make it easy for the students to answer the sub-questions by providing the evidence needed. Other justifications for liking the testing stage was the perception that there is less thinking involved in the testing stage and it was therefore easy for the students to carry it out. The demand for thinking seemed a "distressful" experience for certain students. The excerpt below illustrates some of the reasons for preferences for planning and testing stages of the tasks.

> Planning and Testing: Planning – we liked planning because we knew that planning would help us learn or share the knowledge we had individually which we brought together and helped each and everyone of our group members and this knowledge we had really gave a positive out-coming, it was very courageous to us. Testing – We liked testing because it is a much easier task to carry out and it needs less brain usage which is a very distressful task and connecting the given apparatus was easier as compared to planning. It also helped us understand the steps and the aim of the experiments and they gave a very understandable conclusion helping us understand what this topic is about (1 M4 G5).

Also implied in the comprehensive response by M4 G5 above, was that collaborated contributions by individual members of the group facilitated the mind challenging planning stage towards rewarding results. The perceived role of the tasks in supporting student learning of procedural skills and subject content is also evident in this excerpt.

Students perceived the planning stage as a crucial stage for success of the tasks. Even though planning may have been an intellectually challenging stage for the students, the use of group assessment allowed them to share ideas and develop good plans that led to their success in the task. Supporting the perceived contribution of effective planning towards producing quality answers, is the following excerpt:

> The planning part has been the best task to deal with because all good results come from the planning. It simply means the planning is the first approach towards the good results (1 M3B G1).

Furthermore, the display of equipment and apparatus to be used for the task was perceived to help reduce the difficulty of planning for a task. Students developed their plans according to the equipment provided and without much need to recall what equipment to use. For instance, one group indicated that they liked the planning stage because they could see the equipment needed.

> It is the planning because when planning you can see all the apparatus you will be dealing with. The teacher also helps us to know what to do with the apparatus (1 M3A G4).

Teachers' support in instances where students experienced difficulties was also appreciated by the students.

There appeared to be some differences in views regarding the stage of answering questions. Whilst some students felt that the questions were straight forward, others felt that answering questions was a challenging.

> We liked the answering of questions because the questions are straight forward and they are about our daily life (1 M4 G3).
>
> It is easier for us to plan and test than answering difficult questions (1 M1 G3).
>
> *We disliked* the questioning part. Some questions we were not able to understand (3 M3B G7).

Difficulties students experienced in answering questions seemed to be also associated with interpreting and understanding the sub-questions.

ii) Task difficulty and topic content

There were also indicators that perceived level of difficulty of the tasks might depend on how easy it was for students to conceptualise the concepts in the topic, the cognitive demands questions made on students and familiarity with the performance assessment model. The following interview discussion illustrates this observation.

Victoria:	How do you feel about the questions that were asked after the activity?
Norah:	For the first test it was difficult.
Victoria:	The first test which was on?
Norah:	Conductors.
Students:	Electricity.
Victoria:	Conductors. Ya.
Norah:	It was very difficult as in for everyone (*sic*) because it was our first time to be asked in this way and… It was very difficult.
Victoria:	What was difficult, was it the questions or the …
Norah:	The way the questions were asked. It was like we were not used to that way…Like this one, "which of the materials you tested would be most suitable for replacing part B in the picture" then you had to use your mind, look at part B, know what it is used for, know what its function is. So it was like it was very difficult.
Victoria:	It was very difficult.
Norah:	We had to think, it was not like "what is photosynthesis" we just know it from the notes. Here you had to use your mind plus what you know from class (M4 Interview).

The perceived high demand for cognitive processing and integration of information from notebooks, data from Stage 2 and the context provided in the questions seemed to make the task difficult. The task referred to in the excerpt above was perceived to require more than the recall of information that the students were used to. The excerpt also reflects that students did not seem to favour engaging in deep thinking about the concepts assessed. However, there were indications that students did value the opportunity to think deeply about the content and context and felt that the tasks provided appropriate ways of learning in Science.

Perceived task difficulty depended on the topic and the similarity of the task to class activities. This was demonstrated through students' indications of enjoyment of the task on testing gases (see Appendix IIB), but not those on electricity. Those students who indicated having pleasant experiences in working on the tasks, described them thus:

Norah: That one was very nice …That one was, like most of the class enjoyed it. Let me say it was nice because most of the things were from our notes. We also learned about Electricity but the way the questions were asked … But with this one Air and Life and testing gases was very nice.

Victoria: But also with that one the questions here are not straight forward [*using task sheet as a reference*].

Nandi: But then you, they don't require too much thinking like the other one on electricity because you had to know a lot about electricity and you also had to think. And this one you just applied general knowledge in most of the questions and the things you studied. This one here…

Norah: And testing of gases is something we enjoyed like when we did the experiments we enjoyed them in class. (Victoria: Ok.) Ooh! It was very nice doing it in a test (M4 Interview).

In addition, perceived difficulties of the tasks were attributed to mismatched expectations of the tasks and students' preparation for the tasks. Students indicated that they studied for the task in the same way they did for other tests. There were perceptions that the tasks were insufficiently linked to the way they had prepared themselves for the tasks.

Victoria: Alright. Is there anything that you did not like about the practical tests? ...

Nandi: The test required too much thinking.

Norah: And what upset me the most was that all the things I studied did not come out (some support from other interviewees). It was not straight forward you had to use a lot of your mind, more than the things I had studied like if the wires are that that is that, those things didn't come out. (Victoria:

> They did not come out that way) Yes. In the way I expected them, most of us expected them they did not. …
>
> Norah: And we also didn't like the part about calculations because most of us are not very good in maths… We were not expecting all these things. We studied things like static electricity what is static electricity? (M4, Interview).

The difficulties students experienced with the performance assessment tasks were perceived to be due to the cognitive demands of thinking, analysis of experimental results and use of the data to answer context-based questions, mathematical skills, as well as students' unfamiliarity with the assessment model used. There were indications that perceived expectations about the tasks and experiences from previous assessment models influenced how the students prepared themselves for the tasks and the effort they invested in performing the tasks. Norah's and Nandi's verbalisations in the interview excerpt above, seem to indicate that the students were used to low level cognitive questions, so they may have found the tasks especially challenging. Self-efficacy in mathematical skills affected perceived difficulty of the electricity tasks involving the calculation of resistance. Teachers' perceptions, discussed in Section III, were in line with the above students' experiences.

Some students acknowledged that they experienced confusion in arranging the ammeter and voltmeter in their circuits or selecting material to use in the testing of gases. Recalling from previous activities was not always easy for students. During the planning stage some students tried to recreate what they had done during lessons, drew diagrams from their workbooks but could not select the appropriate setup for the ammeter and voltmeter required in the task. An example of students' dislike of the performance assessment tasks due to confusion, was presented as follows:

> The confusion of some experiments because some look alike when they are not, which is tricky e.g. confusing the meters (voltmeter and ammeter) when connecting them oppositely (3 M1 G7).

Students' apparent confusion about arranging the instruments was also observed when the task was administered. The excerpt below supports the students' observation of confusing the instruments.

> The planning took longer than anticipated – possibly because there were several trials of circuit diagrams- recalling a number of the circuit diagrams they had drawn during

> lessons. Many of the recalled diagrams were not suitable for the circuit required for determining the resistance of a current carrying wire (Field notes M4 Task 2).

Students did not show any sign of confusion when they worked on Performance Assessment Task 1 on the electrical conductivity of materials. These observations could imply that different tasks stimulate different reactions in students and a possible dependence of the student confusion on the task.

iii) Easy understanding of task and group assessment

Perceived ease with which tasks were understood was also associated with group assessment, where students could share ideas. The excerpt below illustrates an additional factor which was associated with students' perceptions that tasks were easy and beneficial.

> It is easy to understand and it is fun, we get to know how to handle the apparatus, we also share ideas among ourselves (1 M3A G8).

Factors that contributed to perceptions that tasks were not difficult were directly linked to the task, as well as to the use of group assessment. Student also indicated that they had fun as they worked on the task. Developing manipulative skills was an added benefit of the task.

Students' perceptions that were associated with task complexity are summarised through the following points:

- Students' perceptions varied and depended on factors such as:
    - level of difficulty of topic content on which the tasks were based;
    - nature of tasks whether practical or theoretical;
    - expectations of assessment demands – cognitive demand of questions;
    - mode of administration – individual or group assessment;
    - similarity of task to previous class experiences.
- Students perceived task difficulty was reduced through:
    - the use of practical tasks that allowed students to interact with apparatus and generate data that also provided trusted evidence for answering the task questions;
    - the use of a group assessment approach where students collaborated to produce successful plans to obtain a single quality response;

- similarity of task to previous in-class or out-of-class experiences facilitated easy recall;
- improved familiarity with performance assessment model.

- Task difficulty was associated with:
  - unfamiliarity with the assessment model;
  - unfamiliar assessment demands of thinking over and above recall;
  - confusing task activities with similar activities previously encountered in lesson;
  - perceived incompetence in mathematical skills.

b) Task importance and value

Perceived importance, utility and value of engaging in a task play an essential role in motivating students to embark on that task (Brookhart 1997). The task importance and value (TCIV) category was used in this study to describe aspects of the tasks that students considered to contribute to their learning of subject knowledge and skills, as well as in the judging of such learning. Students' responses placed in this category reflected some acquisition of, or improvement of students' knowledge and skills through the following:

- Opportunities for developing skills (cognitive, intellectual, and practical).
- Studying and revising their work.
- Improving understanding of the subject.
- Improving memory and retention of learned material.
- Using knowledge in new situations.
- Problem-solving (investigations and observations).

The frequency and distribution of citations that are linked to perceived task importance and value are shown in Table 4.4 below.

Table 4.4 Frequency and distribution of citations on task importance and value by question and school

Code	M1	M2	M3	M4	Total
TCIV(Q1)	5	3	20	10	**38**
TCIV (Q3)	0	2	0	0	**2**
TCIV (Q5)	5	7	11	4	**27**
TCIV (Q6)	1	1	0	0	**2**
TCIV (Q8)	3	5	6	2	**16**

Table 4.4 shows that perceptions of task importance and value were reflected in six of the seven questionnaire items and by all schools. Citations on TCIV raised in Question seven are discussed under recommendations in section 4.3.7. Students' perceptions relating to task importance and value are discussed below.

i) Improving learning and understanding

Not only were performance assessment tasks perceived by students to be easy to some degree, they were also perceived to help students improve their understanding of the scientific concepts embedded in the tasks. Responses reflected perceptions that performance assessment tasks provided learning and assessment opportunities. The distinction between these two opportunities was, however, not always clear from the responses. Even in instances where students mentioned the performance tasks as a test, the rest of the response portrayed the tasks as learning activities, as the excerpt below demonstrates.

> Using practical test tasks to test you on what you know and can do in science, helps us understand topics we are dealing with, and it helps one know the steps on how apparatus and experiments are set up in order to have more knowledge in them (1 M4 G5).

Performance assessment tasks were perceived to support learning through improving students' understanding of task content, acquisition of knowledge and procedural skills through first hand experience of using equipment to generate useful data. Below are examples which reflect students' feelings:

> At the end of it all you find yourself having more knowledge. For example when you are given a number of objects to test whether current is able to flow through them and at the end of it all you will know which are conductors of electricity because you have tested them all (1 M1 G2).

and that:

> Using practical task tests is good because it enables us to know more about what we have been doing as we see the things being done, unlike when we have to learn and write tests afterwards without doing practicals and see what we were learning (1 M4 G2).

The excerpts above also show a perception that the tasks consolidated and provided further learning of lesson content by demonstrating concepts with regard to the topic.

Students also perceived the performance tasks to promote the development of cognitive skills, thus:

> It improves our way and standard of thinking. It helps us understand better because we see and touch the things (1 M4 G4).

Not only were tasks found useful for enhancing learning with understanding and deep thinking processes, they were also perceived to improve retention and recall of concepts and skills.

ii) Improving memory and recall

Students perceived the performance tasks to improve their memory of the practical activities and the content and skills learned, enabling them to recall these later without much effort in studying. An excerpt illustrating this perception is given below:

> Feel good because when doing P tasks we know much better than compared to when writing topic tests. In that way PT are better because we remember the things we did when writing a topic test (*sic*) (5 M2A G1).

Perceived enhanced retention and recall was also attributed to the testing stage of the performance assessment task.

> Testing:-It helps us when we come to theory; we are able to write it down because we have done it before (1 M3A G7).

> Testing because you get to know how to handle apparatus, you do not have to think too much and you do not forget easily. Planning helps you to use right apparatus for the correct method (1 M3 B G4).

Through performing the assessment tasks it seems students were able to visualised and conceptualise scientific concepts for better retention and easy recall in later assessment. Also conveyed in the second excerpt are ideas that the testing stage was easier, the planning stage promoted the use of appropriate practical procedures and handling of equipment. These ideas were already raised above and will be noted below in discussing perceptions associated with psychomotor benefits of performance assessment tasks.

iii) Relevant knowledge

The performance assessment tasks were specifically designed to integrate real life situations. They were thus aligned to the science applications-led teaching approach used in this study. This intended purpose was positively perceived by the students who recognised the relevance of the content in the task to their everyday life situations. The excerpts given below illustrate these perceptions.

> It would help you in future if ever one day you face a problem you will be able to solve it e.g. electricity (6 M2B G3).

> It is challenging because you have to be observant about what is happening around and what you did for us was good and it will help us in future (8 M4 G7)

> It was interesting because Electricity and Air and Life are the things we observe in our environment (8 M3 A G1).

Students also perceived the tasks to be increasing their awareness of the uses of science in their environment. Students indicated and appreciated that the tasks required them to be more observant and aware of the uses of science in real life. Furthermore, students indicated that the skills required for answering the questions and acquired through the same process would be useful in future endeavours.

Much of the students' perceived importance and value of the tasks seemed to be associated more with student intellectual development. There were, however, some students who recognised the assessment value of the performance assessment tasks.

iv) Assessment value

There were students who acknowledged that the tasks could be used to check the extent of students' learning. Performance assessment tasks were thus perceived as general assessment tasks through which a teacher could tell whether students understood or not.

> It was the best assessment we have ever experienced (8 M3 B G5).

> It is good to do performance tasks and it is good to access (*sic: assess*) student after lesson because you will pass and gain more knowledge (8 M2 3B G1).

In summary, students' perceptions about task importance and value related largely to educational gains from the tasks. Learning benefits from performing the tasks assessment comprised improved understanding of the subject content and more effective construction of knowledge facilitated by observations from the practical demonstration of concepts and development of procedural skills. These gains resulted in improved retention and recall of subject matter.

Students also felt that through the use of the performance assessment tasks they acquired relevant knowledge and skills as they became more observant and aware of the uses of science in their environment. Such a perception indicated that the performance assessment model complemented context-based science learning and assessment.

Students recognised the learning benefits, as well as the assessment benefits of the performance assessment tasks for the students and the teachers.

c) Task requirements

Perceived task requirements (TCREQ) were concerned with requirements from instruction and preparation by individual students prior to performing the tasks, that is, pre-task requirements (TCREQ1), and requirements for performing the tasks - in-task requirements (TCREQ2). Included in TCREQ2 were activities that reflected students' abilities such as thinking, assembling equipment and performing tests on materials, making and recording observations or drawing meaningful conclusions from the data.

To ascertain students' perceptions on task requirements students were asked what they thought was necessary for them to be prepared and ready for conducting the tasks (Question 4). Their responses reflected both pre-task preparation and in-task demonstration of their abilities. Students' responses that related to task requirements were found in Questions 1, 3, 4 and 8. Those expressed in Question 7 are dealt with in section 4.3.7, which focuses on recommendations, below. Table 4.5 shows the frequency of citations for pre-task and in-task requirements by question and by school.

Table 4.5 Frequency and distribution of citations related to task requirements by question and school

Category	M1	M2	M3	M4	Total
TCREQ (Q1)	0	0	0	1	**1**
TCREQ (Q3)	0	0	1	2	**3**
TCREQ1 (Q4)	5	8	13	7	**33**
TCREQ2 (Q4)	0	4	4	3	**11**
TCREQ (Q8)	0	0	2	2	**4**

Table 4.5 includes thirty-three citations involving pre-task instructional requirements and eleven involving in-task requirements.

Figure 4.1 below shows a graphic presentation of the data in Table 4.5 on the distribution of citations indicating students' perceived task requirements by question and by school.

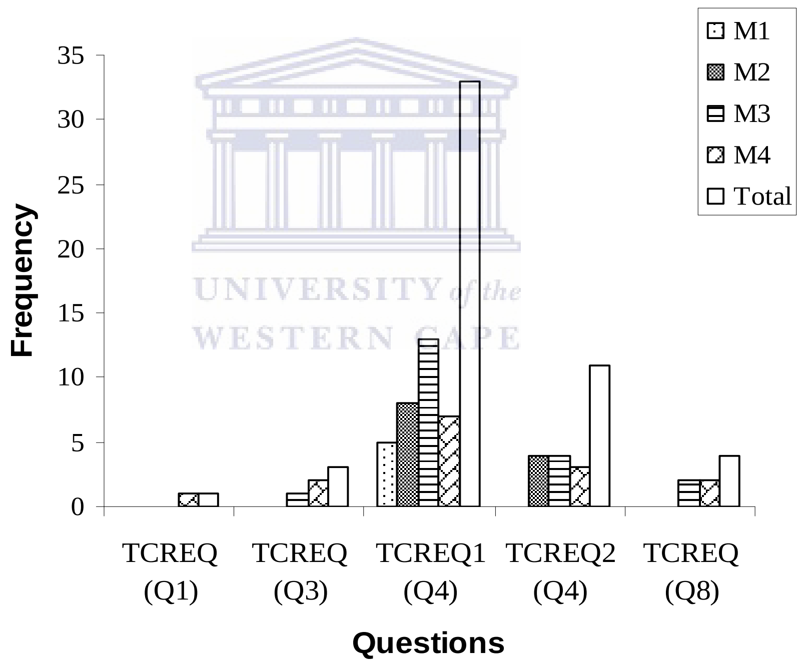


Figure 4.1 Frequency and distribution of citations related to task requirements by question and school

From Figure 4.1 and Table 4.5 it is evident that there were more citations indicating a student perceived importance of pre-task preparation for readiness to perform the assessment tasks. Adequate in-task requirements were also considered important.

i) Instructional preparation (TCREQ1)

Instructional preparation was used in reference to the initial classroom events and acquisition of knowledge and skills that were perceived necessary for the successful completion of each performance assessment task. Students felt that it was necessary to be adequately prepared for the task during lessons and prior to performing the tasks. Students indicated different approaches to pre-task preparation such as practical experience, direct involvement in hands-on practical activities to develop the necessary practical and procedural skills. The following exemplary excerpts illustrate these ideas.

> We must have experiments first before having the practical test (4 M1 G5).
>
> We needed to be taught as to how practicals are carried out and what is required. Information from the teacher is helpful (4 M4 G7).

Students' recognition of the need to have direct hands-on experience of practical work as indicated in the above excerpts was also raised in an interview suggesting that teachers need to involve students in the demonstrations. The interview excerpt below gives an idea of the students' wishes.

> I think the teacher can somewhat demonstrate, but also we need to take part in the demonstration, not just to listen to the teacher. We do need to take part in the demonstration (M3A Interview).

Demonstrations were, however, expected to be visible to the students and successful in what they were intended to show.

Students recognised that understanding of concepts was another important aspect of their pre-task preparation. Students acknowledged their active role in achieving such understanding through full participation and attentiveness during lesson activities. These views were illustrated through responses like the excerpts given below.

> Make sure you participate and get all the information during class time (4 M3A G5).
>
> We need to understand when the teacher is teaching and should study and understand the notes she gives us (4 M4 G1).
>
> It is all about paying attention during the lesson time and when the test comes you simply kindle the information by browsing over the book, ask yourself questions as you do it try to say the teacher's important words (4 M3B G2).

The excerpts above indicate a perception that students and teachers play a complementary role in getting students ready to carry out performance assessment tasks. Teachers needed to provide the opportunities for students to develop appropriate knowledge and skills to meet the demands of tasks, while students needed to be more attentive and strive for understanding.

Perceived need for attentiveness during lessons was also encouraged by the knowledge that students would be assessed through performance assessment tasks at some stage before the end of the unit. The following interview excerpt illustrates this point. It follows a discussion on students' feelings about the inclusion of performance assessment tasks as part of students' tests.

> Jabu: Yes it becomes better because in class you don't concentrate when we are doing experiments. If we do practicals [*performance assessment tasks*] we will know that after this experiment there's gonna (*sic)* be a practical that we will do. So we concentrate on what is going on so it would be easier in the test.
>
> Victoria: Ah! So you pay more attention?
>
> Jabu: We pay more attention to what the teacher is doing because you know what's going to happen afterwards (M4 Interview).

Students perceived the performance assessment tasks to encourage them to concentrate more during the lessons. Students, in addition, needed to concentrate on practical activities and also pay attention to the teachers' deliberations during lessons to ensure their understanding of the content and preparedness for the tasks. A high degree of concentration and effort towards understanding lesson deliberations was perceived to be a better way of getting ready for the tasks than studying.

To summarise the section on pre-task requirements, the data shows that students' perceptions on pre-task preparations for the performance tasks involved input from both the teachers and the students. Perceived instructional requirements thus included:

- practice in practical work to acquire the necessary skills, knowledge and procedures;
- effort by students towards attentiveness and concentration during lesson activities; and
- effort by students and teachers to ensure understanding of the science principles taught.

Post-instruction pre-task input from students by means of revision and studying of the work done was also perceived a necessary preparation strategy, although some felt that more understanding and less studying was necessary.

ii) In-task requirement (TCREQ2)

In-task requirements represented the expectations and the demands of a task on students. In-task requirements involved the abilities (intellectual and physical) students perceived were necessary and important in carrying out the task.

In addition to the perceived importance of pre-task understanding of science concepts embedded in the task, understanding of the task and its procedure(s) were also perceived important for successful completion of the task.

> Students had to understand what we are doing (4 M3B G7).
>
> You need to know what you have learnt before you do that performance task (4 M2A G2).
>
> You should have a bit of theory in science (4 M3A G3).

Understanding of the task and use of previously acquired knowledge were perceived necessary in carrying out performance assessment tasks. The students recognised the balance in the competencies assessed, as students needed to recall and use knowledge in new situations and also display inquiry and practical skills assessed in the task.

The perceived need to possess and display inquiry and manipulative skills is shown in the excerpts below.

> To take the performance tasks you need the knowledge and the steps to take and which chemicals to use when preparing your performance tasks (4 M2B G7).
>
> You should be able to handle the apparatus (4 M3A G3).
>
> He/she must be able to use the apparatus by himself/herself and make a good plan before taking on the task (4 M3B G1).

Other students identified the need for some kind of guidance on expectations and procedures required for the tasks. Thus some students felt that they needed to be given clues on the plan, rather than develop the plan on their own, for example:

> We needed to be given the idea or plan so that we would get the clue of what was expected from us (4 M4 G2).

Students' perceptions of in-task requirements included the intrinsic factors shown below, which were recognised by the student to be essential for successful completion of the tasks:

- the need for understanding the tasks and to use what they know,

- the need to think deeply about the task before responding,
- the need for practical and procedural skills and the ability to use these in a given situation; and
- the need to have theoretical and procedural knowledge to match the balanced competencies assessed in the task.

Students also acknowledged the useful presence of a tutor to help them develop the ability to use the task material and understand the task.

iii) Performance assessment task resources

Students' perceptions of performance assessment task requirements were not only limited to instructional preparation and in-task demands. They extended to the resources required for meeting pre-task and in-task demands of the tasks. Thus perceived requirements also included time and equipment as resources required for performing the tasks. Concerns relating to resources were raised in Questions 3 and 8 by Schools M3 and M4. Most of the concerns about resources were raised by School M3.This information is illustrated in Table 4.6 below.

Table 4.6 Frequency of citations on resource requirements by question and school

Question/Code	M3	M4	Total
Q3-PATRES –E	3	1	4
Q3-PATRES –T	7	1	8
Q3-PATRES –TE	1	0	1
Q8-PATRES –T	2	1	3
Q8-PATRES –TE	1	0	1

E = equipment; T = time, N = notes and school trips

There were students who perceived that performance assessment tasks required a lot of time to carry out to completion. This perception was reflected in responses such as the one given in the following excerpt.

> You find that you are not able to finish the task in the given time because they demand a lot of time (3 M3A G1).

Perceived time constraints were associated with discussions encouraged in group assessment. Group assessment required students to express, justify and convince others about the ideas shared, which students felt required time, as the interview excerpt below show.

Student: The performance tasks take long than when the teacher teaches us in class.
Victoria: What do you think contributes to the taking up of time?
Student: Because, as we are many, there are many ideas so we have to be sure of every idea so we debate (M3A interview)

Thus, while the sharing of ideas and its benefits were appreciated by students, they also recognised the implicit time requirements.

Some students felt that the performance tasks increased the amount of practical work done by the students. This view seemed to confirm the perception that students need practice with practical tasks to be prepared for carrying out the performance assessment tasks. The performance tasks also added to this perceived excess of practical work. Other students perceived the tasks to be slowing down the learning pace. The excerpts below illustrate some of these ideas.

> Wastes time as you do a lot of experiments (3 M3B G5).
>
> It is a slow way of learning (3 M3B G4).

The perception that performance tasks demand a lot of time to carry out seemed to also relate to the duration of the task and the discussions in groups. The frequency of administering tasks was seen by some groups as a poor way of utilising school time.

While the performance assessment tasks were perceived to be time consuming by some students others viewed their use in a positive way. They acknowledged the need for equipment, time and commitment from group members as other factors that may increase time spent on tasks. The excerpts below demonstrate this perception.

> It is challenging. It teaches us that we should be able to achieve something at a specified period; we do not rest until it is done. If we could also have apparatus, little bit more time and may be also disciplined and tamed group members (8 M3B G2).
>
> It is good but the practicals take long although we learn a lot from them (8 M3B G7).

Students acknowledged that adequate materials were necessary for the effective administration of performance assessment tasks. They also recognised that equipment problems, such as adequacy and poor working conditions could also lead to delays in completing the tasks.

The effects of mal-functioning equipment were more pronounced in the performance tasks on electricity due to the delicate nature of the equipment. The main problem was constructing complete electric circuits in order to get the readings required for the other parts of the tasks. Connecting wires sometimes got disconnected from the clips and needed to be repaired. Circuit boards sometimes had loose contact points. Both these faults led to incomplete circuits and therefore no results for some groups and increased time spent on task. It should be acknowledged that students' handling of the equipment also contributed to problems like disconnections of clips from connecting wires as students were not so careful when changing the wires.

4.3.2 Readiness for the task

In addition to establishing students' perceived preparation requirements for the performance assessment tasks they were asked about their perceived readiness for the tasks. Students perceived their readiness or lack of it in terms of formal instruction during lessons and out of class everyday life experiences. Thus students' perceived readiness for the tasks involved the following factors:

- prior experiences (direct or observed);
- task familiarity;
- skills possessed;
- level of understanding of topic; and
- time available to prepare for the tasks (study or instructional).

Each of the above factors is discussed below.

a) Prior experience

The perceived importance of prior instructional and out of class experiences in perceived readiness for the performance tasks can be demonstrated through the following responses.

> We were ready because we were already taught about it, also we have some ideas. Some we experience them (4 M1 G3).

> We were ready because we studied and the teacher told us about the practical test that we were going to have so we prepared everything and got ready for the practical test (4 M3A G6).

Perceived readiness also seemed to be an outcome of the amount of effort invested in preparation for the tasks.

b) Understanding the topic content

Students' perceived readiness for embarking on the tasks seemed to be linked to their understanding of the topic content. For some groups the topic on electricity was rather difficult to understand and therefore, their readiness was affected.

> We were not ready for the two practicals on electricity because the whole chapter on electricity was hard to understand, it would still be difficult for us if we would be given a test even now but the one on Air and Life was okay, it was really easy and we understood it (4 M4 G2).
>
> We were ready (very) on the practical tasks about Air and Life than the one about Electricity because we understood the one about Air and Life clearly (4 M3B G6).

The two excerpts show that the students in the two groups were not ready because they had difficulty understanding the unit on Electricity, although they were ready for the task on Air and Living Things. These observations seem to indicate that students' readiness for tasks was topic- (or task-) dependent.

c) Task familiarity and expectations

Perceived readiness for the performance assessment tasks seemed also to be affected by students' familiarity with or prior exposure to the task content or similar task format.

> We were ready but the problem was that on the Electricity practical we were not sure of how the practical tasks are conducted (4 M3A G5).
>
> We were not ready because we did not know what the practical tasks were all about (4 M1 G6).

The perceptions illustrated in the above excerpt concur with those already raised earlier regarding the need for adequate pre-task instructional preparation.

As students became exposed to more performance assessment tasks their level of readiness seemed to improve as they became more aware of what to expect from the tasks. The following excerpts provide an example of the evolution of experiences from working on the performance assessment tasks as described by one group of students.

> As for the first one, we were nervous because the first people who took on the tasks took a long time as a result our turn had to be postponed (4 M3B G1-Task 1).
>
> We were not fully prepared because we were still not sure of attending to these tasks alone (4 M3B G1-Task 2).
>
> We came fully prepared for this one. We forgot about what happened with the first two as a result we scooped maximum marks (4 M3B G1-Task 3).

The three excerpts above show a possible development of familiarity with the assessment approach and the tasks that led to students' perceived readiness for later tasks. However, since perceived level of difficulty of the task may be content or task dependent, students' readiness for the tasks could also be a factor concerning how easy it was for them to understand the content of the unit on Air and Living Things.

d) Availability of preparation time

Being informed about taking the performance task on time was perceived by students to be important for their readiness for the tasks. There was, however, some discrepancy as to what constituted being informed "on time", as the first two excerpts from groups in the same class indicate below.

> We were ready because we studied and the teacher told us about the practical test that we were going to have so we prepared everything and got ready for the practical test (4 M3A G6).
>
> We were not ready because we were not told in time (4 M3A G8).
>
> We were definitely ready because we were told in time about the coming practical tests (4 M4 G1).

However, the group responses show an agreement about the relationship between readiness and being informed of the tasks on time. It seemed important for students to know in advance about the performance tasks.

Students' perceived readiness for performance assessment tasks seemed to depend on factors such as adequacy of preparation in prior instruction, task familiarity, and understanding of unit/topic content, task expectations and the availability of time to prepare for tasks. Students gave the performance tasks the same status as tests and examinations for which they needed to organise and prepare.

4.3.3 Metacognition

Metacognition can be described as a self-monitoring approach that can help students develop the ability to take control of their own learning. Donovan and Bransford (2005) note that teaching practices that emphasise self-assessment tend to encourage students to be more metacognitive about their own thinking and learning. Providing support for self-assessment is thus an important component of instruction.

The performance assessment tasks used in this study provided opportunities for students to plan and test their own ideas in understanding and solving given problem tasks. Through interacting with each other and sharing ideas in groups and analysing results from task activities, students could reflect on their ideas and understanding, self-assess and modify their understanding. Discussions within groups allowed individuals to express and examine their ideas, and select those ideas that made sense and provided the best approach for tackling the performance assessment tasks.

Students also appreciated the opportunity to assess their progress on the tasks. The excerpt below indicates one way in which students could evaluate their experimental procedures.

> You are able to remember all of the points because you are able to see that you left out something if what you are doing is not a success (1 M4 G3).

Students perceived the performance assessment tasks to provide opportunities to go back and check their procedures to ensure that they obtained good data for answering the task sub-questions. Students were able to retrace their steps in instances of unsuccessful outcomes of the investigation step and repeating data collection procedures as necessary.

Students perceived the tasks as an alternative lens through which their level of understanding could be gauged by their teachers and themselves. The excerpts below illustrate this view.

> It helps boost one another in their marks and also helps in knowing your knowledge towards the subject (science) (6 M2B G7).
>
> It is excellent because you can verify from other groups (6 M3B G5).
>
> It made us realise our potential and sector**(sic)* which will help us improve or try something else if it's not our sector unlike forcing matters (2 M3B G2). *(sector* seems to be used in relation to employment or career/vocation)*

Tasks helped students to reflect on their knowledge of science and confirm what they knew against what other students or data from the task presented. Self-assessment of understanding and capabilities for future self-development was another perceived benefit of tasks for students.

4.3.4 Group assessment

Cooperative group assessment was used to administer the performance assessment tasks. Responses from different questions indicated that students believed that group assessment was a characteristic of the performance assessment model. Group performance assessment was associated with good aspects and problems. These are discussed under the sub-categories of peer collaboration and support (PCS) and group assessment problems (GAPRO).

a) Peer collaboration and support

Peer collaboration and support (PCS) described responses that indicated views about the task that involved students working as teams and helping each other while carrying out the tasks. It also involved ideas that reflected the mutual sharing of ideas and collaborative efforts towards the best approach to tackling the task. Table 4.7 below shows the frequency and distribution of the responses that reflected perceptions of peer collaboration and support by question and by school.

Table 4.7 Frequency and distribution of citations related to peer collaboration and support by question and school

Category	M1	M2	M3	M4	Total
PCS (Q1)	2	9	7	2	**20**
PCS (Q6)	5 (1)	3 (2)	12 (7)	7 (2)	**27 (12)**
PCS (Q8)	1	0	1	1	**3**

(n) = the number of groups citing PCS in both Q1 and Q6.

The frequencies shown in the table indicate that peer collaboration and support was perceived to be a good aspect of performance assessment tasks, shown by twenty citations. Twelve of the twenty-seven citations in Question 6 were made by the same groups in Question 1.

Students liked working in groups for reasons such as the sharing of ideas and helping each other, cognitive benefits, social benefits and motivational benefits.

i) Sharing ideas

Sharing ideas during group assessment was perceived to reduce task difficulty, as noted in section 4.3.1 above. The excerpts below illustrate further students' views on benefits of sharing ideas.

> Using practical test tasks is good because you are able to share ideas, maybe if you didn't know how to do a certain practical, your partner can help you (1 M4 G3).
>
> We feel good about working on tasks as a group because it helps us to share different opinions and ideas (6 M1 G3).

Through sharing ideas students learned from each other through reminders, coaching or observing others perform certain skills or processes or from what they said during discussions. Students could also check their own knowledge against what group members contributed during discussions.

ii) Promoting understanding

Working in groups assisted students through collaborative interpretation of the tasks so that they had similar understanding of the task and its requirements. This shared interpretation of tasks was perceived to help those who had difficulty in understanding the task to follow the activity with better understanding. The excerpts below reflect students' perceptions that group assessment facilitates better understanding by group members and reduces the difficulty of the tasks.

> It is because working together gives us more ideas or we share ideas. It helps those who did not understand about the topic and helps them to understand more better when discussing (6 M1 G6).

Norah:	It is good because even those who don't know or understand the topic they understand it better when they hear from their peers, like from us they can understand it.
Victoria:	… (Redirects) So you see this as improving your understanding (Students: Yes) and coming from explanations of peers. (Students: Yes) Anything else that is good about using these? [Pause] Now suppose you were doing it by yourself, that you were not working as a group, would they still be good, a good way of testing?
Students:	No, no.
Victoria:	No!!
Ann:	It's hard when you do it alone. … It's too hard when you do it alone.
Victoria:	Yes.
Ann:	Because I for one am not good in Science when I work in a group what I think is hard is not as hard as it seems because I hear it from a friend. Doing it alone is hard (M4 Interview).

Working on tasks in groups allowed students to support one another under test conditions. Students appreciated sharing ideas, peer-tutoring, reduced task difficulty resulting from peer explanations and the improved comprehension of the tasks and topic content.

iii) Quality responses and success

Group effort in assessment was perceived to lead to good quality responses for tasks. Excerpts reflecting this perception are provided below. The first three excerpts come from the same group of students in responding to different questions: Questions 1, 6 and 8. These three responses show some consistency in the group's perception regarding collaborative effort for success in performing the given assessment task.

> For bringing up different ideas and come out with the best answer (1 M1 G7).

> We feel satisfied because ideas are brought up and one final answer which is best is eliminated (*sic*) (6 M1 G7).

> It is an infective (*sic*) (*effective)* way of studying since we learn through experience and the combination of ideas from different group members get the best answer (8 M1 G7).

> It is good because we get to share the ideas and come out with one answer (6 M3A G6).

Through team work quality answers were produced leading to many students receiving a good score in the tasks.

Students perceived several benefits from the use of the group assessment approach. Benefits that supported intellectual development of the students comprised improving their knowledge bases through peer tutoring and tapping on the knowledge and experiences of group members. Group interpretation of the task reduced task complexity so that more students could understand the task better and know its requirements. Group assessment also improved chances of students passing as quality responses were more likely to be produced through collaboration.

Group assessment was also perceived to depend on the collaborative effort of all students, if they were to obtain optimum benefits. Amicable participation, equal opportunities to participate and respect for one another in the groups were perceived essential for the functioning of group work. They described these views as follows:

> We think working on tasks in groups was okay and it helped us because we would get the knowledge from the other members of the group and we would correct each other and even argue at times, which was part of learning. At the same time it was giving us problems because some people would just relax and expect others to work for them and others who claim to know everything and would just look down upon others because they thought they were the only ones who know and anybody else's idea is wrong (6 M4 G2).

Students recognised the importance of team work for the mutual benefit of all, and that it was more beneficial to participate actively than to rely on others. Support among group members was perceived to be a two way system where members benefit from each other and not have a "host-parasite relationship". These views were indicative of perceived group assessment problems, which are described below.

b) Group assessment problems

Perceived problems associated with group performance assessment varied between schools, with some schools presenting fewer citations on concerns regarding group assessment problems (GAPRO). Problems included variations in contribution and participation by students in the groups, inattentiveness during group activities, dependence on enthusiastic group members, dominance by some members, sharing of marks from group task and conflicting arguments.

The distribution and frequencies of citations that were linked to GAPRO responses are shown in Table 4.8 below.

Table 4.8 Frequency and distribution of citations related to GAPRO by question school

Category	M1	M2	M3	M4	Total
GAPRO (Q1)	0	0	0	1	**1**
GAPRO (Q3)	0	9	11	1	**21**
GAPRO (Q5)	0	2	0	0	**2**
GAPRO (Q6)	0	9(5Q1)	4 (3Q1)	4	**17**

NB: The numbers in brackets indicate the number of groups that cited GAPRO in other questions, for example 9(5Q1) means that of the nine citations five groups made similar citations in Q1.

The data shown in Table 4.8 indicate that there were no complaints about group assessment from School M1. There were several citations on GAPRO for Question 3 and Question 6 and very few in Question 5.

i) Non-participation and sharing of marks

Students complained about non-participation by certain group members during group discussions. The differentiated participation was also observed whilst groups worked on the tasks. The excerpt below was taken from observation field notes written during the administration of performance assessment tasks in School M3 to illustrate this observation.

> The students worked in the groups as best as they could. A few pupils however did not seem active, even when asked to make contributions they would not say anything. There was co-operation in all the groups, though more evident in the smaller groups. There was one group of ten at the front, where all the pupils showed enthusiasm, paying attention. Both boys and girls were recording in the different groups, but in some groups it was evident that the boys were more active in making the connections of electric circuits (Field notes M3).

Students presented varied levels of motivation to participate in the tasks. They also had varied opportunities to participate. Thus, some students were concerned about the differentiated participation they observed. Students perceived factors such as laziness among students, students' lack of knowledge or understanding of what was being done and sheer lack of motivation by group members as being responsible for non-participation.

Awarding of group marks to each group member seemed to be an important source of worrying for some students. Students were concerned that marks were awarded to students who were seen as non-participatory during the group discussions, or not participating enough. The following excerpts reveal how students expressed their discomfort regarding the allocation of marks in group assessment.

> Nandi: I did not like the practical test because when working in groups, so like other people in the group did not participate yet they are going to get the same mark as us (M4 interview).

> It is not good because you find that there are six people per group and you find that only two people only know what we are to do and the rest have no ideas and when the papers are taken to be marked we get the same marks while only the two were working (3 M2B G6).

> Working together as a group on a task where each gets the same mark is not good because some students do not want to share their ideas with others and at the end the whole group gets the same marks whereas only one or two was working (6 M2B G6).

What is also of interest in these excerpts is the seriousness of students' concerns for marks being awarded to non-participating students, even though each concerned student was not deprived of the marks awarded to the other students. The excerpts show that students were concerned that some students benefited even though they did not participate, refused to participate, participated partially, lacked the necessary knowledge to participate or understand what was done, or were "just lazy" to play their part in group assessment.

ii) Dependence on others

Non-participating students were seen as depending and benefiting from other students in their groups, not only in terms of marks but also with regards to the ideas and knowledge shared by contributing students. Different descriptions were used for students who were perceived to be non-participating. The following excerpts illustrate aspect of the tasks that students viewed as not good.

> Do not think the test is good because some of the students do not participate they just relax and use other students' ideas without his own ideas to help the group (3 M2B G4).

> Some pupils are not concentrating when we do practicals and they expect you who is *(sic)* concentrating to give ideas or view points from them (3 M3A G7).

> Other people do not contribute but see it good to lean on other people's strength (3 M3B G1).

Students were concerned that certain students were lazy and lacked commitment towards working on the tasks. Such students were perceived as unfairly depending on willing, committed and enthusiastic students. One group described dependent students as 'stealing their ideas', whilst another group said:

> We feel it is unfair because others do not participate. They are parasites they depend on our knowledge (6 M4 G8).

Other students rationalised non-participation by fellow group members in a more constructive way, thus:

> It is not good because some of us do not concentrate on the work they will not know what is happening and also that they will not see the results of the experiment (3 M2B G5).

> When making the performance task you do it in groups and when coming to writing what you have done in the performance tasks you will find that you don't know anything. This is because when some of the group members work together not all the other group members understand (6 M2A G1).

These students acknowledged that those students who did not participate in group discussions were likely to miss out on learning from the discussions and outcomes of the assessment task.

iii) Dominance by certain students

Alongside the perception of dependence by some members on others, were perceptions that certain individuals in the groups dominated. Some group members were confident in their knowledge and skills so that they took the lead and performed the task. In responding to Question 3 (what was not good about tasks) and Question 6 (how students felt about group assessment) students raised concerns on dominance as illustrated in the following excerpts.

> Other people ignore ideas and views of the other students (3 M3B G1).
>
> Not good because some maybe they did not understand the practical test and then even do something is not expectable (*sic*) and we can lose marks (6 M2A G6).
>
> It was unfair because some of our colleagues would not take our points for they thought their own points were that good to contribute every stuff (*sic*) (6 M4 G4).

Dominance in group assessment was perceived by students to lead to some students not benefiting from discussions because their needs were not attended to in the group. Some students thus felt that they were not always treated with fairness in such groups as ideas from some group members were sometimes disregarded. There was also a concern that dominating students may misunderstand the task and perform procedures that were not required for solving the problem presented in the performance assessment tasks. These concerns seem to indicate that students recognised that participation in groups was in the control of all group members and contributions were to be discussed and accepted by all members in the group. Sometimes no agreement was reached during the discussions. Students' views on disagreements during group discussions are presented below.

iv) Conflict of ideas during discussions

Students seemed aware that group discussions allowed them to express various ideas about the assessment tasks for the purpose of generating one view that they all agreed on. However, sometimes the students experienced lack of agreement. Students expressed their disappointment and frustration regarding disagreements during discussions as follows:

> Practical tests (tasks) are a waste of time and they suck because other people in that group did not come to one same conclusion. Each one wanted her voice only to be heard (1 M4 G8).
>
> We sometimes quarrel over an answer and end up not writing the correct answer (6 M3A G8).

Conflicting discussions seemed to be seen to arise from misunderstanding of ideas from different students. A discouraging experience for the students was that they did not produce answers for the tasks, which they felt reduced their chances of receiving passing grades in the task. Also conveyed in the excerpt by 1 M4 G8 is the perception that discussions without consensus were a waste of time.

All in all, students perceived group performance assessment in encouraging ways. These were improving students' understanding of the task and subject content, generating quality responses and improving performance or achievement. Knowing that they were being assessed seemed to have encouraged students to be more serious and committed in their discussions in order to succeed in the tasks. Students also perceived group performance assessment in ways that reflected their discontent, such as allowing:

- non-participation of students;
- students depending on others;
- students dominating others;
- unfair and undeserving allocation of marks to students who did not make any or enough contributions during group discussions or who refused to cooperate; and
- delays in the generation of answers because of disagreements in groups.

4.3.5 Affective dispositions

The theme dealing with affective dispositions comprised views that related to perceived affective aspects of performance assessment tasks. Affective dispositions comprised intrinsic (ADMI) and extrinsic (ADME) motivational aspects, briefly described below. The distribution of responses relating to affective disposition among schools and questions is provided in Table 4.9 below.

Table 4.9 Frequency and distribution of citations related to ADME and ADMI by question and school

Question	M1		M2		M3		M4	
	ADME	ADMI	ADME	ADMI	ADME	ADMI	ADME	ADMI
Q1	0	0	2	2	2	2	0	1(1^{-}ve)
Q2	0	1	0	0	1	1	1(1^{-}ve)	1
Q3	1	0	2(1^{-}ve)	0	2(1^{-}ve)	0	1(1^{-}ve)	0
Q5	1	1	5	1	6	2	5	0
Q6	0	0	11(6^{-}ve)	0	0	2	0	0

($^{-}$ve represents responses that carry a negative sense such as complaints)

ADME and ADMI aspects were raised in six of the eight questions and to varying degrees by the different schools, as Table 4.9 shows.

a) Affective disposition motivation extrinsic (ADME)

Students' perceptions comprised motivational aspects that reflected influence by extrinsic factors such as marks. Marks indicate the level of success and are usually the driving force of student motivation.

Students felt that performance assessment tasks afforded them better opportunities to pass. Opportunities to pass were believed to be due to the awarding of marks for making observation, relevance of content of the task to students' everyday life experiences and the practical format of the tasks.

> We get marks for correct observations (1 M3A G6).
>
> There are more marks in practicals than theory (2 M3B G3).

In addition to sharing ideas and responsibilities in carrying out the tasks, each member of the group received the same score (group mark) irrespective of input or effort. Some students indicated that they were satisfied with the approach. The excerpts below reflect students' views regarding group marks.

> I like performance tasks because in a group we get equal marks (1 M2B G4).
>
> Good for somebody who do not know anything with the test because s/he is going to obtain the same marks with those who know (6 M2A G4).

Others indicated displeasure in the awarding of equal marks for unequal contributions thus:

> They get free marks because they do not concentrate (3 M3A G7).
>
> It is not good because some members do not co-operate while others contribute and at the end we share the same marks (6 M2B G5).

The concerns that are raised in the above excerpts focus on the non-participation of group members, or less knowledgeable members, who benefited in terms of their marks from the efforts of those students who engaged in active performance of the tasks. Students were concerned that less informed students could get away with deceptive grades from group assessment. Teachers were also concerned about undeserved marks, but acknowledged that not all students could participate or contribute because of problems beyond the students' control (see Section III below).

Performance assessment tasks allowed teachers to interact with students and help them with the tasks. Some marks were withheld when teachers contributed directly to the performance of the task, such as telling the students exactly how to proceed. Students perceived such deduction of marks to be discouraging, as shown by the excerpt below.

> We lose simply, like when you are on the right track but you are slowly *(sic)*, the teacher contributes and takes your marks (3 M3B G2).

Students perceived performance tasks to be a way through which students could improve their grades. It was easy for students to pass because marks were given for specific aspects of performances by students, that is, manipulation of equipment, observations and cooperation, or for being a member of a group of good performing students.

b) Affective disposition motivation intrinsic (ADMI*)*

Responses relating to motivational aspects also dealt with the influence of performance assessment tasks on students' personal interests and fulfilment, anxiety and commitment to performing the task such as attentiveness and concentration on the task, uncertainty, feedback and need to succeed.

There were indications that some students perceived the performance tasks to be just fine, good, interesting and that these improved the enjoyment of Science. Students said, for example, that

> We liked the testing part because it was interesting (1 M3A G3).
>
> We gain more interest because if you are learning about something you have already done and seen before it is easier to understand than learning about something you haven't seen (2 M1 G2).
>
> Norah: That one was very nice …That one was, like most of the class enjoyed it. Let me say it was nice because most of the things were from our notes. We also learned about electricity but the way the questions were asked …. But with this one, Air and Life and testing gases was very nice (M4 interview).

Some students perceived the task to be interesting and improved their interest due to some prior experience with the content of the task and the structure of the questions.

Perceptions relating to affective disposition also emerged in responses dealing with the use of multiple assessment tasks. Both intrinsic and extrinsic motivational aspects were identified. Normally, students in Swaziland secondary schools take monthly topic tests or regular tests. This study used two performance-based assessment tasks for Electricity and one for Air and Living Things together with one unit test for each of the two units. Students perceived the use of more than one assessment opportunity beneficial in motivating them to study, improving their knowledge and understanding of the subject, improving their scores, as well as developing familiarity with tasks and their content.

> We prefer several tasks reason being that: We get maximum marks; we acquire more knowledge; we improve the skill of helping other people in Science (5 M3B G1).
>
> We think that this would be a good idea, we would understand better because we would be seeing the things that we learnt about and we would be familiar with them (5 M4 G2).

Students also felt that they got useful and regular feedback about their progress and could invest more effort to improve previous grades.

> We feel it is a good idea to take several tasks that are used to give us marks because the teacher easily tell us if we understand or not when she gives us a test to write (5M2B G6).
>
> It is good because if you make a mistake in the first tests you can be able to fight and improve your aggregate with the following ones unlike one or two whereby you have no chance of vengeance and it also gives more experience such that you can develop your own skills (5 M3B G2).

Initially students were surprised by the new format for the performance assessment tasks and the unfamiliar practice of taking several assessment tasks for one unit. However, students perceived the use of performance tasks and unit tests to be complementary. The interview interaction below illustrates the perceived complementary nature of the two assessment approaches used in this study.

Victoria:	… currently you write one topic test (Students: yes) but here you wrote a topic test and two other tests and for the other topic you wrote one topic test and a practical test.
Norah:	A lot of us were not happy about this because we were not used to it (Victoria: About this?) three tests. Yes we are not used to three tests in one topic …
Victoria:	Yes, but, yes...
Nandi:	But I think it was good, because, like if the other questions were left out in the other test then they'll be included in the next test. (Victoria: Ok) Then you get to use your mind in the next test, next questions now then you get to exercise your mind in that way.
Victoria:	So you feel that writing so many little pieces of work kind of tests you on different things in the topic?
Students:	Yes.
Norah:	Yes. Like everything that we learned about. Like, yes if we just had one test the teacher picks up something there and there and there, but if we write the three tests we get tested on everything that we did (M4 Interview).

The ideas advanced in this interview excerpt support those presented by 5 M3B G2 above. There is concurrence with Nandi's views on the complementary nature of performance assessment tasks and unit tests.

Other students presented alternative views regarding the use of multiple assessment tasks. They were concerned about the load accompanying the use of such tasks and the possible lowering of final scores.

> One thinks it's too much work because she doesn't specialise in Science only, one thinks it helps to test you as to how much you understand, think and use your brain when learning Science. …. One feels its okay because we have to learn more about Science (5 M4 G7).

> We do no like taking a lot of tasks as tests because we sometimes fail and these marks let us down (5 M4 G1).

Students' views on the use of a multiple assessment approach acknowledge strengths of the approach in enhancing learning and achievement in Science, but also noted the potential for an increased work load for students and its effect on the students' achievement.

In summing up this section, one can note that students expressed different views regarding extrinsic and intrinsic motivational issues. Students' perceived performance assessment tasks to improve their enjoyment of Science and motivation to participate in the tasks and lesson activities with greater commitment.

Students also appreciated the perceived increased chances of passing their Science through better grades acquired from the performance tasks. However, some students were not happy that non-participating students, perceived to be undeserving of the credit, were awarded marks that were the same as those of students who worked hard on the tasks.

Some students appreciated the use of a multiple assessment approach. The students perceived positively the gains in knowledge resulting from the various ways of interacting with the topic content through the assessment tasks, teacher appraisal of students' work and possible remediation, extensive coverage of learning goals in the topic, as well as opportunities to improve poor scores. Workload was, however, a concern considering the number of subjects students do in schools.

4.3.6 Handling apparatus

Handling apparatus and procedures (PSYCHAN) described responses that indicated that students engaged in some form of psychomotor operations through performance assessment tasks. Included here were ideas relating to seeing, touching, manipulating equipment and engaging in other experimental procedures.

Table 4.10 indicates the distribution of the frequency of citations of PSYCHAN related responses by question and school. Statements on handling equipment were made in Questions 1, 6 and 8.

Table 4.10 Frequency and distribution of citations on PSYCHAN by question and school

Question	M1	M2	M3	M4	Total
Q1	1	0	20	4	25
Q6	0	0	1	0	1
Q8	2	0	0	0	2

As can be noted from Table 4.10, opportunities to handle apparatus were perceived to be a good aspect of performance assessment tasks receiving twenty-five citations. Most citations were presented by groups from School M3.

Perceived psychomotor value of performance assessment was providing students with experiences and practice with practical work and procedures, as well as following instructions. The excerpts below help illustrate these views.

> It helps us get more experienced on that practical test (1 M1 G6).
>
> This taught us how to do experiments (8 M1 G6).
>
> It gives us the experience of following instructions and doing tests which we will need at higher institutions (1 M3B G2).
>
> We like the testing stage of the practical task. When we are testing we handle the apparatus, read the observation and come out with answers (1 M3A G6).

Students also perceived the performance assessment tasks to help in the development of psychomotor abilities. Students felt that they learned science process skills of experimenting, measuring, reading instruments, observing, making inferences and predicting. These perceptions were more conspicuous in the case of School M3, but not much mention of these perceptions came from School M1. Both these schools had poorly equipped laboratories. This observation seems to indicate an inconclusive link of perceived opportunities for developing manipulative, procedural and process skills to equipment conditions in the school.

4.3.7 Recommendations on performance assessment task

In Question 7 students were specifically asked for suggestions or recommendations about the future use of performance assessment in Science. Justifications for the recommendation, where given, were found to match the different sub-categories. Unjustified or non-committal responses, like the ones in the excerpts given below, that portray a positive point of view of the task, were taken to imply a recommendation for continued use of performance assessment.

> These practical tasks are just fine (7 M3A G5).
>
> Performance tasks enables us to write what we see (7 M2A G4).

The distribution of responses to Question 7 into the different sub-categories is given in Table 4.11 below.

Table 4.11 Frequency and distribution of citations for recommendations by category, sub-category and school

Category	Sub-categories	M1	M2	M3	M4	Total
Affective disposition	ADME	2	4	3	0	**9**
	ADMI	0	0	1	2	**3**
Group assessment	GAPRO	0	2	1	0	**3**
	PCS	1	1	0	0	**2**
Metacognition	MET	0	0	1	0	**1**
Task characteristics	TCC	2	1	11	6	**20**
	T CIV	8	9	5	8	**30**
	TCREQ	1	0	0	1	**2**
Performance task resources	PATRES-E	0	0	8	0	**8**
	PATRES-N	0	0	1	0	**1**
	PATRES-T	0	0	11	0	**11**
Psychomotor	PSYCHAN	0	0	1	0	**1**
Number of sub-categories with recommendations		**5**	**5**	**10**	**4**	

PATRES-E: -E = equipment; -N = notes, T= time

From Table 4.11 it can be noted that School M3 cited recommendations in ten of the twelve sub-categories while the other schools made recommendation in four and five sub-categories. Also, all schools made recommendations relating to the complexity of the tasks (TCC) and their importance and value (TCIV). Both these categories have higher frequencies of citations. School M3 was the only school to make recommendations relating to performance task resources (PATRES) although only one citation on recommendations on PSYCHAN. Although there were a number of complaints about group assessment problems (GAPRO) yet there were very few citations on GAPRO related recommendations. All these observations do not necessarily imply that the recommendations represented by more citations are more important than those represented by fewer citations. They only show that more groups noted the recommendation.

Students' perceived justifications for the recommendations are described below and their nature is illustrated by means of students' excerpts.

a) Task complexity

Students' recommendations that were associated with task complexity focused on ensuring reduced difficulty of tasks related to task instructions and content focus, as shown by the excerpts below.

> The instructions given as to how to go about when doing these practical tasks should be clear and understandable (7 M3A G3).
>
> Practical tasks have to be simple and easy to understand and not tricky (7 M3B G6).

Recommendations encouraged the integration of everyday life experiences into questions, as such experiences were perceived to make questions easier, as reflected in the excerpt below.

> I think it was good because we are asked about things that we see day in and day out so we are able to answer with ease. At the same time it was difficult because we answered and we got it wrong. The teachers expect answers that are too scientific yet we just give simple answers as we see the things occurring (7 M4 G2).

Seemingly, there was an awareness among some students that their answers needed to use more scientific knowledge, even if task questions were linked to familiar everyday life situations.

b) Task importance and value

The data in Table 4.11 show that thirty citations were made recommending the continued use of performance assessment tasks due to its importance and value for meaningful learning. Such recommendations were based on the perceived role of performance assessment task in improving students' acquisition of knowledge and understanding, retention and recall of information, as well as the development of appropriate skills.

i) Improved learning and understanding

Students recognised that the tasks provided a variety of ways in which they could interact with the subject and present information. The use of performance tasks seemed to produce rewarding learning experiences for students, who recommended their use in other schools. Students' recommendations were presented in excerpts such as those provided below.

> We would recommend that these things (practicals) are done continuously even in the future because they help us learn simpler, better and faster (7 M4 G2).
>
> Give us the performance tasks so that we can understand what the topic is all about (7 M2A G6).
>
> They are fine but they should be enforced in most schools to help us as students understand better (8 M3A G5).

Students perceived the tasks to simplify the learning and understanding of key concepts in science topics.

ii) Retention and recall

Improved learning and understanding lead to improved retention and recall. Some recommendations to continue using performance assessment tasks as part of regular assessment in Science were justified by a perceived improved retention and recall of learned material.

> We prefer practicals test tasks because they help us to recall what we did during exam time (7 M4 G1).
>
> You get a chance of doing the practicals yourself in that way you won't forget in future or it's not easy to forget (7 M3A G8).

The tasks were perceived as opportunities for students to engage fully with the science content and procedures. They enabled students to engage in deeper learning of the subject and therefore the ease with which they recalled information in other assessment experiences.

iii) Skills for the future

Recommendations for continued use of performance tasks were also based on perceived long term benefits to the students. The development of science practical and procedural skills for use in future endeavours for students in all schools were another justification for the continued use of performance assessment tasks, as the excerpts below show:

> The practical tasks, mainly because of the experience, lets say you become less successful in life, this experience can make you something as you can go to the varsity or hospitals to work in the labs since you can handle the apparatus, you know the

> apparatus unlike when you are saturated with information but find yourself at the bus ranks or garden since may be you have no certificate. (No money for degree) (7 M3B G2).
>
> They should be applied in all schools so that all students will be something in future, do not rely on other people for knowledge, e.g. doctors. Every student should be educated for the benefit of her own country (8 M3B G6).

Students recommended extending the continued use of the tasks and therefore their benefits to students in other schools. Such a recommendation illustrates how appreciative students were and their desire for other schools to experience the use of the tasks.

Cognitive and psychomotor benefits seemed to be the basis for students' recommendations with regard to the importance and value of the tasks. The use of performance assessment tasks was recommended on the basis of their contribution to improve learning, understanding, retention and recall of scientific concepts, as well as the development of career skills.

c) Performance task resources

Time and equipment were perceived to be essential resources in preparing for and implementing performance assessment tasks. Students' concerns about their availability were raised in three questions as shown in Table 4.6 above. In line with these concerns, students recommended the provision of sufficient time and equipment for tasks.

More time seemed to be required for two purposes to allow for more and frequent performance tasks and to allow students to finish assigned tasks.

> They should provide many periods for practicals (7 M3A G4).
>
> The practicals tasks should be done more frequently (7 M3B G6).
>
> The time should be increased so that we are able to finish the practicals (7 M3A G2).

Students' recommendations regarding equipment were captured in excerpts such as the following:

> You should make sure that the apparatus are enough and they function properly (7 M3B G7).
>
> They should provide each and every student with his/her apparatus (7 M3A G4).

There is consistency in the perceived problems regarding opportunities for handling apparatus (PSYCHAN), concerns about resources and recommendations given by the students from School M3. This school had eight groups for each class because the classes were large. It had limited equipment as a resource for science teaching. Students in School M3 seemed in favour of building up equipment stock to allow continued use of performance assessment tasks. Students from School M1, another school with limited resources for science teaching, did not express similar ideas.

d) Group assessment

As far as peer collaboration and support were concerned, students recommended the continued use of group performance assessment.

> The use of performance tasks should proceed because as students *we* understand each other when we work together and this contribute to the knowledge of science (7 M2B G5).

Students found the sharing of ideas and the explanations given by fellow students favourable for their learning through performance assessment.

Negative experiences from group assessment led students to suggest and recommend other ways of administering performance assessment tasks. Some students seemed against the use of group assessment because of the way marks were allocated, as the excerpt below shows.

> Performance tasks are not good because group members give one point and when we are given marks they get the mark which they don't deserve (7 M2A G2).

The above statement does not state clearly that group performance assessment tasks should not be used, but implies that. The students felt very strongly about marks and were displeased about other students acquiring marks they had not worked for. Thus, other recommendations focused on strategies for grouping, monitoring and supervising students when administering performance assessment tasks.

> Those who don't talk should be separated and write on their own instead of just sitting and know at the end s/he will get marks (7 M2A g5).
>
> You should make sure that the people in the group are all concentrating and they all contribute (7 M3B G7).

Students' advice was to assign students to groups according to their levels of participation and seriousness. The perceived feasibility of these suggestions is revisited in Section III below on teacher perceptions about performance assessment.

e) Motivation

Students still suggested the continued use of performance assessment tasks even though they were unfamiliar to them. Excerpts that illustrated these ideas were:

> It was the best assessment we have ever experienced (8 M3B G5).
>
> They must continue using the performance tasks because it is easy for students to pass (7 M2B G1).

The continued use of performance assessment tasks was justified for reasons such as enjoyment and improved grades.

4.3.8 Summary of Section II

Described in this section of the chapter are various perceptions and experiences of students, regarding the use of hands-on performance assessment model in Science. Perceptions associated with each sub-category, category and themes are summarised below.

Cognitive disposition

Dealing with task characteristics involving

Task complexity: level of difficulty of the task.

Task importance and value:

- o learning and understanding, skills development, retention and recall, relevance of knowledge, assessment.

Task requirements:

- o pre-task instructional level (adequate exposure to and practice in practical activities, student understanding, attentiveness, concentration in lessons);
- o in-task administration level (understanding task; deep thinking and use of knowledge in new situations; practical, inquiry, procedural skills and their application; balanced knowledge competencies - theoretical and applicable);
- o readiness for task depended on prior instruction and experiences, familiarity with assessment model, task content, knowing in advance about carrying out tasks, expectation from the tasks;
- o task resources (time and equipment).

Metacognition

Dealing with self-assessment, reflection on own knowledge of science, confirming own knowledge against task data and information from peers and teachers.

Affective disposition

Dealing with

- extrinsic motivational aspect of improved grades and passing;
- intrinsic motivational aspects involving interest, improved knowledge from multiple assessment tasks.

Psychomotor disposition

Dealing with the development of practical, procedural, manipulative skills and process skills.

Social disposition (Group assessment)

Dealing with

Peer collaboration and support: sharing ideas, learning from and tutoring each other, improved understanding of task, and quality responses;

Group assessment problems: unfair sharing/allocation of marks, non-participation /dependence /dominance, conflicting points of view resulting in no answers.

Recommendations

Complementing perceived positive aspects of performance assessment and suggesting strategies for minimizing perceived negative aspects from different categories.

Students' perceptions in the different categories concurred. Perceptions that the tasks were easy were supported by perceptions on peer collaboration and support from group assessment. Students perceived tasks to be easy to understand due to their practical nature, collaborative interpretation of the task and generation of group quality answers. Quality answers improved grades and therefore chances of passing.

Perceived problems of performance assessment tasks were associated with degree of participation by group members and allocation of group scores. Problems of malfunctioning of equipment were also identified.

Some additional evidence regarding students' perceptions was also provided by the focus of their recommendations for the future implementation of performance assessment in Science at the secondary school level.

Section II provided a brief preview of teachers' perceptions of performance assessment in relation to students' perceptions. Teachers' perceptions of the use of performance assessment are discussed further in Section III.

SECTION III

This section of the thesis deals with the experiences and perception of the four participating teachers about the use of performance assessment in Science at the junior secondary school level. The four teachers were given an opportunity to experience administering performance assessment tasks and their rubrics from units on Electricity and Air and Living Things. These topics were taught at Form II/III (Grade 9/10) level in Swaziland. As mentioned in Chapter One, the teachers taught these topics using an applications-led science teaching approach, referred to as a contextualised teaching approach in this thesis. Teachers administered the tasks to their students and used rubrics as a guide to observe and grade students' performances according to criteria specified in the rubrics. Pre-task and post-task discussions were also held with the teachers to discuss how to administer and score the performances in the tasks. Some in-task consultations also took place. In addition to discussing the scoring of students' performances in the tasks, observations made by the teachers were also discussed.

4.4 TEACHERS' PERCEPTIONS OF PERFORMANCE ASSESSMENT

Perceptions of teachers are discussed under the sub-categories of task complexity, task importance and value, task resources, group assessment (peer collaboration and support, and group assessment problems), and affective disposition, as used in Section II above. Additional views are also discussed as they emerged from the interviews. The additional categories relate to teachers' perceptions on class size and teacher observations, the use of rubrics, teacher-student interaction, use of multiple assessment tasks and alternative ways of administering the performance assessment tasks.

Quotations are provided as examples to support and authenticate the identified perceptions and to enhance the validity of interpretations made about teachers' perceptions. Quotations from field notes are used to support teachers' assertions and to highlight other important aspects of performance assessment tasks and their administration.

4.4.1 Task complexity

Task difficulty was perceived in terms of pre-task instruction and degree of students' preparation for handling the tasks. Insufficient equipment and use of delicate equipment during lessons were perceived to contribute to inadequate preparation of students during lessons. The shortage of equipment led to the use of class or large group demonstrations for activities that involved delicate equipment or equipment in limited stock. Class demonstrations did not allow students to develop practical skills and sufficient knowledge of the use of laboratory equipment. For example, the performance assessment task on electrical resistance required students to assemble and use a voltmeter and an ammeter in a circuit to generate data and use the data to calculate the resistance of given wires. This meant that they needed to read the two instruments and perform the appropriate calculations. Students were observed by teachers to exhibit signs of confusion and uncertainty in doing the task. During this task students tried many permutations in arranging the two instruments in relation to the resistance wires. This uncertainty resulted in the use of more time than had been anticipated and therefore a perception by students that the tasks took a lot of time. When working on the task on testing for gases some groups of students wanted to use all the materials they had used when testing gases during lessons, without realising that limited samples of gases were available - two test tubes of each of three gases (see Appendix IIB). The teachers' recollections of these observations are captured in the excerpts below on their observations during the performance of the tasks as challenging for the students.

Inna: What I can say is that this activity in the actual fact is a bit demanding. Ya, it is a bit demanding. For instance the students have a tendency of forgetting to connect the voltmeter parallel to the wire, they remember that the voltmeter should not be connected in series, they were told that it should be parallel to the cells, should be connected parallel to the cell or to any resistance wire, but now what comes to their minds is the cell, forgetting that now we are finding the voltage passing through the wire, in that case I'm saying it is demanding. … And No 2, when doing this activity, in most cases we do a demonstration and then in a class as big as this one... then because of that when doing a demonstration, not everyone is paying attention, and then when it comes to the practical now it is then

> they begin to realize then that they don't know most of the things, like when they are connecting, you keep on reminding them, "you remember you are going to break the ammeter, if you connect for instance positive to negative you are connecting wrongly" just reminding them.
> And another thing again it's the reading. For instance a group would choose a scale from 0-15 only to find they need to take their readings from 0 to 5. They would use a wrong scale instead of the one that they have used (Inna post-task discussion).

Reflected in the excerpt above is the recognition of the negative effects of teaching conditions on students' preparation for the practical tasks. Another point is the importance of teacher input in guiding the students during the tasks to counter poor preparation. Such input from teachers was seen as helping reduce the difficulty of tasks and helping students to learn more information from the tasks. Similar concerns about poor visibility of class demonstrations were raised by students.

Further instances of student confusion observed during performance tasks were related to the selection of appropriate testing procedures for identifying the given gases. Students' knowledge of the tests for gases and their ability to use this knowledge in a live hands-on situation was also tested. Students' apparent confusion and implications on time available for the task were captured in Lorraine's observations as shown in the excerpt below.

> Lorraine: Other causes (*of students not finishing tasks on time)* like they tend to confuse ideas. So you first carry out practicals during the learning sessions, and when it comes to the test, the performance tasks they are asked to do something, they run to the conclusion using previous ideas not using what they are carrying out (Victoria: Ok). Take for example, are we following topic by topic, that Electricity, Air and life?
> Victoria: No I think we can just talk.
> Lorraine: … where-by we had to test for the type of gases in the performance task. Somewhere they've got to plan and then write down the materials that they are going to use. They included all the tests that had been done whilst they were learning yet in the test they were given, in my test, only three gases to find out and only two sets to use. But some of them would request for five reagents for testing. They would need some litmus paper, limewater, bicarbonate indicator, splints and so on, yet they had to use one for all three and the second one for all three, instead of asking for only two they asked for all that were used in the lesson (Teachers' FGI).

The observations made by the teachers in the above excerpts were based on content related examples. Students appeared confused about the orientation of the instruments used in circuit measurements and in selecting appropriate gas testing reagents and procedures. Students tried to recall as much information as possible that related to the tasks without selecting what was appropriate and necessary for carrying out the tasks. Support for the views that the tasks were challenging for the students and that students seemed somehow confused about how to do the tasks, is shown in the excerpt below that was taken from field notes.

> The planning took longer than anticipated – possibly because there were several trials of circuits - recalling or drafting of a number of the circuit diagrams they had been drawn during lessons, many of which were not suitable for the circuit required for determining the resistance of a current carrying wire (Field notes M4)

Students did not seem to find it easy to select the appropriate circuit models or gas testing reagents for the tasks.

In summarising perceptions about the complexity of tasks, one notes that teachers perceived the performance assessment tasks used in this study to be challenging and confusing to some students in that students tried to recall and use whatever they had encountered during lessons without proper selection of relevant information for the given task. Teachers linked task complexity to instructional preparation of the students. Thus, pre-task instruction was perceived important in preparing and training students for performance assessment tasks and for making tasks less difficult for students.

4.4.2 Task importance and value

Whilst the tasks were seen as demanding and confusing to some students, teachers felt that they also presented assessment opportunities for teachers and learning opportunities for the students. Performance assessment tasks enabled students to display skills like communication and manipulation of equipment allowing teachers to assess students' abilities regarding those skills.

> Inna: … Performance tasks create a free environment for the learners to show their skills, that is, communication, handling apparatus, drawing, etc. Through performance tasks a teacher is able to make assessment of the skills exhibited by the learners.

And later on said

> As I have said previously that one is able to assess or to find out, to assess their skills. We have talked about communication, which is a skill; we have also talked about the handling of apparatus which is also a skill. So

this type of test helps the teacher to be able to assess (Victoria: even those skills) those skills yes (Teachers' FGI).

Tasks were also perceived to promote students' learning by encouraging them to participate more actively and become more attentive during lesson activities and the tasks. Lorraine noted that students were compelled to participate in the assessment tasks, as seen in the excerpt below.

Lorraine: Another thing good about it is that most of the time, students had to carry out the practical themselves, finding out about the concepts rather than a teacher carrying out a demonstration only for them, so that they only observe. Cos at times you find that if you ask them to come around for observation some would be busy with (Inna: their own things) their own things, pricking others from behind. Only a few would benefit from them, yet in this approach where they are somehow forced to carry out the practical themselves most of them do the practical (Teachers' FGI).

The excerpt indicates that teachers seemed to associate students' motivation for carrying out the tasks with some form of coercion or obligation. The fact that the tasks were assessment activities seemed to be the motivating factor for students to take the tasks more seriously than they normally did in lesson activities. In response to ascertaining whether the observed improvement in participation was during lessons or during the performance tasks, Lorraine responded thus:

Lorraine: The use of performance tasks. In no way that one would sit back and not participate in the performance tasks. Maybe because they know that they are being assessed on that I'm not sure but somehow it helps them to take part (Victoria: Ok) (Teachers' FGI).

The improved motivation of students when they carried out the tasks observed by Lorraine was also noted by Josephine as shown in the excerpt below.

Josephine: Ya because they know it's a test they normally take it seriously. If it's just practical work in class the may not be as serious.
Victoria: What did you notice? Did you notice any difference?
Josephine: Yes there was a difference. They were actually more committed when they were doing the test, because they know that it will go with marks. They are usually more serious if it's a test than when it is an activity in class (Josephine Interview).

Increase in students' level of seriousness through the use of performance tasks was also observed by students. Another perceived contribution of performance assessment tasks to learning was justifying students' engagement in practical activities during lessons. Students were, therefore, not just doing the practical activities as part of lessons but they knew they would be assessed on them.

> Victoria: Do you think they [*performance assessment tasks*] contribute in any way to their learning of the subject matter?
> Josephine: I think they do.
> Victoria: Or do you think they have learned it during their preparation time and now we are just assessing them?
> Josephine: No, it helps them. I think it helps them because now they can see that what they have been doing all this time when they are experimenting and now they are being tested. It helps them I believe (Josephine Interview).

Although the view expressed by Josephine in the above excerpt did not indicate a direct link to the learning of subject content, it seemed to indicate a motivating factor for students' engagement with the tasks. This view concurs with that expressed by students from School M4 during their interview.

To summarise this sub-section it can be noted that teachers presented three main perceptions relating to the importance and value of performance assessment tasks. The tasks were perceived to be useful in allowing students to display abilities and skills, such as planning investigations, manipulating apparatus and communicating, thus enabling teachers to assess such competencies.

Teachers also identified the tasks to be useful in improving students' level of motivation as observed by the increase in their:

- participation and learning; and
- seriousness and commitment to the assessment tasks and lesson practical activities.

4.4.3 Performance task resources

Teachers acknowledged the need for providing more resources such as equipment, time, laboratory space and personnel for the successful implementation of the performance assessment tasks.

a) Time requirements

Time constraints imposed by the use of the performance assessment tasks were perceived to impact more on pre-task instruction. Teachers' submissions regarding time requirements tended to revolve around the contextualised teaching approach and less on the time taken for students to do the tasks. In the performance tasks on electricity teachers felt that students needed more time to finish the tasks. The link between the time requirements and the tasks per se was unclear, though the time available for Science in the schools' time-tables was perceived to be one reason for the shortage of time. This information was presented during a discussion of perceived weaknesses of using the tasks in the teachers' focus group interview.

> Jabulane: What I observed was that it was the timing. You find that the time is less in our case we have only two periods with our Form IIs, which is just one hour ten minutes.
> Victoria: One hour ten minutes a week. (Jabulane: no per period) Oh! per session … OK.
> Jabulane: It's two periods so you can't cover all the activities which must be done by the students per session.
> Victoria: So this time, is it time for teaching or for taking the practical?
> Jabulane: Even if it is time for taking the practical you cannot finish, so you need to actually borrow some time from others so that you can cover (Teachers' FGI).

Teachers felt that the contextualised teaching approach in which the performance assessment tasks were embedded required more time than was available. The perceived impact of the teaching approach and performance assessment tasks on the time available could be observed from the teachers' statements given in the excerpt below.

> Victoria: What about the performance assessment? ... I mean the use of performance assessments for assessing students... How do you feel about the time that you have spent?
> Inna: During the tests.
> Victoria: Yes taking the tests, giving students the tests.
> Inna: Well I think what we were referring to, I think generally we were saying that this method this contextualised method though good it consumed a lot of time. I don't know.
> Victoria: But the performance assessment tasks?
> Inna: The performance assessments no, during the test, no. The time is not that [*Lorraine goes on to discuss the teaching approach*].
> Victoria: I still want to go back to my performance assessments. (Lorraine and Inna: Ah! Ah!) Don't you feel that they are also taking a lot of time?
> All: No, no, no (Teachers' FGI).

While teachers felt that the tasks did not take long, they identified certain student factors that could account for the need for more time by students to complete a given task. These factors were confusion experienced by students, delays in starting on tasks and low students' confidence. The excerpt below illustrates these views.

Victoria: What do you think contributes to the students not being able to finish? Is it the time being limited or it's some other cause? …

Inna: Some groups take a long time to initiate their work. They sit you know and read as if now they are meeting something they have never met before. They kind of you know what can I say, they drag in such a way that most of the time is now wasted; they don't jump into their work and do it quickly. And then this mostly occurs more especially if now if there is only one student in the group, one or two or sometimes you know, they are kind of not, what can I say; they've got low intelligence. They just drag in such a way that one can think that they don't finish because you know that maybe the time was limited or what. The problem lies with them I can say.

Lorraine: They are not confident.

Inna: Ya they are not confident

Lorraine: At times they wait for the others to get started so that they can check (Inna agrees with Lorraine)

Inna: [*After checking the transcript she added*] Some groups take a long time to finish because they take a long time to start what they are supposed to do. The reason might be because they were not confident enough as they were doing this for the first time (Teachers' FGI).

Evident here is the effect of the uncertainty and confidence level on student initiation of the task, which caused delays and therefore a lack of time to finish some tasks. Waiting for others to start on the tasks also resulted in less time for working on the tasks. These observations do not seem to indicate that the tasks took long to finish, but rather that other student factors such as lack of familiarity with the task contributed to related problems.

In Section II above it was mentioned that there were students who recommended more performance assessment tasks. Josephine, however, felt that one task was enough because students needed a lot of preparation for the teacher to be certain that they would be able to respond to the tasks. The number of performance assessment tasks that could be given in a unit was perceived to be dependent on the pre-task instruction time.

Josephine: One [*performance assessment task*] is okay because there are so many things to be done in the syllabus. If we give them a lot of the tests and we delay a bit unless we make our own time outside the timetable and, ya I think they do need time because by the time we do them we must be very sure that the students have really been learning, they are grasping the ideas that you are teaching them well before you can actually give them. So I don't know. I think its, I don't know how I can put it but before you can give them the practical test you have to be really sure that they really know, … you do all the work that you should so that by the time they write they write it well. So the more you give them the more time you need to prepare for them.

Victoria: Okay. So that's where you are coming from. You need to prepare them (Josephine: … before the test and it needs time too …) Yes. So the more tests you give them the more preparations (Josephine: … you will need …)

Josephine: The only thing you could do if you need more time is to arrange for some other time besides the teaching time (Josephine Interview).

The submissions in the excerpts above indicate that teachers spent some time working on developing students' practical skills to enable them to do the tasks. Being aware that students would take some form of practical test seemed to influence pedagogical strategies teachers used to ensure that students developed the skills expected for the assessment tasks.

The summary of perceived time requirements on the use of performance assessment tasks is as follows:

- Time constraints on the use of performance assessment tasks were perceived from three perspectives namely, science teaching time allocation, student factors contributing to incompletion of the tasks and pre-task preparation of students. All perspectives were in favour of the view that the performance assessment tasks were less time consuming than the contextualised approach in which the tasks were embedded.

- The time allocated in the time table for teaching Science, and therefore, for the tasks was thought to be insufficient, thus some students could not finish the task.

- Students' inability to finish the tasks was also attributed to students' delays in starting to work on the tasks. The tasks were perceived by teachers as less time consuming provided students did not delay initiating their working on the tasks, were confident and familiar with the task procedures, and not confused by it.

- Preparing students to embark on performance tasks by developing their competencies during instruction took a lot of time. This was because of teacher awareness of the nature and criteria of the assessment tasks. Teachers found themselves spending more time on the teaching of skills and concepts required in the task to prepare students.

b) Equipment and class size

Securing adequate and proper equipment for implementing performance assessment was perceived to be a challenge for the teachers. The excerpt below illustrates teachers' perceptions of the situation in their schools regarding equipment and class size for implementing performance assessment.

> Jabulane: The equipment. You need to have proper equipment to do the activities. But if you don't have then you are in trouble…
>
> Inna: … We do need some apparatus and chemicals for this to be successful otherwise (inaudible). And another thing I can say we should have the right numbers of students per group. If the groups are big (Lorraine: agrees) [*sentence not completed*].
>
> Lorraine: And we always have large numbers in classes and if you try to make the groups smaller you are making the number bigger. Instead of having five, six groups you end up having eight, nine group (others agree)
>
> Jabulane: Ten or more than ten per class.
>
> Victoria: You mean reducing the number of students per group?
>
> Inna: Yes. We are saying that if the numbers are big per group, it becomes difficult, because you will find out (*sic*) that even if all the members of the group would like to try to manipulate the apparatus they can't due to the great number of pupils per group (Teachers' FGI).

The teachers recognised that student numbers and equipment quality (working condition) and quantity may interfere with their implementation of performance assessment. According to the teachers a large number of groups per class made it difficult for the teachers to monitor, help, observe and score students' work, as Jabulane noted in the excerpt below.

> *And the other thing,* I don't know; besides *there are too many of them* for, *to be followed at each stage of their performances.* They are too much. *Yet if they are in a few groups you can be able to monitor that now* they are doing that. (Jabulane post-task discussion: Translations from SiSwati are in *italics*).

It was a challenge to set up for the classes even for group performance assessment as the excerpt below shows. The observation presented in the excerpt further supports the perceived effect of class size and equipment on administering performance assessment tasks as observed above.

> The teacher already had six test tube racks on the work benches. I also briefed the teacher on what we were going to do in setting up. The teacher had already pinned up the list of group members (ranging from 6 to 10) for a combined class of about 57 pupils. Two Form III classes were combined for purposes of teaching Science because they were short of teachers in the school. I felt the groups were too large and thought of having 13 groups of 5 pupils per group but this meant 78 test tubes of gases and a lot of test tubes of gases to collect. This could not be done in the time available to set

> up. So I settled for 8 groups for which I had test tubes and rubber stoppers. I could not split the classes because that was the time the teacher had organised with the students (Field notes M1).

Teachers acknowledged that equipment requirements were task dependent. Some tasks, like the tasks on testing gases and electrical conductivity, required basic school laboratory apparatus and equipment that is usually available. But there was still a limit to the availability of the equipment for combined classes. The task on electrical resistance required voltmeters and ammeters that were in short supply - particularly in the non-INSET schools.

In summary the following perceptions were identified from teacher expressions and observations.

- Equipment and class size had a controlling effect on administering performance assessment tasks.
- Equipment availability determined the number of groups and therefore the group size for group assessment and opportunity for students to participate.
- Perceived lack of student participation by students may be due to large groups and lack of opportunities to participate.
- Large numbers of groups per class may reduce time available for the teacher to monitor, interact, guide and score students' work.
- Effects of equipment shortages may be task dependent. Some equipment is used for specific content areas (ammeters and voltmeters), some are delicate and in short supply while others are basic and are used for several content areas (glassware) and are usually part of the fundamental laboratory equipment.

4.4.4 Affective dispositions

Besides the tasks being perceived to contribute to improved student participation, the tasks were also perceived to improve students' interest in Science.

The teachers indicated that they observed students displaying a high level of interest in the tasks. Students were observed to be enthusiastic and keen to take on the tasks as noted by Inna in the excerpt below. Teachers expressed the view that students were motivated to work on the tasks during a discussion of students' preparedness and readiness for tackling the tasks.

Lorraine: I think they were (Victoria: were they ready? Were they prepared for this?) They were. …
Jabulane: Ey! Mine, they were not. In most cases they were not. …
Inna: I don't know how can I say they were ready or not? I don't know but they were enthusiastic to see what they are expected to do. They wanted to do the practicals. In that way I can say that they were ready (Teachers' FGI).

Later on in the interview while discussing the future use of performance assessment in Science teachers seemed enthusiastic about the continued use of the tasks because of their contributions to increase

ing students' interest. The excerpt below demonstrates teachers' enthusiasm towards the use of performance assessment tasks.

Lorraine: Ya we should continue.
Inna: How can we assess their skill you know, without doing this? (Mmm) It's very interesting.
Victoria: You found it interesting?
Inna: Very interesting. More especially if you've got enough apparatus if apparatus are enough. And my kids also kind of liked the tasks, more especially the second one; they were waiting for your coming, in the second one. The first one no, they were ready but they did not know what kind of thing they were going to meet. So, with the second one *(sentence not completed*) (Teachers' FGI).

Jabulane and Lorraine also indicated to have observed an increased enthusiasm for laboratory work in students. It was not clear, however, if the interest was due to the context-based teaching or the performance assessment tasks or both.

Varied perceptions by the teachers regarding student readiness for the task were evident. However, the only explained readiness was psychological readiness. No elaboration on conceptual or skills related readiness was given by the teachers, neither was it explored by the interview. The non-committal responses regarding content and skills readiness could be linked to the perception that pre-task instruction might not have been adequate for the students to be ready to meet the cognitive demands of the tasks.

Motivation to work on tasks with greater effort was improved by the allocation of marks. According to Josephine, in order to sustain students motivation, the marks given for tasks needed to be reported in a way that the students would be able to see the tasks as important for their school work and therefore grant the tasks the seriousness they deserved.

> Victoria: I was wondering about just keeping it as practical test marks and not computing it together with the (Jo: the students real mark) student's real mark.
> Josephine: But that means the practical test won't be a test anymore.
> Victoria: No it will reflect there that in a practical test the group in which this student is in got this grade. But this grade is not really put in the report. I mean it's not really added to the marks of that individual student.
> Josephine: But they won't take the test seriously (Victoria: the practical seriously) they won't. If it won't be reflected in their report they won't take it seriously (Josephine Interview).

Emerging from this excerpt is the idea that if marks for performance assessment tasks were not used in the final mark or grade, then the students were less likely to put effort when working on the assessment tasks. The desire to reflect the marks in students' reports was clear. However, it was countered by worrying that marks were not fair to a student working with low performing students as, a low mark could spoil the students' term-reports by lowering the overall Science. The excerpt below illustrates Josephine's concerns.

> … but it also has a disadvantage for the good student. You find that her own participation wasn't taken so seriously and so the group as a whole is not doing very well and so the child's mark is lowered (Josephine Interview).

Josephine was however, against segregating students on merit. She felt that such a move was likely to be psychologically distressing and more detrimental for students of low-ability or who were shy. She expressed her concerns as follows:

> Josephine: You mean putting the good ones in one group? I think that would make it worse.
> Victoria: How?
> Josephine: Because it means that the slow ones will not have anyone to help them, they would just be worse. When they have someone who is good, somehow it helps them. When they are not doing well they know. Like they know each other, that if they are just the ones that don't perform well. Because I remember there is a group that was doing very badly that side [*pointing to desk at back of lab where the group worked*] and they just didn't do it well, 'cause I think in their minds they know it's just us, and on the other hand most of the good ones were here and there [*pointing to front desks*] (Josephine Interview).

Josephine recognised the positive effect high performing students had on low performing students, tutoring them and giving them hope.

Performance tasks were perceived by teachers to increase students' motivation and interest in Science, as well as in class practicals and assessment tasks. Students were observed to be interested and enthusiastic in attending laboratory sessions and the assessment tasks.

Marks and students' achievement were perceived important for encouraging serious participation of students in the tasks. Furthermore, there was a perception that marks for high-ability students may be lowered by low task marks resulting from working with a group of poor performing students. The lowering of students' marks was seen as unfair on affected students and that it could spoil their report cards. Unfortunately, grouping students working on performance tasks on merit, was not seen as a solution, although such a suggestion was made by students.

4.4.5 Group assessment

As mentioned earlier, a group assessment approach was used to administer performance assessment tasks. Similar perceptions to those of students discussed in section 4.3.4 above, emerged. Teachers also identified peer collaboration and support, and group assessment problems as important aspects of group performance assessment.

a) Peer collaboration and support (PCS)

As noted above using heterogeneous groups based on merit was perceived useful for low-ability students. Low-ability students benefited from peer tutoring by high-ability students.

Teachers also recognised that students developed good social relationships during group performance assessment. The excerpt below shows a teacher's perception relating to PCS among students.

> Jabulane: And also the co-operation they need to co-operate as they are doing the activity so there is some communication skills in that… they co-operate on that because somebody will be saying lets do that and then somebody will ask why and then give the reasons (Teachers' FGI).

Students co-operated with one another and improved their oral communication skills as they explained their points of view to their peers. Students also helped each other. Worth noting was that Josephine's view seemed to focus on the high-ability students helping the low-ability students to improve their understanding of the task, learning of subject content and acquiring of skills.

b) Group assessment problems

Group assessment problems perceived by teachers were associated with dominance by some students, dependence on other students and the allocating of marks to non-working students or students who were perceived to be less-knowledgeable about the task or content. Students were also unhappy about these issues.

Teachers observed that students' home background played some part in the students' confidence in working on the task and in dominating other students in the group. The dominating students were thought to have had exposure to science related out-of-class experiences, such as electrical activities. Some students were observed to lead the group and to do all the circuit constructions and connections with less consideration for the students who were not so familiar with electrical activities or who were less knowledgeable. In the following excerpt, Lorraine shares her observation of confident students in their groups.

> Lorraine: Like in electricity, those who are usually involved in, what shall I say, motor mechanics at home … you find that they have some small workshops in their residential places and students would be involved, that is, the children around would be involved when the father is doing some of these electrical things in cars or other appliances. So you find that they do have an idea and as a result those [*students*] would tend to lead or somehow challenge the teacher.
>
> Victoria: Does this happen even in the performance assessment tasks?
>
> Lorraine: Yes, most of the time these students are the ones who do the wiring (Teachers' FGI).

The enthusiasm shown by the students who possessed background knowledge that was relevant to the performance tasks was perceived by teachers as a challenge. Some students perceived this as domination by these "knowledgeable" students and the teachers recognised this drawback as well.

> Inna: Ya they do challenge the teacher (Victoria: Ok) Or lead the other members of their group. Ya and tell them how things are done.

Lorraine: Like in wiring circuits those tend to be much more faster.
Victoria: So wouldn't you say that those tend to dominate the group?
Teachers: Somehow they do, somehow they do.
Lorraine: Because they are the ones who would be very vocal, challenging the teacher in some of these, so most of them would tend to withdraw (Teachers' FGI).

A noticeable issue from the above excerpts is that the expertise displayed by some students may have led to the intimidation and withdrawal of other students from making contributions.

Students' high confidence was also observed when the tasks were administered. Some students simply overlooked the instructions and went on to connect the circuits.

> Many groups started with connecting the circuit even before reading the instructions or discussing and planning how they were going to test the electrical conductivity of the six objects (Field notes M3).

While some students were perceived to challenge teachers and dominate other students in their groups, others were perceived as participating at an inappropriately low level, depending on the other students, as expressed in the excerpt below.

> When it comes to the practical assessment (Victoria: yes) you find that in a group not everyone is participating you find that maybe one or two students who are very good in Science and they'll be the ones taking an active role. And in fact the other students would tend to respect those and sort of leave everything to them, like "this one is good so let's leave her to do everything". And so you find the other people more-or-less observing because they've trusted those people, probably, and participation sometimes is not so good (Josephine Interview).

The observations by the teachers reflected in the excerpts above, reveal two extreme cases of differentiated student participation during group performance assessment. At one extreme, certain students took on dominant roles and performed the task with little input from other students in the group. At the other extreme, there were students who seemed to be completely dependent on a few more knowledgeable students. It was difficult for teachers to distinguish whether non-participation was due to lack of knowledge and skill, or due to domination by others and therefore a lack of opportunity to take part. In all three instances, there was some perceived unfairness of student behaviour on other students. Whilst teachers recognised that some of the students did not deserve the marks, they had difficulty deciding how to differentiate the awarding of the grades.

Victoria: And the giving of the same mark to all members of the group, what do you think of that?
Inna: Some do not deserve the marks I think
Jabulane: No, they don't …
Victoria: Some do not deserve the marks?
Inna: I think no …
Jabulane: You may find that even in the free participation they are not involved. You find that they are the "push type", they are being pushed
Inna: [*After checking the interview transcript she added*] It is difficult to come up with a specific response here because … though they might not be manipulating the apparatus they do you know contribute by giving ideas and suggestions of how particular things should be handled and be done. So in that way though we are saying not all of them deserve the marks, do we mean that they should not be rewarded for what they have contributed? (Teachers' FGI).

Inna's view seemed to be that a student's effort of whatever nature or degree needs to be recognised in the way that students understand: the awarding of grades.

There were also concerns about the accuracy of group scores in representing student performance. The interview excerpt below illustrates a concern for the lowering of grades for high-ability students due to a lack of consensus of ideas and the raising of grades for low-ability students.

Victoria: So you want to comment on the group work.
Josephine: On the group work. That is, sometimes it doesn't really give a true reflection of the child's performance (Victoria: okay) because you find that like I said there are these active people and there are those who are sort of a little bit behind but at the end they will share the same mark. And sometimes you find that a good student doesn't participate a lot and so the performance of the group is not so good and they will share those low marks.
Victoria: So a good student may decide not to participate?
Josephine: She's part of a group that is not performing very well. And sometime they don't really reach an agreement when they are doing this so they just do the task for the sake of doing it. Someone may decide lets do this, this way and you find that a child whose normally performing very well alone is sort of dropped, and the other way round also happens. A child who is not performing very well because she was working with people who are doing very well may end up getting a mark they normally do not get (Josephine Interview).

Awarding the same score for each group member was perceived to be misrepresenting the students' achievement levels.

In summary, it can be noted that group assessment presented both good aspects and drawbacks that teachers had to deal with. The main perceptions teachers conveyed were that:

Through peer collaboration and support:

- Students can learn from and support each other through cooperation and communication in group assessment activities.
- Students can learn from one another and should not be grouped on merit.

Group assessment problems:

- Presents problems and difficulties for teachers to observe and score groups efficiently due to variations in participation levels, namely:
    - dominance: where certain students take over and do everything without sharing the responsibilities of the task;
    - dependence: where some students take advantage of willing students and sit back and watch them work; and
    - inability to participate although willing because of limited workspace, equipment or dominance.

- Presents problems and difficulties for teachers in interpreting group scores where:
    - undeserving students received marks;
    - good students received low marks because they were part of an underperforming group;
    - low performing students' marks were boosted by working with high performing students.

- May involve conflicting and non-resolution of arguments during group discussion and a perceived non-beneficial effect of such arguments on students' learning and scores.

Although there were perceived problems associated with the use of group performance assessment, it was still perceived to be helpful in assessment supported learning and in administering tasks in situations of insufficient equipment.

4.4.6 Class size and teacher observation

The number of students in Form II ranged from 39 to 52, and 26 to 49 in Form III. Equipment available and student numbers per class determined the group sizes and the number of groups. Six to eight groups were used in this study. Teachers stated that they experienced difficulties in observing adequately all the students in the groups due to the number of groups per class.

> Victoria: Ok. That was on the performance tasks during the test. So it's good that it motivates them to participate (Teachers: Mmm) (Inna: Ya). Did you observe all the students; I mean most of the students indeed participating in the activities of the test when you were administering the test?
>
> Inna: eeeya not all of them (Jabulane: but most of them) but most of them.
>
> Lorraine: It was difficult to observe all that because of the numbers (Jabulane: Ya.) So if you have seven groups per session it becomes difficult for one to note in all of them that every student in the group is actually carrying out the tasks. (Victoria: Ok).
>
> Inna: And what I can say is that because of the numbers as you have already said you know that in this particular group, you know so and so is good at handling the apparatus. Ya. You can be able to spot, can see that three out of five maybe are good, not all of them. (Victoria: Ok) Due to the large numbers of learners in each group it was sometimes difficult to assess their individual skills (Teachers' FGI).

The teachers indicated that they found it challenging to assess students' skills individually as it was difficult to observe all the students in a group. However, Inna noted that teachers could strategically observe some of the students.

Similar experiences and projected challenges were also noted by Josephine. Her concerns were in regard to following the detailed scoring criteria in the rubrics to observe and score students while they carried out Stage 2 of the tasks.

> Josephine: It becomes a bit difficult with one teacher or two and there are six groups if we have to follow all these details properly.
>
> Victoria: These details in the scoring guide?
>
> Josephine: Yes the scoring guide… So for the teachers to really assess well and follow the details of the scoring guide, I feel the need for a group to start, although it will take a lot of time, one group comes in they do it the teacher assesses and checks everything, the same teacher. Another group comes in same teacher (Josephine Interview).

The use of highly detailed rubrics required the teachers to spend time with a group to see if they met the procedures described in the rubrics and allow grading. For Josephine the best way was to work with one group at a time.

Observations made during the administration of the task illustrated further these concerns as shown in the excerpt below.

> Challenges experienced were that
>
> - group plans were not adequately checked for accuracy, students were just given the testing material without ascertaining whether they had indicated them in their plans or not;
> - groups asked for different things at different times despite being asked to include everything they needed in the initial request - a sign of inadequate planning;
> - groups could not be given the appropriate attention and monitoring they needed;
> - the laboratory was crammed without much space for teacher movements between the groups (Field notes M1).

The excerpt above shows possible effects of large class size on administering performance assessment tasks – teachers not being able to check students' plans properly, time delays in getting material for groups at different times, difficulty of attending to students' needs in understanding or performing the tasks. Further complications were perceived to be due to the use of a detailed scoring guide. Following the detailed criteria in the rubrics delayed the teachers' progress in observing students working on the tasks, so that teachers could not fairly score students skills and competencies as required. Josephine's concerns are shown in the excerpt below.

> It was hard for one teacher to go round the groups. Like if you want to check this, say they have to be careful how they open the test tube carefully you find they've already opened it and they have already tested. They find the lime water turned milky, you didn't see them maybe you were watching the other group. You check the other group you didn't see the other groups. So by the time you go to them they have already done that part (Josephine Interview).

Problems of scoring arose in instances where students needed assistance while the teacher was still busy with another group of students. In some instances, as noted in the excerpt above, students could not wait for the teacher to finish working with one group but went on to perform the tests or their investigations without the teacher approving their plan or observing them carry out the tests or making observations. Students also expressed concern about waiting for the teacher to come and observe them test the electrical conductivity of the material or perform gas tests, or read from the ammeter or voltmeter, because they felt it was a waste of time. The teachers' experiences demonstrate difficulties encountered in observing and scoring the students' handling of apparatus and demonstration of science procedures and skills, which affected the scoring of such performances by the teachers. Further difficulty was due to the non-synchronisation of the distribution of material (gas testing reagents) on request and scorer observation. However, such synchronisation required more time for the tasks, particularly if there were many groups to be checked.

To summarise the teachers' experiences regarding their observation of students it can be noted that teachers found it challenging to take on the role of mentor and assessor - even though group assessment was used to administer the tasks.

Class size seemed to affect the scoring of student performances. Assisting students during their performance tasks caused delays for some students' progress due to having to wait for the teacher to complete working with one group. Students adversely affected were those who were able to produce their plans fairly quickly and proceeded to test them without being observed by the teacher. As a result teachers experienced difficulties in executing that crucial step in using performance assessment, observing and scoring student performances.

4.4.7 Using rubrics

The teachers found pre-task and post-task discussions useful for using the rubrics to focus their observations on the students' actions as they worked on a task and for grading students' competencies on the task. The use of the rubrics throughout the task also allowed the teachers to make notes on their observations so that the grading of the final report became easier, as Jabulane notes below.

> Jabulane: It does help in a way. If I remember well you get, some of the points you get them as they are doing the activity, then you come and sit down for the marking, you are marking the other things it becomes simpler, it becomes simpler even with the marking for the teacher …
>
> Lorraine: It was because once you look at the plan err you, it would be easier for you to allocate the rest of the marks in the scoring plan (Jabulane agrees) A group which could not plan well did not know what to do until the teacher gives them help-so it becomes easier to tell their scores as they proceeded (Teachers' FGI).

It was possible for the teachers to use the rubrics to assign scores even after helping the students by withholding marks intended for the corresponding item that involved direct assistance from the teachers. Examples of direct assistance included specifying constructing a table of results with appropriate entries as opposed to asking students how they would record their results. However, the decision was not always an easy one as noted in Section 4.4.8 below, on teacher and student consultations.

A different view was that the detailed scoring guide was useful to some extent except that the amount of detail in the rubrics used for scoring students carrying out the testing stage of the task might have been excessive. In discussing her experiences of the scoring guide and its usefulness in guiding the teachers, Josephine shared the following views:

> Josephine: It [*scoring rubric*] is useful except that I got a bit mixed up somehow. I had to put the marks, to be specific that describing what to do for each gas, open the test tube … Oh is it three, 1½ for each gas, ya. I think it's a good idea to be specific the way it is. I found it easier to assess them because the guide was there and how to allocate the marks. I don't know whether it's too much to expect all the details like that they should open the test tube carefully.
>
> Victoria: But that is on the planning. (Okay) Yes that was on the planning that is they are planning how to do it. So they are basically giving a kind of procedure or steps that these are the things they are going to do this way. Anyway this is just one particular one. Your point is that the amount of detail that you have.
>
> Josephine: It becomes a bit much, you find that they don't really follow it, probably maybe because when we teach them we don't tell them to "open it carefully" we just tell them to put lime water, insert a glowing splint. *They too,* they think you just open and put lime water. So maybe those details (Victoria: they may not be …) probably as teachers we should stress them so that when they do them they are careful to follow….
>
> Victoria: Anyway I'm not explaining this, but I'm on its use by a teacher, the experiences of a teacher in using it.
>
> Josephine: It's good, it looks - its organised it helps the teacher to assess very well, because it is divided for each of the parts (Josephine Interview).

The format of the rubrics and detail of mark allocation were appreciated by the teachers. Unfortunately, instructional practice was perceived to marginalise the procedural details outlined in the rubrics. Josephine's view seems to suggest that their instructional practices tended to omit such procedural detail and that teachers need to be encouraged to incorporate procedural detail in developing students' laboratory skills. Whilst she seemed concerned about the amount of details in the rubrics, her views concurred with an earlier assertion that she spent time getting students ready for the tasks.

Josephine made submissions that indicated that although students may have been required to use data from the testing stage to respond to task questions she accepted responses that did not show a direct link to the data, as long as they looked acceptable. An analysis of students' task scripts for the task on testing gases showed that some groups produced good plans but did not indicate what their observations were. Three of the six groups from her school had no record of observations at all but they identified the gases correctly and answered the other questions well. She scored the correct identification of the gases and the questions independently of the data section, on the grounds that students' plans reflected what they would do and what the expected observations of those actions would be. The students were not awarded the marks for the data. Two other groups had tables of results, with one group showing all the information (tests conducted, expected observations, actual observations) while the other group had very scanty information recorded in the table of results. The sixth group recorded their observations in point-list form. From these experiences Josephine was not convinced that it was necessary to expect students to record their observations in table format (as expected by the rubrics in this study). Her view was that the students should be allowed to record their observation in whatever way they were comfortable with, as she advised in the excerpt below.

> And then the other thing I wanted to say, we shouldn't necessarily let the students write the results in a table form because sometimes we stressed that… They can write it any other way as long as it is clear (Josephine Interview).

Josephine thus modified the rubrics to fit the situation in her class. Further support for Josephine's point of not insisting on a data table at the planning stage, is provided by the reflection notes in the excerpt below.

> Group 6 conducted the tests without deciding how or if they needed to record their observation, but once they had performed the test they then came up with a method of recording their observation (Field notes M4).

Indeed stressing a particular format, the table, for recording data delayed students' progress on the tasks. Many groups experienced difficulties producing a table for their results. They appeared uncertain about how to structure and label the data table. The use of tables to record observations was stressed in order to encourage students to think about the variables they would work on before generating data and to report these in a systematic way. It was also useful for the teachers to see if the students' plans included the necessary variables. The table could also serve as a part of the summary of the plan. Many groups were omitting the formal table. The demand for a table of results was relaxed in the third task on gases as students were taking some time to come up with tables for the results. They were allowed to do the investigation and then decide how to record their data.

In summary the following teacher perceptions on the implementation of rubrics to observe and score students' performance were identified.

- The teachers found the pre-task, in-task and post-task discussion useful for focusing their observations on the students' performances as they worked on the task and grading of the entire task.
- Teachers perceived the rubrics to be organized and useful in guiding and scoring student actions. However, sometimes students' actions did not match the criteria in the rubrics, making it difficult for teachers to check student performances against the rubric criteria.
- Teachers questioned the necessity of having in-depth procedural details as criteria, while acknowledging a possible need to modify pre-task instruction to match the criteria in the rubrics.
- Teachers felt it was not necessary to insist that students produce a table of results in their plans as long as the students achieved what they had set out to do. They recommended some allowance of teacher modifications in the use of the rubrics.

4.4.8 Teacher-student consultations during performance tasks

As mentioned in Chapter Three the administration of the performance assessment tasks in groups allowed students to interact with each other. Students could also consult with their teachers. The teachers could raise questions with the students in order to redirect them towards meaningful plans and investigations through those plans.

Students were observed to be taking advantage of these consultations with teachers. Some students seemed to depend on the teacher for every step of the task and this posed problems for teachers in grading performances by students. While these students attempted to ensure they did the “right thing” as approved by the teacher, the teachers’ understanding of the requirements of testing conditions presented conflicting emotions, particularly with regard to the “scoring” of students’ work. In discussing their views concerning the coaching of students by the teachers during the performance tasks, teachers shared the following experiences:

> Lorraine: Hoo! That gave me some problems in allocating points. Because with one group you’d find that they would not go a step further having not consulted with the teacher. So in that way you look at the points, even where they could have done this they are just not confident that they could proceed. Maybe they, just because it was a test they thought they had to be correct in everything, you know. (Inna: Ya) But you note that they do one step, they want the teacher to come, they move on to the next step they want the teacher to come, they do one step they want the teacher to come as a result they ran short of time. Although they were prepared in that during the teaching sessions they would carry out practicals whether in larger groups or smaller groups. But then there was that feeling in them, they were not free. We couldn’t make them feel very free that they can carry out the practical without having to call on the teacher now and then, “Teacher! Teacher!” you find the teacher moving up and down to that group then to that group. In fact with one teacher carrying out the assessment during practicals it becomes very difficult. It becomes very difficult. (Inna: Agrees) My groups, they are not confident. I don’t know why.
>
> Victoria: Ok. Any additions? I think that was the experience of one school.
>
> Inna: Ya the tendency is that they would call us as teachers. Sometimes personally I would observe that maybe they are not approaching whatever they are supposed to correctly. Then I would pose a question which somehow is going to help them. Then they would say “Oh Okay!” then they would do the correct thing. But as far as awarding marks, it wasn’t a problem because I would just subtract a mark if now there is something that you know (sentence not completed)…
>
> Jabulane: In my case it wasn’t all that bad but some would actually call, some would not until they actually finished. So that was something which, and you find that most of those who would not call me are the ones who would get more marks. …
>
> Victoria: How did they feel about that [*marks being deducted*]?
>
> Inna: Well they were feeling, I remember in my case they would call you and say you [*Inna*] keep on writing something once we call you, we won’t call you we will call the other madam to come and help us. I think you must remember that they also kept calling you to come.
>
> Victoria: … In your case?
>
> Lorraine: They would call me now and then.
>
> Victoria: Were they aware that you were (Lorraine: Ya) that you were taking away some credit? (Teachers’ FGI).

The experiences from Inna and Lorraine’s classes show how teachers perceived students’ needs for teacher support during the assessment tasks. From the teachers’ perspectives students wanted to do things in a way that would guarantee that they got as many marks as possible, and were going to use the teachers to reach that goal. In Lorraine’s situation the students’ behaviour was ensuring that their plan and procedure were acceptable to the teacher, and they were willing to trade off some marks for the teacher’s guidance. In Inna’s case students wanted assistance but did not appreciate the deduction of marks for contributions made by the teacher and they made it known to the teacher. Teachers attributed the students’ behaviour to low student confidence and possible anxiety. It is, however, encouraging that students made use of the teacher to chart their course in conducting the tasks.

Even though teachers had to deduct marks for specific help given to the students, they still felt that their presence and assistance was useful for the students’ learning, and certainly preferable to no assistance at all.

> Inna: One thing I like is that we are there to help them its not just that we leave them on their own until they are confused, we help them.
>
> Victoria: So this approach of giving them performance assessment tasks is better than the approach used for the final exam, where they are just given the ….
>
> Inna: And without anyone guiding them, here we are there with them, sometimes I go around and subtract marks from them and tell them their mistakes like… “Can’t you see you are using a scale from 0-15 instead of 0-3?” and the students tend to see their mistake then they realize their mistake (Inna post-task discussion).

Teachers anticipated that the behaviour of dependence on teachers by students would change. Students were expected to get used to the assessment approach and therefore improve their confidence in working on the tasks if the use of the tasks would continue.

There was, however, a perceived need for adequate equipment to develop familiarity with the tasks and students’ practical and procedural skills in order to improve student preparation for, and confidence in, the performance assessment tasks. Teachers also needed prolonged exposure to the use of tasks to be more confident in using them.

To summarise, teachers were expected to assist students while they carried out the tasks. While teachers felt that their presence was a good strategy that ensured that students received the necessary help with the tasks, they felt that students took excessive advantage of the opportunity to consult the teacher.

The perceived dependence on teachers was expected to decrease with continued use of the tasks. This seemed to indicate that teachers were linking the dependence on teachers to students' familiarity with the assessment format.

Teachers experienced conflict in their dual roles of mentoring students as they worked on a task and of grading such assisted work, particularly in instances where students consulted a lot.

4.4.9 Use of multiple assessment tasks

Normal assessment practices used in Science in schools in Swaziland involve giving students at least one formal assessment per month that is used to compile their continuous assessment scores. This test is usually the topic test, unless the topic takes more than a month, in which case the students would write their test based on part of a topic. In this study students were given three assessment activities in the Electricity unit and two in the Air and Living Things unit. Teachers' perceived the use of multiple assessment activities to provide a balanced assessment of student competencies and to cater for varied assessment preferences by students. Multiple tasks were also seen to give students opportunities to make up for low scores. Below is an excerpt that illustrates the above interpretation.

Victoria: Earlier on you said we should continue with the use of performance tasks and you felt they were not taking too much time, and maybe it would depend on how many they are. In this particular case we had a maximum of two in a topic. (Inaudible)... So using many ways of assessing students, how do you feel about that?

Inna: Ya.

Lorraine: I think it is good. Like she said some students would perform better in these tasks compared to their theory part (Inna: the theory part). So it's somehow balancing up the students skills. (Inna: Yes.) So if the students have not performed well in the theory part these would make up (Inna: Ya, ya) for them. So in the end the student would have passed, which is very encouraging.

Inna: In actual fact with one assessment a student may get a zero and then that can imply that the students know nothing, whereas that is not correct. With this number of assessments you kind of as a teacher get the different skills from each pupil.

Jabulane: I think she has just said it all because that is the main thing. When you do these parts then you check whether they are part of this whole, you are able to grade that one until the whole. Then you actually find out from the whole test that this one didn't do well here but she was able to do well on that one. So when you add the three then you come up with something (Teachers' FGI).

Teachers' concerns seemed to have been with regard to grades for students' work and the opportunities for possible improvement of such grades. Josephine also agreed that multiple ways of assessing students could be advantageous for students. Her arguments, however, seemed to favour the usual theory-based end of topic tests. Discussing the use of multiple assessment tasks indicated that Josephine felt it might be an over-assessment of a single topic at the cost of valuable time.

> Victoria: How do you feel about the use of these several tasks for a given topic as part of the assessment of the students? They use context-based test and the performance tasks per topic to assess pupils for a given topic. Do you feel it's too much, do you feel its okay?
>
> Josephine: I think with time it, we would probably get used to it. But for now I think it takes a lot of time, having to prepare the test if you have to prepare the test and the requirements, they do it, that's time taken. Then you give the test at the end. It's like assessing them on the same thing, on the same topic but in different ways. So it looks like its testing on one point. But at some point it helps those that maybe didn't do well the one part maybe the practical test, then they can do better in the (Victoria: the end of topic) yes, test. Otherwise I think its, the only problem is also, these are good ideas, the only problem is that looking at the setup we are in I think it takes a lot of time.
>
> Victoria: The set-up we are in being?
>
> Josephine: I mean like in the school setting where we only have six periods per week and one period is thirty five minutes so to cover the topic and to assess them in these different ways it takes a bit of time. You would have to arrange for extra time if you want to make good progress (Josephine Interview).

Multiple assessment tasks were perceived to have a potential for allowing for a balanced assessment of student competencies, catering for different assessment preferences among students and allowing students to make up for low scores. However, there were perceptions that multiple tasks could pose a danger of over-assessing certain aspects of a topic, although this depended on the number of tasks used.

Preparing and setting up for the tasks and preparing students for multiple performance assessment tasks were perceived to be time consuming for the teacher and students.

The use of multiple assessment tasks, while advantageous, required more time than that available for Science in the school time table.

4.4.10 Summary of chapter

Experiences and perceptions about the use of performance assessment in Science by students and teachers reported in this chapter show some similarities and some variations. These experiences and perceptions about the hands-on performance assessment model and the mode of implementation of the performance tasks, are summarised below.

Participants perceived performance assessment in terms of:

- Task complexity: Tasks were perceived to have aspects of being easy or difficult. Students found the tasks easy because of the support of group members and teachers in interpreting and formulating answers to the tasks. The practical format of the tasks and their similarity to previous activities, also contributed to the students' perceived reduced difficulty of the task. Students felt that they were able to write what they saw from the testing stage of the task or real world contexts and had to recall less. Tasks were also found to be intellectually challenging requiring thinking and analysis. Students associated task difficulty to lack of familiarity with the assessment format and the level of understanding of task content. Tasks from the Electricity unit were perceived to be difficult, while the one from Air and Living Things was perceived to be easy, indicating a possible dependence of task complexity on the topic and task content.
- Task importance and value: The tasks were perceived to have both learning and assessment benefits. Students perceived the tasks to contribute to their intellectual and psychomotor development. Students, in addition, felt that the tasks helped them understand, consolidate and improve their retention of knowledge. The assessment tasks also helped students to develop science procedural skills and to learn from the sharing of ideas with group members. The integration of tasks with real life contexts made students aware of the uses of science in their environment and encouraged them to be more observant in their surroundings.
- Task requirements: Input from students and teachers were perceived to be important in the pre-task preparation of students for the tasks. Pre-task instruction involved adequate practice with practical activities, effort by teachers towards

developing students' knowledge and understanding, as well as manipulative skills. Students' input involved attentiveness and concentration during lessons. In-task requirements involved students' understanding of the tasks, balanced theoretical and procedural competencies and the use of knowledge and skills when working on the tasks.

- Task readiness: Students perceived the following to be important for them to be ready and prepared for the tasks: prior learning, task familiarity, expectation of task, studying, level of understanding of topic, as well as type of task.
- Metacognitive influence on students: Students perceived the tasks to encourage them to monitor their knowledge level against that of students and the outcome of the task.

Supporting the characteristics of the performance assessment tasks and their perceived effects on students' development were contributions associated with the use of group assessment to administer the tasks. Students perceived these contributions in terms of intellectual and social development through peer collaboration and support. However, they also perceived problems associated with differentiated participation and contributions by different students to the achievement of the task goals. Perceived differentiated participation involved low or non-participation, dependence and dominance by certain students. Differentiated participation impacted on perceptions relating to mark allocation and, to a small extent, learning benefits. It also caused problems for the teachers.

Students' recommendations for the use of performance assessment tasks addressed ways of minimising the perceived drawbacks of the tasks and their administration.

Teachers' perceptions about the performance assessment complemented students' perceptions. Teachers perceived the tasks to be:

- useful in assessing students' practical and social competencies;
- increasing students' participation, learning, and commitment to task and lesson activities;
- justifying the use of practical work;
- increasing student's interest in Science.

However, they were also perceived to:

- be challenging and confusing for students;
- require proper training and preparation of students to carry out the performance assessment tasks and to reduce students' confusion.

Teachers perceived the use of group assessment to administer performance assessment to:

- improve students learning from the support of other students although inconclusive arguments in groups may counter this perceived improvement;
- have problems of differentiated participation;
- have problems associated with mark allocation for differentiated participation by students and the representativeness of allocated marks for the ability of the students; and
- be associated with problems concerning appropriate resources and time.

Teachers perceived the use of rubrics designed in this study to score students' performances on tasks to:

- help them organize their observation, guidance and scoring of students' actions; and
- contain too much detail in the criteria to facilitate meaningful student action or to allow the teacher to identify those criteria.

Regarding student-teacher consultations during the performance tasks,

- students were found to take full advantage of the opportunity to get feedback from the teachers; and
- teachers perceived a conflict in their roles of mentor to students and scorer of students' performance.

The use of several assessment tasks in a topic or unit was perceived by teachers as beneficial in providing a balanced assessment of students' abilities and catering for different assessment preferences among students. However, preparing students for the tasks and setting up for the tasks, had time related implications.

In the next chapter students' and teachers' perceptions are discussed in regard to the categories of task characteristics: task complexity, task importance and value, task requirements; metacognition; group assessment and the use of multiple assessment tasks. Relevant literature is also used to link the findings of this study to current studies in performance assessment and perceptions of students and teachers.

5. CHAPTER 5

DISCUSSION OF RESULTS I: PERCEPTIONS OF PERFORMANCE ASSESSMENT

5.1 INTRODUCTION

This chapter discusses the experiences and perceptions of students and teachers regarding the use of performance assessment in Science at secondary school level. It begins with a brief overview of the conceptual framework. This study treated the participating teachers and students as one case. Thus, views are reported and discussed as those of the participants irrespective of whether one teacher, all four teachers, or one group, or all groups of students raised the views.

The conceptual framework for this study is based on a view that assessment is undergoing a paradigm shift from a predominantly psychometric measurement model towards an educational assessment model (Gipps 1994). This view of assessment attempts to respond to conceptions of learning that are based on cognitive and constructivist models of learning. These conceptions of learning are calling for the use of alternative models of assessment where assessment is integrated with instruction to become part of the learning process, a concept referred to as assessment for learning (Gipps and Stobart 2003). Assessment for learning emphasises more support for learning and less for achievement. This call encourages assessment practices to shift from emphasising the quantification of educational achievement to displaying educational growth in students' competencies (Gipps and Stobart 2003). Assessment is also viewed as a dynamic process. Students participating in an assessment task engage in learning procedures while being tested. At the end of the testing exercise the student's state is changed intellectually, physically, emotionally and socially (Goldstein 1994). Assessment for learning is therefore an interactive assessment model that recognises and acknowledges the active engagement of students with assessment tasks, and that such interaction results in the modification of a student's schemata.

In using assessment to support learning and monitor students' progress, teachers can effectively promote educational growth of students by assisting them while they are being tested. Learning, as a social-process of knowledge construction, may be supported by group assessment where students discuss and formulate responses collaboratively. As Griffin *et al.* (1995) note, when students work in groups they explain their thoughts and reasoning, and receive feedback from peers and the teacher and in the process improve their intellectual competencies.

Performance assessment models have been viewed as assessment strategies that can be integrated into instruction to support students' learning and understanding. Both students and teachers in this study perceived performance assessment to provide a range of learning and assessment opportunities. The use of group assessment enhanced these advantages but also introduced some displeasing experiences for students. Thus, students' recommendations for the continued use of performance assessment were supported by perceived learning and assessment benefits from the tasks and their mode of assessment through groups. Recommendations also advocated for reduced group assessment problems through closer monitoring of the groups and a pedagogy that was geared towards performance assessment.

Perceived learning and assessment benefits and drawbacks were associated with the following categories and sub-categories:

1. Task characteristics: a) Task complexity
   b) Task importance and value
   c) Task requirements
2. Metacognition
3. Motivation
4. Group assessment: a) Peer collaboration and support
   b) Group assessment problems
5. Multiple assessment tasks

5.2 TASK COMPLEXITY

Perceptions of how difficult or easy performance assessment tasks were, were associated with topic and task content difficulty, the nature and format of the tasks, prior experiences and expectations of assessment demands, as well as the mode of administration of the tasks. Task complexity was perceived to be reduced by the practical nature of the tasks, the use of group assessment, familiarity with the task content and assessment format. Group assessment is discussed in Section 5.5.

According to Brookhart and DeVoge (1999) factors such as perceived self-efficacy, perceived effort required for the task, amount of assistance available to the students and past assessment experiences and achievement, tend to influence students' perceived task difficulty. Moni *et al.* (2002) also agree with this assertion and note that students' prior experiences of assessment may have positive or negative effects on students' views about an assessment task and their reactions to pedagogy.

5.2.1 Task level of difficulty

Students' perceptions of task difficulty varied between the groups. Some students felt that the tasks were easy while others felt they were difficult. Students and teachers in this study perceived the performance assessment tasks to be challenging and somehow confusing to some students. Tasks required students to combine their cognitive and manipulative skills in working on the assessment tasks. Students needed to know how to manipulate the equipment and what data to generate and collect, find patterns in the data and relate it to previous knowledge. Thus, each performance assessment task involved a number of processes through which students were to display their knowledge and skills. The performance assessment tasks were also embedded in a real world context, which made them more complex for some students. Teachers associated the observed confusion with inadequate pre-task instruction and a general low students' confidence in their abilities in practical work. Unfamiliarity with the assessment model was another source of the perceived confusion. Moni *et al.* (2002) contend that students tend to demonstrate less confidence in unfamiliar tasks.

a) Pre-task instruction

Task difficulty was perceived to be linked to pre-task instruction. There were students who experienced some confusion in setting up and using equipment for the electrical resistance task. In the other task students wanted to use all the tests they used during lessons rather than select the relevant materials to distinguish oxygen, carbon dioxide and air. Both these experiences from students were associated with inadequate pre-task instruction. Teachers indicated though, that they made some attempts during lessons to help students get ready for performing the tasks by working towards developing students' practical skills as required in the tasks.

b) Task and content difficulty and familiarity

Tasks on the Electricity unit were perceived to be cognitively more demanding and therefore more difficult than the task on the Air and Living Things unit. The tasks required students to think deeply about concepts on electricity, something they were not used to, did not expect, and therefore did not prepare themselves to tackle those kinds of tasks. This perception indicates that perceived level of difficulty of a given task was dependent on the task or the topic content. This observation can be expected as topic content differs in levels of abstractness. Its level of difficulty and length of time required to prepare for it will also differ for different students (Brookhart and DeVoge 1999). Students reported that they were not ready to carry out the tasks from Electricity because the unit was difficult. Webb *et al.* (1997) and Chang and Chiu (2005) reported similar findings on topic related difficulties involving electricity concepts. In a study by Chang and Chiu (2005) Ninth grade students found Electricity hard to grasp although the students were able to conduct calculations involving electricity. These students also performed better on hands-on activities but did poorly on questions that were concerned with concepts related to everyday life. Unlike the students in Chang and Chui's (2005) study, students in this study had difficulty with the calculations of electrical resistance, although the difficulty seemed to be both conceptual and mathematical.

The perception that the task on testing gases was less demanding than those on electricity, could be viewed from three perspectives. The content from the unit on Air and Living Things was easier for students to understand compared to that from the

Electricity unit. Another reason could be that the tasks were similar to activities done in class, including the context-based sub-questions in the task. Students perceived the task to link well with information from their lesson notes, which they understood. Students were therefore familiar with the content and skills required to carry out the task. A third reason could be that students had developed some familiarity with the performance assessment model, knew what to expect from the task and therefore prepared for it. The task on testing gases was the third performance assessment task to be performed by the students. Therefore, perceived difficulties with the tasks on electricity could also be attributed to low familiarity with performance assessment tasks in general and the level of difficulty of concepts from the Electricity unit, and therefore, the tasks.

c) Practical hands-on format

Although students and teachers perceived the tasks to be difficult to some extent and required students to think deeply about the concepts in the tasks, presenting a task in a hands-on practical format was perceived to help reduce their difficulty. to some extent. For some students the display of the equipment made planning easier as there was less need for the recall of appropriate equipment. The practical tasks also allowed students to manipulate the equipment in different ways and to generate data and provide clues for answering task questions. Students were also motivated by positive observations during the testing stage, such as the lighting of the electric bulb, relighting of glowing splint or change in lime water; even though the other parts of the task - the planning and answering questions - might have been more challenging.

d) Real world context

Another aspect that contributed to the perceived low difficulty of tasks for some students was the integration of real world contexts in the sub-questions. Some students, who recognised some contexts as familiar, although not fully understood scientifically, perceived the task to be easy. However, they still demonstrated difficulties in using the data to explain events in the context, opting for answers based on informal knowledge. Everyday experiences seemed to induce a perception that the tasks were easy and prohibited further thinking concerning the use of the data and scientific explanations required by the tasks. Thus, students formulated answers that lacked scientific knowledge or that had no link to the data from the tasks.

e) Previous assessment experiences and expectations

Students' previous experiences and expectations of assessment tasks seemed to influence their perceived level of difficulty of the task. Assessment tasks that students recognised as similar to activities encountered previously in lessons and/or elsewhere, were perceived to be easy to carry out. Familiar out of class experience was also perceived to improve students' confidence in conducting the tasks as observed in the electricity based tasks. Such students displayed high levels of enthusiasm and confidence in conducting tasks.

Sometimes prior experience was not compatible with the performance assessment format. Teachers and students acknowledged that students were normally given content based (theory) tests that demanded less understanding and processing of questions. Prior experience of recall assessment tasks developed expectations of similar intellectual demands from the tasks, for which students had studied and prepared. The students had not anticipated tasks that required deep thinking and understanding of concepts and therefore perceived the questions to be difficult because they were asked in unexpected and unfamiliar ways. Anderson (1980 cited in Hall 1993) noted that students who are continuously exposed to tasks of low cognitive complexity and challenge tend to prefer easy, clearly defined tasks that require less time and effort to work on. Such students tend not to recognise tasks that would have been easy if some thought was put into them.

Herman *et al.* (1997) indicate from various literature sources that, due to the testing of content in traditional standardised tests, teaching to the test led to the distortion of the curriculum for many students by narrowing it to basic low level skills. Students' perceived difficulty of the task seemed to be partly due to the unfamiliar intellectual demands they placed on the students. Students in this study indicated that they were not used to being assessed in this way and therefore their preparation for the tasks followed the familiar rote learning of facts. These findings support Wiggins' (1989) suggestion on selecting and teaching those capabilities considered essential. Performance assessment seemed to provide such a step.

The influence of prior assessment experiences on students' perceptions of task complexity and preparation for assessment tasks have also been noted by Sambell and McDowell (1998). Assessment experiences strongly influence the ways in which students approach different assessment models *(ibid)*. Students' views of the nature of learning of subject matter also influence the meanings they make in assessment tasks. The views also determine whether students adopt approaches to learning that are likely to increase understanding of their subject or not.

Students in this study indicated preparing themselves for recall type tests. They acknowledged changing their mode of learning. They attempted to learn for understanding by increasing their level of participation, attentiveness and concentration during lessons and during the performance assessment tasks. Furthermore, students became more serious about their learning in anticipation of the performance tasks and acquiring good marks. Students play an active role in constructing knowledge about assessment through their experiences with assessment tasks, interactions with each other and with teachers (Moni *et al.* 2002). Thus, students with little or no prior experience of the performance assessment were likely to find them challenging, although the difficulty of the tasks would also depend on the difficulty of the content assessed.

Tittle (1994) asserts that in the constructivist perspective teachers and students use assessment experiences to construct knowledge and beliefs about the self, teaching and learning. They use such knowledge and beliefs to interpret assessment, their intentions to use and the actual use of assessment for instruction and for learning. Assessment is sometimes used to motivate the investment of effort in preparing for the task (attentiveness during lessons and revision) or in doing the actual task. Teachers also construct knowledge about students' representations and conceptions about subject matter or beliefs about the self, the teacher and learning, which they use to judge the level of students' confusion, confidence and their abilities to succeed in a given task. That is, teachers hold beliefs about the capability of students that influence their interpretation and reactions to situations involving those students (Dweck 2000 cited in Deemer 2004). Students' knowledge and beliefs may include learning strategies, metacognitive activities, expected intellectual demands, self-regulation and their motivation towards working on the assessment tasks.

f) Teacher mentorship

Teacher mentorship while students carried out the performance assessment tasks allowed teachers to interact with students to guide and direct them towards producing appropriate plans for their tasks. Good plans enabled students to make meaningful observations and produce useful data. Improved understanding and performance of tasks were also achieved through consultations with teachers.

Teachers' presence and assistance provided to students during their performance of the tasks contributed to simplifying the tasks. According to the teachers some students fully exploited this assistance to maximize teacher input and assurance of good scores. The extent to which students consulted the teachers reached levels where teachers felt their input approached doing the tasks for the students. Teachers became concerned about how they would score students' competencies in instances of high teacher input in students' performance in the tasks. This study has shown that the teachers and students placed a high value on scores in motivating students' participation, seriousness and commitment in working on and in preparing for the assessment tasks.

Teachers' help and support during the assessment of students' work seem critical in the successful use of performance assessment to support learning. In the process teachers gain immediate insight to students' knowledge, understanding and thought processes and identify help needed for proceeding with the tasks (Shepard 2000). Teachers can discover gaps in students' thinking, as well as help to close those gaps. According to Shepard (2000) teacher mentorship during assessment constitutes an occasion for teaching and the scaffolding of subsequent task steps that lead to success concerning the task's requirements. A drawback of such scaffolding is its limited applicability to the non-scoring of assisted or independent work. In this study the differentiation between the scoring of assisted and independent performance was by means of considering the nature and degree of assistance provided to the students. Scaffolding through the use of questions to direct students to identify and correct errors in their work was not penalised. However, direct coaching, after guidance from the teacher had failed, cost the students credit that was aligned to the criteria for which assistance was provided. The reason for differentiating in scoring procedures of assisted work, was to make the task fit into a measurement paradigm where marks or grades for students' work are emphasised. The teachers held a predominantly traditional assessment view where marks are important. At pre-task level, teachers spent time preparing students for success in performing the tasks. During in-task coaching of students and scoring of students' performance of tasks, they found the dual role of mentor and judge challenging.

In a view of assessment that supports learning, performance assessment takes two roles: assessment of learning and assessment for learning, or learning through assessment. Teachers found reconciling these two roles challenging. Excessive consultation by students was perceived to be responsible for the difficulties experienced by the teachers in merging the two functions. Deciding on the extent to which they could provide guidance to the students without making significant adjustments to the scoring process was difficult for the teachers. Students objected to the non-scoring of assisted work and some developed some reluctance to consult with the teachers. Adjusting scoring according to the nature of assistance provided was necessary for maintaining consistency in assessment conditions. Some students received more assistance than others even if it was getting the teacher to check their progress regularly. Grading students' work and the purpose of the grade are important aspects of assessment. According to Nitko (2004), performance assessment tasks become assessment tools when they are designed primarily for assessment procedures, not as practice and enrichment opportunities. Assessment tasks must therefore include criteria and a scoring guide and, obviously, scores.

g) Use of scoring guide

Teachers appreciated the use of the rubrics to focus their observations and in scoring particular skills or performances. However, they found the rubrics to be too detailed to follow easily. The excess detail was perceived to slow down the teachers when scoring performances according to all the criteria. Difficulties were particularly experienced in instances where students did not display the expected procedures or when scoring assisted work. Waiting on students to demonstrate those procedures tended to delay teacher movement between groups.

Difficulties experienced by teachers in mentoring students and scoring their work were not only experienced at a professional level, but were also a problem of practicality. Implementing performance assessment tasks under the prevailing classroom environment in the schools was a challenge. Teachers felt overwhelmed by the duties they had to perform in implementing the hands-on performance assessment tasks in large classes. They had to assist students to interpret and understand the tasks, approve task plans, conduct general student monitoring and observe and score students' performances according to the rubrics.

Students' consultation with teachers and teacher mentoring of students led to extended waiting periods for the approval of students' plans by the teacher. It also led to the non-observation of performances by students who did not wait for the teacher to approve their plans, but moved on to the testing stage of the task. Students who collected data without being observed by the teacher were placed at a scoring disadvantage. No marks were given for Stage 2 of the task without the teacher's observation.

Teachers' concerns about the amount of detail in the rubrics, and its scoring applicability, have implications for rubric construction. The rubrics criteria need to be teacher-friendly to facilitate a comfortable observation and scoring of students' performance of assessment tasks. It may, therefore, be necessary to focus criteria on skills and procedures deemed important in the students' present and future endeavours.

While clear rubrics were essential for focusing teachers' scoring of skills and procedures they also felt that the amount of detail in the rubrics should support and not hinder scoring. In addition, the teachers suggested that rubrics for use in science classrooms need to be flexible to accommodate needs in individual classrooms, particularly so if they are intended to support learning through assessment. Situations and experiences that justified flexibility in the use of rubrics in classrooms where time was not enough for the various school activities were:

- Students omitting certain criteria such as not indicating the procedure for recording data before conducting the test, yet report all their observations after the investigative stage.

- Students assembling the circuits, then drawing the circuit diagram from the actual circuits rather than drawing up the plan first and then doing the activity.
- Students not waiting for the teacher to observe them execute the testing stage of the task, but proceeding to take readings or make and record observations.
- Students waiting and losing time while the teacher monitored and scored performances of other groups.

Sometimes the omission of certain procedural details from the plans by the students, was attributed to omissions during pre-task instruction. For classroom instruction to adequately prepare students for the assessment tasks, there was a perceived need to match instruction to the rubrics, a suggestion also advocated by Wiggins (1989) to promote testing that serves teaching and learning. He notes that if tests determine what teachers teach and what students study then it is only logical to test capacities and habits thought to be essential. In this way practicing for the test is likely to enhance rather than impede education. Such a match seems to be a way of encouraging teachers to develop important practical skills in students. Dochy and McDowell (1997) also argue that assessing higher order skills through authentic assessment, such as performance assessment, is likely to lead to the teaching of higher order knowledge and skills.

The perceived need to match instruction to rubrics, or rubrics to instruction, seems to favour the practical assessment of students' performances and the support for the learning of the selected capabilities and habits. The performance assessment tasks also influenced teachers' classroom practices. Teachers in this study indicated that they made deliberate attempts to ensure that students were ready with the relevant knowledge and skills for the performance assessment tasks. According to Shepard (1995), it is normal for teachers to teach specific strategies for producing acceptable answers, unfortunately this move tends to reduce the conceptual challenge of the assessment problem. However, Wiggins (1989) argues that in teaching to the test, a conceptual challenge can be maintained if essential abilities and habits are carefully incorporated into tests if those tests are then used to channel teaching.

Borich and Tombari (1997) and Nitko (2004) note that carefully constructed performance assessment tasks and rubrics need to be used by trained scorers if performance assessment models are to be effective. This advice implies the importance of reliable scoring in performance assessment as a primary concern of assessment. It seems to be at variance with the vision of assessment for learning where scores take a secondary role (Gipps and Stobart 2003). The place of performance assessment scores in this study was questioned on certain grounds of group assessment. The use of scores from group performances to account for and report individual students' achievement was perceived to be a problem for the students and the teachers. They perceived no match between group scores and individual performance in instances of non-participation and dependence by some group members. Boud *et al.* (1999) also raised concerns about the relationship of new forms of assessment, which support peer learning with regard to the need for marks and grades. The findings from this study indicate a possible need for carefully designed strategies that may allow the merging of the new role of assessment, assessment for learning, with group scoring and the use of scores for reporting students' progress. Such strategies may also make the mentor-assessor role of teachers easier.

5.3 TASK IMPORTANCE AND VALUE

Three main perceptions regarding the importance and value of performance assessment tasks were presented. The students and teachers perceived the performance assessment tasks to possess learning, motivational and assessment properties. Perceived importance, value and future functionality of assessment tasks motivated the students to invest effort in their work. If students identify some value in a task they are likely to devote more time and effort to it (Brookhart and DeVoge 1999). Students also have strong views about different formats of assessment (Struyven *et al.* 2003) and these views are important driving forces for the effort students invest in executing assessment tasks (Maclellan 2001; Brookhart and Bronowicz 2003).

5.3.1 Performance assessment and learning

Performance assessment tasks were perceived to provide opportunities for students to learn. Students' and teachers' perceived benefits for learning involved increased acquisition and consolidation of students' knowledge.

Learning opportunities were in the form of making observations, obtaining readings from instruments and generating meaningful data that were interpreted and used in answering the questions. For this to happen, tasks need to generate appropriate and meaningful data to enable students to complete the task in a meaningful way.

In the new assessment paradigm where assessment aims to support learning, the performance assessment tasks lead to meaningful teaching and learning activities. According to Shavelson *et al.* (1991) tasks used for assessment of learning can also be used as exercises through which students could further explore their understanding and their ability to use their knowledge in new situations. Since performance assessment models focus on both assessment of achievement and developing understanding and skills, they also help students to learn (Gipps and Stobart 2003; Moorcroft *et al.* 2000). According to Shepard (2000) viewing performance assessment as having this dual role makes it a useful occasion for teaching where teachers use interaction in scaffolding subsequent steps in the task. In the treatment of performance assessment tasks as learning activities and measurement tasks, the teachers need to have an open relationship and interact with students to be able to guide them while working on the assessment tasks. On a similar note Nitko (2004) advises that not every performance learning activity is an assessment activity. An activity that is to be used as a performance assessment task needs to focus on specific learning outcomes and the evaluation of these against established criteria.

The importance and value of the performance assessment tasks were perceived in terms of student leaning and measurement of student learning. Students' perceived the tasks to provide direct learning opportunities involving acquisition of knowledge and skills (procedural, practical, intellectual) from engaging in the tasks. This benefit corresponds to the aims of practical work - demonstrating scientific concepts to promote their understanding by students and the development of laboratory skills. When students work with the task content material they internalise the material, think about it creatively and can therefore remember it better and longer (Struyven *et al*. 2003). Learning by students also involved awareness of the uses of science in the environment through the inclusion of realistic tasks

Other ways that provided learning opportunities were the interactions that took place during the administration of the assessment task in groups and teacher mentorship. Indirect contributions to students' learning involved pre-task instruction. Students also perceived gains in knowledge resulting from the use of multiple assessment tasks. All these are discussed further in sections 5.4.1; 5.5; 5.7 below.

5.3.2 Performance assessment and motivation

The tasks were found to be interesting to both teachers and students. Perceived motivational aspects of performance assessment were of intrinsic and extrinsic nature. Intrinsic motivational aspects involved students' perceived personal interests and enjoyment. There were students who perceived the tasks to be interesting and those who felt that the tasks improved their enjoyment of Science. Thus, some students were observed to be very enthusiastic about carrying out the assessment tasks.

Students and teachers perceived performance assessment to motivate students to participate more actively, and with more attentiveness, seriousness and commitment while working on the tasks, as well as during regular lessons. Students' serious behaviour during normal class activities seemed to be encouraged by their anticipation of performance assessment tasks. Students and teachers also perceived performance assessment to provide a rationale for students doing practical activities during lessons.

Extrinsically inclined perceptions were dominated by achievement scores. Perceptions regarding the allocation of group marks to individual students were mixed. These perceptions were presented at three levels. These were marks from different aspects of the tasks, marks allocated for group effort, and 'free marks' that were acquired by non-participating students.

Students perceived the allocation of marks for different parts of the task (planning, testing, analysing and answering questions) a favourable aspect of performance assessment. They appreciated the recognition of their efforts in different stages of the tasks through marks and their contribution to a good overall subject grade.

Other perceived contributions to higher scores, apart from the practical nature of the tasks, were:

- Collaborative group work.
- Teacher mentoring during tasks.
- Increased motivation that resulted in the improved, committed and serious participation of students.

Performance on tasks improves if students interact among themselves and with the teacher while they are engaged in the task (Slater (not dated)). This assertion indicates that students can achieve higher grades if human support, such as peer or mentor support, is available to the students during the assessment process. Such support is not permitted in traditional assessment models. While students' seriousness and effort may have been stimulated by the desire for better marks, teachers perceived that such actions were likely to increase students' learning prior to and during the performance assessment tasks.

Perceived motivation through marks was also demonstrated by teachers' cautiousness when withholding marks while scoring assisted performances. Assigning scores and checking performance against rubrics were acts of facilitating accountability of students' learning (Brown 2004). There were feelings among the teachers in this study that the marks obtained from the group performance assessment needed to be used in accounting for students' learning in their school reports. Teachers perceived that non-use of scores in students' term reports was likely to reduce the level of motivation, seriousness, and commitment with which students prepared for and participated in the performance tasks. These perceptions corroborate Boud *et al.*'s (1999) assertion that the use of formal assessment through assigning scores is an important component of assessment. If no grades are assigned to a task both, students and staff see those tasks as less important, and tend to divert their efforts to learning goals that are assessed.

5.3.3 Performance assessment model in assessment

When used as a measurement tool performance assessment can measure an array of skills and abilities. The opportunity for students to use knowledge in a variety of realistic situations and contexts, and to use what they have learned in a myriad of hands-on activities, can thus be assessed (Aschbacher 1991; Shavelson *et al.* 1991;

Gott and Duggan 2002). The perceived role of performance assessment in increasing students' awareness of the uses of science in their environment strengthens the use of performance assessment tasks in combination with contextualised or applications-led science teaching and learning.

Students and teachers recognised the role of performance assessment tasks in enabling teachers to judge students' learning. The teachers also perceived the performance assessment tasks to be important in the measurement of practical and social skills. Teachers in this study appreciated observing students perform practical tasks for the purpose of assessment, enabling them to focus their assessment on both the processes and products of the practical activities. The opportunity for teachers to assist students during their performance of the assessment tasks enabled teachers to gain insight about the students' knowledge and understanding and the ways in which they could enhance this knowledge and understanding, as observed by (Shepard 2000).

Through group administration of the performance assessment tasks, teachers could also gauge the level of social skills such as cooperation and communication among students. Teachers could also encourage cooperation and participation of students in discussions and in operating equipment for the successful completion of a task. Viewing performance assessment tasks as facilitating learning and as appraising learning, and provides supporting evidence for the successful integrating of performance assessment and learning activities.

5.4 TASK REQUIREMENTS

To administer performance assessment tasks successfully, students must possess the competences being assessed and teachers must have the means of implementing the tasks. Linn and Gronlund (2000) advise that students must be provided with the necessary intellectual and physical tools to embark on performance assessment tasks.

Students experience specific expectations each time a particular assessment task is to be administered (Brookhart and DeVoge 1999). They formulate some notion of what the task might require from previous experiences of assessment tasks. In this study the participants perceived the required tools to be related to pre-task instructional preparation and in-task demands. Adequate time and equipment were perceived necessary resources for successful implementation of the performance assessment tasks. Time requirements were more in favour of pre-task preparation to allow for more practice with practical activities and development of understanding. Additional time was also needed for more assessment tasks.

5.4.1 Pre-task instruction

Performance assessment tasks used in this study were constructed to ensure that the materials to be used by students in carrying out the tasks were similar to those used during lesson activities. Students, however, indicated that their pre-task exposure to the material was sometimes limited, an observation confirmed by the teachers and attributed to limited equipment.

Pre-task instruction was perceived important for getting students prepared for the performance assessment tasks. According to Herman *et al.* (1997), pre-task instruction is important in preparing students adequately for tasks that require critical thinking from students and which encourage them to draw their own procedures and conclusions to problems. Both teachers and students also recognised and acknowledged the need to develop a greater understanding of scientific concepts and procedural knowledge and skills, as well as thinking skills during lessons. Students perceived their readiness for the assessment tasks in terms of experience in practical activities and the acquisition of skills, knowledge and procedures necessary for performing the assessment tasks.

According to students' perceptions, performance assessment tasks motivated them to be more attentive during class demonstrations. Students were aware that the performance assessment tasks demanded a different level of concentration during lessons. They recognised the need to take on an active role in their acquisition of the intellectual and physical tools required for performing the assessment tasks. Students who did not concentrate during lesson demonstrations experienced difficulties working on the tasks.

Another indirect way in which performance assessment supported pre-task learning of practical skills by students was increasing opportunities for students to acquire task competencies prior to the task. Teachers made deliberate efforts to increase opportunities for students to develop practical competencies as assessed through the tasks. Teachers' actions were in line with Linn and Gronlund's (2000) advice that students must have developed the skills to be assessed before they are assessed. The reaction to the use of performance assessment by focusing on developing students' practical skills seem to indicate the effect of performance assessment on teaching where teaching is matched to desired skills and competencies directed by assessment tasks. However, spending a bit more time on developing students' practical skills was perceived to slow down the teaching pace.

5.4.2 In-task demands

Students' perceptions of pre-task and in-task requirements involved input from both the teachers and the students, although in-task input was more student related. Students' perceptions of in-task requirements included the need for students to understand the tasks; the content embedded in the tasks; and be able to use their knowledge as required. Students recognised the need to think deeply about the task before responding, as well as being able to apply practical, inquiry and procedural skills. They also needed to have a balance of the competencies assessed, including theoretical and applicable knowledge.

Perceived in-task readiness by students was topic and task-dependent. Students needed to understand the topic, and that different parts of a topic have different demands for understanding. The students experienced difficulties in understanding concepts in the Electricity unit and therefore felt not ready for the tasks included under this topic. They perceived the topic on Air and Living Things as being easier to understand and were ready for those tasks. Achievement in performance assessment tasks has been found to be task-dependent and cannot be generalized to other tasks (Popham 1999; Sanders and Horn 1995; Shavelson *et al.* 1991).

Students in this study associated their readiness for the task with the difficulty or ease with which they understood the topic. There is, however, insufficient evidence to link their perceived readiness to the level of difficulty of a particular topic considering that only three tasks were taken for the two units. Furthermore, the tasks on Electricity were done before the one on gases. So, the influence of unfamiliarity with the assessment model and tasks cannot be ruled out in contributing to the perceived lack of readiness for the tasks on the Electricity unit. Nonetheless, evidence from the literature indicates that there is a link between performances and task familiarity. Moni *et al.* (2002) found that students were more confident in tasks that appeared familiar. Unfamiliarity with tasks reduced confidence and generated weariness in students. These effects, however, faded away as students worked on the tasks. In this study teachers not only attributed students' dependence on teachers to students' lack of confidence in their practical skills, but also to lack of familiarity with the performance assessment model. Teachers thus expected students' level of consultation, and therefore their dependence on the teachers, to decrease as they became more familiar and confident in the performance assessment model. These expectations show some hope for the use of performance assessment as an alternative model of assessment in Science in Swaziland.

Linn and Gronlund (2000) also advise that if students do not possess the ability to use task material, then such abilities should be developed during the tasks. The presence of a tutor to help students in the use of the task material and to understand the task was perceived to be a good approach. Consultations with the teachers were beneficial for the students as students could confirm ideas with the teacher or the teacher could guide them towards appropriate procedures. However, this study has shown that some caution is required when assisting students during performance assessment tasks. Some students may rely on the teacher to tell them what to do. In such instances students might not fully utilise learning opportunities presented through regular class activities taken prior to the task. Other students may not see the need for attentiveness and concentration during lessons. There might be a danger of students lazing around during lessons and even during the performance of the task, a behaviour that was perceived to be a problem of group assessment. According to Struyven *et al*. (2003) perceived assessment requirements have a strong effect on the approach students adopt for learning, be it surface or deep learning. Thus, students adapt their learning according to perceived demands implicit in the assessment task.

5.4.3 Task resources

Students and teachers identified the need for adequate resources for the successful implementation of performance assessment tasks. Performance assessment tasks tend to be more complex than multiple choice or short answer structured assessment questions and may need more time and other resources for students to work on (Roberts and Gott 2004; Linn and Gronlund 2000).

Task requirements also include resources that were perceived to influence pre-task preparation of students and were necessary for students to carry out the tasks successfully. Resources that were identified were time, equipment, space and personnel.

a) Time as a resource

Students' perceived their readiness for the tasks to depend on the time they were provided prior to the tasks to prepare themselves. Students needed time to revise their work to become emotionally ready for the tasks. Perceived readiness varied even among groups of students in the same class where one group indicated they were ready while another indicated not to be ready because they lacked the time to prepare for the task. Central to these perceptions was not the variation of students' views about being informed of the test in time, but the relationship of the time available to prepare for the tasks and students' perceived state of readiness. Students perceived the need to allow time to organise themselves for the task, particularly because of the high regard for assessment results.

Time demands experienced in this study were associated with pre-task preparation of students by teachers. Teachers perceived that getting students ready for the tasks necessitated them spending time ensuring that students developed the practical skills needed for performing the tasks or revising work with students for the tasks There were, however, concerns that the use of performance assessment tasks increased the amount of teaching time and slowed down the teaching pace.

The students' and teachers' perceptions regarding time requirements for performing the tasks varied. Time related issues involved the school administration and organisation, teaching processes, as well as other student and task related factors. The amount of teaching time available on the school timetable (six thirty or thirty-five minute periods per week) was considered insufficient for either teaching Science or administering performance assessment tasks. Nuttall (1992) notes that the performance assessment tasks demand a significant amount of teacher time and energy to construct, administer, and grade as compared to the standard assessment models normally used.

Teachers in this study seemed uncertain about the time required for the actual administration of the tasks. They perceived the amount of time required for carrying the tasks to be slightly more than that allocated for Science in the timetable, but felt that other factors contributed to increased time demands for the tasks. These factors were student related; group assessment related; equipment related and/or task related. According to the teachers, performing the tasks did not take a lot of time, except for the time needed to accommodate delays by students in starting on the tasks, resolving students' confusion about tasks, excessive consultation of teachers by students, and unproductive group discussions. Students tended to take sometime before starting on the task and there were students who consulted almost at every step that resulted in delayed progress for other students. Consultations were perceived to be due to students' low confidence and unfamiliarity with the assessment format.

Other factors that increased the time for administering the tasks, were:

- Teacher approval of students' plans before they proceeded to the next stage of the task and observing them while they performed the tasks.

- Teacher mentoring of students who had difficulties with the tasks. The constant consultation of teachers by students caused time constraints and the perception that performance assessment tasks required a lot of time to work on to complete.

- Adhering to certain assessment criteria by insisting that students meet such criteria, in particular producing a table of results at the planning stage. Students seemed to experience difficulties designing such a table. A possible source of this difficulty was that tables of results were provided in most of the activities that students conducted during lesson. Teachers felt it was not necessary to insist on certain procedures as long as the students were able to find alternative ways of reporting their results.

These perceived challenges may be indicative of some of the hurdles that could be experienced when implementing performance assessment models, particularly in under-resourced schools.

b) Equipment as a resource

This study used hands-on practical performance assessment tasks and therefore required laboratory equipment for implementation. Equipment as a resource was perceived important in determining the format of administering the task, whether group or individually administered tasks.

Tasks that required delicate and expensive equipment would tend to be affected more by equipment shortages than those that required less sophisticated equipment.

Equipment needs to be in good working condition for the successful implementation of performance assessment tasks. However, it should also be acknowledged that if students do not have the dexterity with which to handle equipment, even equipment that is in good working condition could malfunction. Electrical conductivity was essential for activities in the Electricity unit and keeping such equipment in good working order required careful handling by the students. The delicate nature of the conducting wires and electricity meters required the support of a technician while the tasks were being carried out, for example, to mend connecting wires or to get the current flowing in circuits. Such support eased the anxiety students experienced when the equipment appeared to stop working due to a loose contact and the thought of losing marks.

5.5 GROUP ASSESSMENT

Perceptions on group performance assessment comprised benefits and problems. As noted in sections above, benefits involved learning from each other and reduced task complexity through peer tutoring and the sharing of ideas, as well as improved scores due to the quality of responses. Perceived problems included non-participation, dependence and dominance by certain group members, non-resolution of different views, and the acquisition of undeserved group marks. Students' recommendations encouraged the use of group assessment on grounds of perceived peer collaboration and support, and suggested strategies to deal with some of the perceived group assessment problems.

5.5.1 Peer collaboration and support

According to Webb *et al.* (1997) group assessment is justified on the grounds that if teamwork and collaborative learning are valued, this must be reflected in an assessment process which emphasises that students are judged on their collective efforts. The same benefits experienced in regular group work activities can be experienced when assessing students in groups. The students support and tutor each other and through the interaction with each other and with the task, students acquire new skills, ideas and knowledge (Webb *et al.* 1997). Thus, group performance assessment enabled students to learn through the tasks and through interaction with group members, as well as with the teacher.

Students and teachers perceived several benefits from group performance assessment. Participants perceived group interaction to improve their understanding of the tasks and the subject matter embedded in the task. It became seemingly easy for students to understand the requirements of the tasks through collaborative interpretation of the tasks. According to the students they also learned from each other as they discussed and shared ideas related to the content and procedures of conducting the tasks. Students who were more knowledgeable about a tasks and its content helped others to understand different aspects of the task.

Students also perceived an improvement in the quality of their responses and therefore an increased possibility of getting good grades and passing. Students working collaboratively have been shown to obtain combined higher performance output than those working alone (Fawcett and Garton 2005). Though task products or answers are important particularly for grades, the real learning occurs during the process of generating the answers. During the planning stage of the tasks students debated different inquiry aspects such as hypotheses and the procedures for testing the hypotheses. They also debated possible answers and agreed on which to use. Fawcett and Garton (2005) and Hall (1993) note that the quality of the interaction, cognitive benefits and the answers are determined by students' motivation, interest, confidence and comparative ability levels.

Students were also perceived to develop important social outcomes associated with collaboration, teamwork and communication (including listening) skills and learning to learn from others. These benefits are related to peer learning that can foster certain types of lifelong learning skills (Boud *et al.* 1999). Discussions during group performance assessment increased students' opportunities to reflect on and explore different ideas, practice articulating their ideas and experience being critiqued by peers. Group work has been known to encourage student learning through collaboration among students. Such collaborative learning also increases social and emotional outcomes such as social skills, self-esteem and attitudes towards others (Webb *et al.* 1997). Students who engage in the collaborative exchange of ideas through talk and active debate are more likely to benefit cognitively than passive ones (Fawcett and Garton 2005). The extent and benefits of such interaction for students in this study were also likely to depend on the actions of dominant students in the groups. Such students were perceived to disregard contributions from other group members.

5.5.2 Group assessment problems

The preceding discussion shows the strengths of group-work during assessment and how it can improve support for learning through assessment activities. In group assessment the students expected to work together towards a common goal, but were disappointed that there were no strategies to ensure compulsory participation of all students and mutual respect for ideas presented. Students were therefore concerned that would be benefits of group assessment through peer collaboration and support were weakened by differentiated participation and conflicts in discussions.

A few factors were identified for the perceived lack of participation and dependence by some students. These factors were student related and non-student related. Student related factors were perceived to be unwillingness to make contributions, deliberate withholding of information by some students or lack of relevant knowledge for the task. There were also concerns that students who did not know much about the subject or the tasks successfully hid behind those who were knowledgeable. These concerns seemed to be aligned to a competitive assessment approach.

Dependence on each other was perceived from two perspectives: underserved marks and 'stealing of ideas from others'. Perceived dependence by non-participating students on willing and enthusiastic students was considered unfair by the students and the teachers. Perceived dominance by certain group members was considered unreasonable. Dominating students took charge and did everything for the rest of the group without any regard whether all the group members understood and followed the processes involved in performing the tasks or not. Dominance was perceived to have a negative effect on slow learning students who were not able to follow and keep pace with the working rate of the group and who, therefore, were likely to fail to benefit from learning through the task or from peers.

Dependence and dominance are characteristic of group dynamics and reflect the complexity of human interaction. They may not be easily and closely monitored by teachers due to the other responsibilities they take on during the administration of performance assessment tasks. To minimise the impact of such complex human interaction during assessment involving group work, Webb *et al.* (1997) note from literature that it may be better to use problems that have obvious answers and that can be easily explained or demonstrated by a single competent group member. However, the presence and contributions of such a competent individual may still encourage dependence by certain individuals in the group and feelings of exploitation of the competent member. It may also be perceived as dominance. Committed and competent students felt cheated by working with non-participating students, which are feelings that may work against motivation in future tasks.

Teachers seemed aware that teamwork and collaborative effort should be judged on the basis of students' collective efforts. Unfamiliarity with collective group evaluation seemed to create collaboration and scoring difficulties. Differentiated participation (non-participation, dependence and dominance) presented challenges for teachers regarding observing students and scoring their procedural competencies. The allocation of the same marks to all group members was perceived to be particularly advantageous to students who did not know much about the tasks. However, grades received by less knowledgeable students were perceived to be unfair to those students who put more effort into the task and deceptive to students who lacked the knowledge required in the tasks. Students who are used to being judged in terms of their own individual effort can resent others gaining credit for work perceived not to have been done (Webb *et al.* 1997) or for which they are perceived to have deliberately withheld contributions, as shown in this study.

The unfair and undeserving allocation of group marks to all group members seemed to influence the motivation of certain students. According to Boud *et al.* (1999) perceived non-participation and dependence may affect students' motivation and the effort invested in preparing for and working on tasks if they believe others will benefit equally from such effort. Students may also be encouraged to either reduce their effort and contributions or continue non-participation, a practice that could be unfavourable for performance assessment and learning.

Not only were teachers concerned that marks for low performing students were elevated by working with high-ability students, they were also concerned that marks for high achieving students were lowered by working with low achieving students. Some high achieving students may not be assertive so that their ideas may not be expressed or considered by the group. Non-expression of ideas may have been judged by others as withholding of information from the group, particularly if the student was known to do well in school. Marks could also be unfairly lowered if the students' contribution was disregarded and/or there was lack of consensus during group discussions.

Despite concerns about the possible non-correspondence of group marks and students' general achievement, teachers still felt the need for heterogeneous assessment groups as opposed to homogeneous groups on merit. Such groups were perceived to be beneficial for poor performing students as they would be helped by the other students, thus giving marks a secondary position. Students of different ability levels working together in a group can facilitate the construction of knowledge through social interaction (Fawcett and Garton 2005). Webb *et al.* (1997) also observe that low-ability students working with above average students tend to perform better.

Success of group assessment seems to depend on how students perceived the contributions of group members, whether there was dominance, dependence or cooperation. Motivation for success may also be affected by students' perceptions of fairness in the allocation of marks to individual members of the group. Strategies for helping students understand how group assessment functions may be necessary. Such an understanding may enable students to demand explanations from dominating students or refrain from quickly judging other students as non-participating or as less knowledgeable. It may also encourage students who normally do not participate to be more active in group discussions. Furthermore, the use of group scores may need to be carefully considered especially because they are important for student motivation.

Students recognised the benefits of group assessment in peer collaboration and support though they still felt that disagreements among group members were not beneficial if no agreement on answers was reached. Students in this study perceived conflicts to be a hindrance to their progress towards finding correct answers and therefore good marks. Long discussions that did not lead to agreeable conclusions were perceived a waste of time and resulted in non-completion of the tasks.

Tao (1999) notes that conflicts arise whenever students disagree on procedures and solutions in problem-solving but such conflicts ensure exchanges that somehow force students to come to some agreement and co-construction of the solution. According to Tao disagreements arising from students' differences in points of view create a disequilibrium that demands resolution by the students. In order to resolve the disequilibrium, students need to reflect on the different views raised in the group discussion and then generate the required response. Furthermore, if students encounter some disagreements in their discussions they usually end with the correct solution, making initial disagreements more beneficial for learning through the reconstruction of ideas.

Students who work towards resolving disequilibrium resulting from differences in points of views and debate the different points find group assessment beneficial in generating quality responses. Studies on collaborative work have also shown that students perform better when in a collaborative situation than when they work alone, although this may depend on learning styles of students (Fawcett and Garton 2005).

Quality answers are important for getting marks, and so is the students' learning that takes place when they work collaboratively towards generating the answer. Students' motivation, interest and the cognitive processes they engage in when working on a task determine the answer to the task (Hall 1993) as well as the amount of learning that is likely to occur.

5.6 METACOGNITION

Teaching practices that encourage self-assessment also encourage students to become more metacognitive about their own thinking and learning. Students engage in a process that can help them develop control of their own learning by a process of self-assessment or self-evaluation of their knowledge (Donovan and Bansford 2005; Gipps and Stobart 2003). According to Shavelson *et al.* (1991) performance assessment requires students to monitor their own performance as they proceed with a given task. Different students perform differently on different tasks with some performing well in one task and poorly in another. Combining students with different preferences for topics or content areas may produce a supporting learning environment for students.

Whilst this study did not take on a complete metacognitive approach to learning by students, perceptions indicated that there were good opportunities for the use of performance assessment tasks to promote metacognitive learning. The performance assessment tasks used in this study and their mode of administration (group assessment) were perceived by students to promote some elements of metacognitive learning. Students found the tasks useful for self-assessment and monitoring what they already knew. Students questioned what they knew, what they read or what they had been told by their teachers as they looked for collaborating evidence from ideas expressed by their colleagues in the groups and from the empirical data obtained from carrying out the tasks and its interpretation. In the process students had opportunities to construct well debated and articulated knowledge.

5.7 USE OF MULTIPLE ASSESSMENT TASKS

Performance assessment and unit tests can be useful in assessing a wide range of different competencies, abilities and skills. The two assessment models used in this study can be used to strategically monitor and consolidate the development of difficult concepts. Performance assessment models employ tasks that are generally complex, evoking multiple performances in students and eliciting complex, higher order thinking and reasoning skills (Baker, O'Neil, and Linn 1993).

Teachers and students conveyed mixed perceptions regarding the use of multiple assessment tasks. The two alternative assessment models used in this study were perceived to allow a balanced assessment of students' competencies and assessment preferences by different students. The different task formats provided opportunities for students to make up for low scores. Different tasks require different levels of confidence from students and different students have different levels of confidence in carrying out a particular task. The level of motivation a student has to work on a task thus depends on the nature and format of the task and learner preferences of learning and assessment style, that is, whether it takes a hands-on practical or a theory format (Honeyfield 1993).

Perceptions that multiple assessment tasks were liable to lead to an over-assessment of a topic and cost teaching time were evident amongst the students and teachers. Teachers were concerned that students could be assessed on the same knowledge and skills several times. Assessment tasks may share assessed skills as can be observed in the tasks used in this study. In the two performance assessment tasks used in Electricity, the circuit diagrams and circuit constructions used differed only in certain circuit components and the purpose served by those components. In fact, the task on electrical conductivity was the basic circuit for the task on electrical resistance, thus constituting a possible perceived excess in assessment of skills in constructing circuits. Another example of a perceived over-assessment was the assessment of the concept of electrical resistance and its calculation in the practical task (see Appendix IIA) and also in the context-based unit test (see Appendix IIIA Question 4). However, this overlap or repetition was intentional.

Concerns about possible over assessment were counterbalanced by the perceptions that students had more chances of correcting mistakes made in previous assessment tasks and improve their grades. Any apparent over-assessment needs to be deliberate and purposeful, such as consolidating difficult or fundamental concepts and skills or paving the way for more complex tasks. In addition, it should be acknowledged that a task may be easy for one group of students, but difficult for another (Shavelson *et al.* 1991) depending on certain student factors such as content preferences and abilities. Tasks may be designed to match students' capacities, or to make new demands yet be feasible for the students (Honeyfield 1993). Rubric criteria may also need to be carefully constructed to avoid unintended multiple scoring of the same competencies.

The use of multiple assessment tasks seemed to favour assessment for learning, but was demanding for assessment for measuring the level of learning, in terms of time and pre-task instruction. Perceived time constraints were associated with the need for special revision of work done. Treating performance assessment as measurement tools motivated teachers to spend time revising and ensuring that students developed practical skills and understood concepts. Such efforts increased time spent on a given topic or unit. The more performance assessment tasks the more time was required for pre-task instruction.

According to Yung (2001) studies on how teachers implement new initiatives in the curriculum show that the philosophy of a curricular innovation may be significantly different from the beliefs teachers hold. In such a situation, teachers may re-structure and modify their beliefs to accommodate initiatives or 'domesticate' the curriculum in order to fit it into their belief system. In the case of integrating assessment and instruction the teachers were more inclined to keep the teacher and assessor roles separate and treat performance assessment tasks as measurement tools. The teachers had a challenging responsibility of merging their roles of teacher and assessor. Furthermore, experiences of the teachers in merging principles of assessment for measurement and testing to assessment for learning, indicated the difficulties teachers may encounter in the use of assessment for instruction and learning, as well as for measurement.

This study has also indicated perceptions that the time spent on tasks may be reduced when students (and teachers) become more and more familiar with the use of performance assessment tasks. Once students are familiar with the procedure followed in conducting performance assessment tasks, have the necessary practical and intellectual skills to carry out the tasks, and the social skills required for group work, they are likely to consult less, work faster and handle the equipment with greater dexterity. Using more tasks would help students develop the necessary familiarity and skills to shorten the time they spend on the tasks. The supervision by the teacher remains a challenge.

5.8 SUMMARY OF CHAPTER

This chapter discussed students' and teachers' perceptions regarding the use of a hands-on performance assessment model, as well as the use of group assessment. The students and teachers, on many of the aspects of the performance assessment model, held similar perceptions. The list below summarises these perceptions.

- Performance assessment was perceived to provide opportunities to assess students' knowledge and procedural skills, as well as social competencies.
- Performance assessment tasks were perceived to be challenging and to require substantial thinking and understanding.
- Prior experiences and therefore expectations of assessment tasks and unfamiliar format of the assessment model, made tasks appear difficult and confusing.

- Group support, collaborative effort and teacher mentorship helped students in the interpretation of the tasks, helped to reduce task difficulty and led to the production of quality responses to the tasks, implying better achievement and motivation.
- Students' motivation was perceived to be improved by the practical format of assessment tasks and their assessment purpose. Students' participation, commitment and seriousness in the task and lesson activities were also perceived to improve.
- Learning of content and practical procedures was perceived to be derived directly from the assessment tasks and from peer tutoring and teacher mentoring and guidance.

- Pre-task instruction was perceived to play a significant role in student preparation for performance assessment tasks. Students and teachers appreciated their roles during pre-task instruction. Students felt that they needed sufficient experience in relevant practical tasks to develop the skills and knowledge necessary to handle the performance assessment tasks. Their attentiveness and concentration during lessons improved. Teachers made deliberate attempts to equip students with the abilities required to perform the tasks.

- Awarding group scores to individual students was perceived unfair to students who worked on a particular task while others made no contributions for reasons such as perceived unwillingness to participate; withholding of information; lack of opportunity to participate due to dominance by some members (or large groups); lack of knowledge and skills regarding the content and procedures of tasks and therefore dependence on others.

- Resources such as equipment and time were perceived essential for both pre-task instruction and for the performance assessment tasks. Time seemed to be required mainly to allow students to become familiar with the task format, build their confidence and to accommodate discussions of students' ideas.

- Merging teacher and assessor roles during the administration of performance assessment was perceived difficult for teachers due to differentiated student participation, high student consultation and dependence on teacher support, as well as a lack of opportunity to observe students due to time constraints and the mentorship role.

- The use of performance assessment in combination with paper-and-pencil unit tests was perceived to provide a balanced assessment of student competencies though caution was necessary to avoid over-assessment of certain competencies.

In the next chapter the data and results on the perceptions and experiences of students and teachers regarding the context-based assessment model are presented.

6. CHAPTER 6

RESULTS II: PERCEPTIONS OF CONTEXT-BASED ASSESSMENT

6.1 INTRODUCTION

This chapter reports on the data and results of the exploration of students' and teachers' perceptions regarding the use of context-based questions in assessing student learning in Science. It describes results that lead to the answering of the second research question, namely:

> How do students and teachers view the use of context-based assessment in assessing learning in Science?

The chapter is divided into two sections. Section I describes students' perceptions and Section II describes teachers' perceptions. Perceptions are presented according to categories generated from participants' statements. Excerpts are presented verbatim to reflect the perceptions, as well as to authenticate and increase the validity of the interpretations made. The following examples illustrate how the excerpts are presented.

> They are simple to understand and have clues of the answer (ALM1 3B 8 Q5)

> Some things are obvious since they are around us, however, some are found only in suburbs which then give us a problem we have to imagine something you have not seen (EM3 2B 20 QQ6).

The above excerpts should be interpreted in the following way:

> "They are simple to understand and have clues of the answer"

is the statement made by a respondent;

AL	denotes Air and Living Things Unit test
E	denotes Electricity Unit test
M1 and M3	denote School M1 and School M3
2B and 3B	denote classes: Form IIB and Form IIIB
8 and 20	denote students assigned the numbers 8 and 20
Q5	denotes reference to test Question number 5
QQ6	represents Questionnaire Question number 6

SECTION I

6.2 STUDENTS' PERCEPTIONS OF CONTEXT-BASED TESTS

As mentioned in Chapter Three students' views on context-based questions were explored through a questionnaire that was completed by each student immediately after they had written each of the two unit tests. The unit tests comprised context-based questions that required students to use scientific information to state or explain occurrences or the behaviour of objects or people in real life situations provided with the question. For the questionnaire students had to indicate the following:

1. test questions they liked;
2. why they liked the test questions;
3. test questions they disliked;
4. why they disliked the test questions;
5. other feelings they had about the test;
6. their views about the use of similar questions in future science tests and explain these ideas. (See Appendix IIIA and Appendix IIIB for unit tests and questionnaires.)

Questionnaire items were open ended to allow the respondents to freely express their views and experiences pertaining to the tests and questions. Section 6.2.1 describes briefly the process of generating the categories of perceptions, students' perceptions in different categories and students' recommendations.

6.2.1 Generating categories

Analysis of students' responses to the questionnaire from each unit test involved the generation of a codebook using ideas from the students' statements, as well as codes and sub-categories used in the performance assessment questionnaire analysis. Statements from each student were coded using ATLAS.ti. 4.1. Questionnaire responses were grouped into three themes. The themes comprised categories and sub-categories as reflected in Table 6.1 below, which presents an overview of the categories, sub-categories and students' perceptions that were aligned to the sub-categories.

Categories in the social disposition theme involved perceived social relevance of the question and context to the lives of the students. These categories include aspects of task characteristics, in particular, task importance and value and task format, content and presentation.

Table 6.1 Overview of students' perceptions aligned to each category and sub-category

Theme	Categories and Sub-categories	Perceptions
Cognitive disposition	**Task characteristics**	
	Task complexity	
	Task easiness	Understandable questions Easy to answer questions Pictures helpful Successful performance
	Task difficulty	Questions are difficult to understand Questions are difficult to answer Questions are tricky Unfamiliar context Uncertainty of response
	Task importance and value	Development of intellectual skills Usefulness Learning of content Contributes to Careers
	Task requirements	Studying (revision of work done) Thinking
	Task format and presentation	Practical-realistic experiences Pictorial presentation
Affective disposition	Motivation	
	Intrinsic motivation	Easy to pass
	Extrinsic motivation	Easy to fail
Social disposition		Empathy

The information in Table 6.1 reflects that most of the students' perceptions were aligned to the characteristics of the questions. It also indicates that students' views about context-based tests varied.

6.2.2 General views on context-based questions

A general overview of students' perceptions of context-based tests and questions are presented prior to the descriptions of the specific perceptions. Students were asked to indicate which questions they liked or questions they disliked, their acceptance or non-acceptance of the use of contextualised questions for their assessment, as well as their general views about the tests. The reasons for their choices reflect the different specific perceptions of the students.

a) Students' liked questions (QQ1) and disliked questions (QQ3)

A general overview of students' liked or disliked questions is presented in Table 6.3 and Figure 6.1 for the Electricity test and Table 6.4 and Figure 6.2 for the Air and Living Things test. Where more than one question was listed as liked (or disliked), the questions were counted separately - as a result the total frequency of citations exceeded the number of students participating in the study.

Presented in Table 6.2 below is a summary of the questions from the Electricity test.

Table 6.2 Overview of questions from the Electricity unit test

Test question	Science content focus	Context
Question 1	Static electricity and its production	Observation of sparks on taking off a jersey
Question 2	Testing battery using voltmeter	Testing radio battery by a shopkeeper (picture used)
Question 3	Wiring circuits, choice and working of light bulbs, drawing electric circuit	Electrical wiring of a house (some diagrams required)
Question 4	Safe use of electrical appliances, interpreting ratings of appliances, electrical resistance, calculations of current and resistance.	Connection of multiple appliances of different ratings on one socket (picture used)

Table 6.3 Frequency of citations of liked and disliked questions by school and question from the Electricity unit test

Questions	Schools					Schools				
	Frequency for Liked questions					Frequency for disliked questions				
	M1	M2	M3	M4	Total	M1	M2	M3	M4	Total
All	1	2	2	1	**6**	4	0	2	4	**7**
Almost all	1	0	0	1	**2**	3	0	2	3	**8**
None	3	0	2	5	**10**	1	3	3	1	**9**
Q1	7	7	27	7	**48**	11	17	17	8	**53**
Q1 components	1	0	1	2	**4**	2	0	3	0	**5**
Q2	2	4	29	11	**46**	11	14	9	7	**41**
Q3	6	8	27	7	**48**	10	28	31	5	**74**
Q3 components	16	7	4	5	**32**	9	6	8	5	**28**
Q4	19	51	34	9	**111**	13	5	21	10	**49**
Q4 components	19	5	23	10	**57**	14	7	10	9	**40**
Drawings	1	0	0	0	**1**	0	0	0	0	**0**
Uncertain	0	0	0	0	**0**	0	0	0	2	**2**
No/non-response	8	1	1	1	**11**	13	5	7	3	**38**

(n=294 (M1 (A/B) =76; M2 (A/B) =74 M3 (A/B) =100; M4 = 44). Some students cited more than one question thus the inflated frequency of citations)

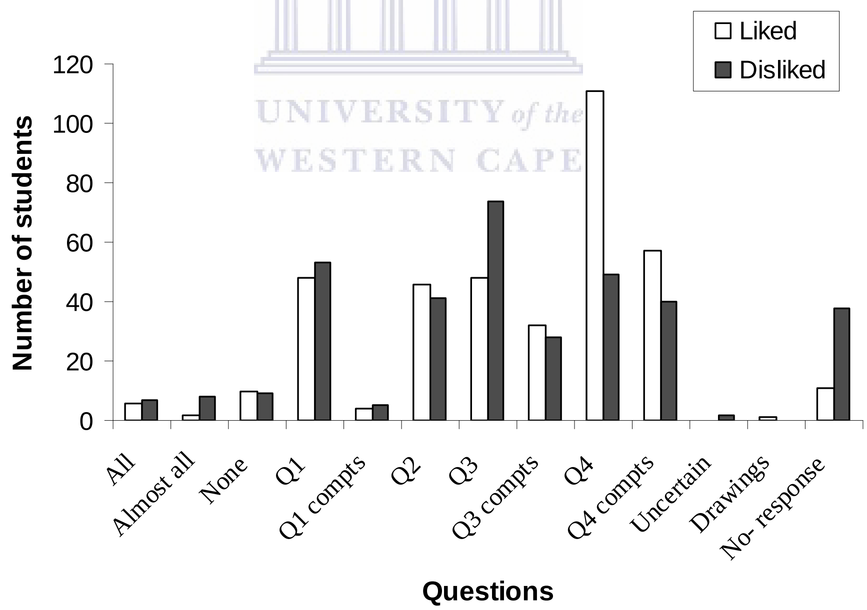


Figure 6.1 Comparison of citations of liked and disliked questions from the Electricity unit test

The data from Table 6.3 and Figure 6.1 indicate that Question 4 and its components were cited as liked more times (167 citations) than disliked (89 citations). Question 2 was cited as liked (46 citations) to a smaller extent as it was disliked (41 citations). Question 3 was only slightly more disliked than liked, with 74 citations on disliking it as compared to 48 citations for liking it. Question 1 and its components were cited as liked 52 times compared to being disliked, cited 58 times.

Table 6.4 below provides an overview of the questions from the Air and Living Things unit test.

Table 6.4 Overview of questions from the Air and Living Things unit test

Test question	Science content focus	Context
Question1	Conditions for seed germination	Delayed germination of oiled seeds
Question 2	Light and photosynthesis; starch test.	Farming practices - no crops under trees (pictures)
Question 3	Effect of exercise on breathing and pulse; differences between inhaled and exhaled air	Student participating in athletics competitions
Question 4	Rusting	Rust prevention in iron roof sheets by painting
Question 5	Properties of oxygen; combustion and respiration	Dangers of taking a Brazier into a house (diagram)

Table 6.5 Frequency of citations of liked and disliked questions by school and question from the Air and Living Things unit test

Questions	Schools					Schools				
	Frequency for liked questions					Frequency for disliked questions				
	M1	M2	M3	M4	Total	M1	M2	M3	M4	Total
All	2	3	6	1	**12**	5	0	0	8	**15**
Almost all	0	1	1	0	**2**					
None	4	1	1	10	**16**	4	4	19	2	**29**
Q1	8	7	31	5	**51**	14	9	7	7	**37**
Q2	11	14	38	9	**72**	3	4	9	3	**30**
Q2 components	5	7	1	1	**14**	0	4	1	0	**5**
Q3	5	7	8	2	**22**	13	13	33	6	**65**
Q3components	2	2	2	3	**9**	1	2	5	1	**9**
Q4	4	4	14	3	**25**	8	12	18	7	**45**
Q5	23	25	43	5	**96**	5	11	7	7	**30**
Q5 components	5	3	0	1	**9**	2	0	0	0	**2**
No/non-response	4	5	9	2	**20**	3	14	12	5	**32**

n =249. M1 (A/B) = 57; M2 (A/B) = 64; M3 (A/B) = 92; M4 = 36. Some students cited more than one question thus the inflated frequency of citations).

Data from Table 6.5 above and Figure 6.2 below indicate that there were more citations for liking Questions 5 and 2, and to some extent Question 1, than citations for disliking the same questions. For Question 3 and Question 4 there were more citations on disliking the questions (65 citations for Q3 and 45 for Q4) than citations for liking them (22 for Q3 and 25 for Q4). Questions that were listed as liked (or disliked) in a single statement were counted separately.

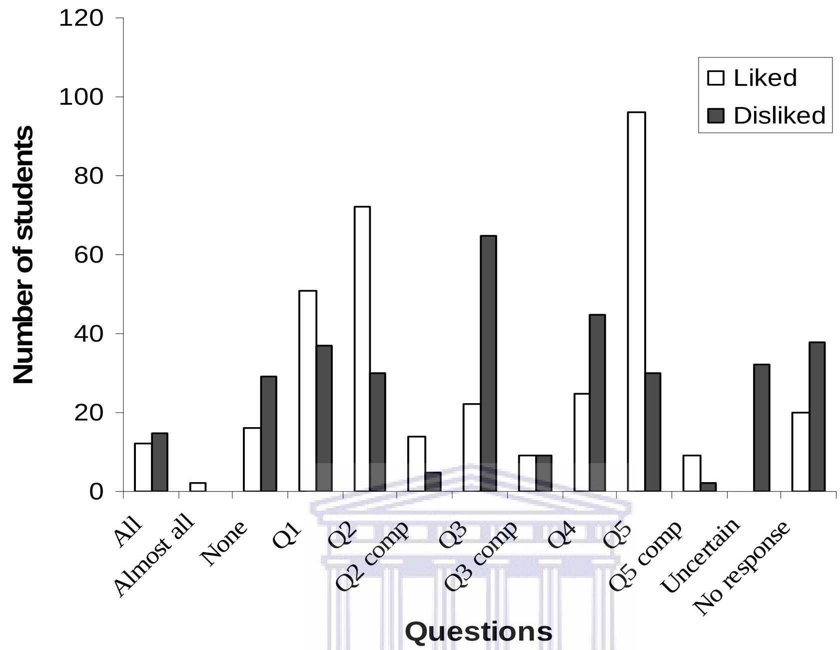


Figure 6.2 Comparison of citations for liked and disliked questions from the Air and Living Things unit test

It is evident from the data presented in Tables 6.3 and 6.5, as well as Figures 6.1 and 6.2 that all questions were liked by some students and disliked by others. Students stated a number of reasons for liking or disliking questions. These reasons matched four sub-categories of task characteristics, namely: task complexity; task format, content and presentation; task importance and value; and task requirements. Perceptions in each of these sub-categories are presented in Sections 6.2.3 to 6.2.8 below.

b) Students' acceptance or non-acceptance of context-based tests (QQ6)

Questionnaire Question 6 required students to indicate their views regarding their acceptance, or non-acceptance, of the use of context-based questions in future assessment in Science. The frequency and distribution of the citations for acceptance, or non-acceptance, of context-based tests are provided in Table 6.6 below.

Table 6.6 Frequency of citations of acceptance and non-acceptance of the use of context-based tests by school

Code	Schools								Total	
	M1		**M2**		**M3**		**M4**			
	E	**AL**	E	**AL**	E	**AL**	E	**AL**	E	**AL**
Accepts use	32	**31**	45	**36**	59	**60**	18	**14**	154	**141**
Do not accept	29	**8**	24	**6**	22	**10**	22	**13**	97	**37**
No comment	0	**2**	0	**0**	0	**1**	0	**0**	0	**3**
Uncertain of position	1	**6**	0	**4**	12	**7**	2	**0**	15	**17**
No Response	14	**10**	5	**18**	7	**14**	2	**9**	28	**51**
Total	76	**57**	74	**64**	100	**92**	44	**36**	294	**249**

It is evident from the data in Table 6.6 above that altogether, students had different views regarding their acceptance of the use of context-based topic tests, although there were more citations in favour of the acceptance of the use of context-based tests. There were 154 citations for Electricity and 141 citations for Air and Living Things that were made in favour of using context-based questions in science tests, while 97 and 37 citations respectively indicated that students did not accept the use of such tests. There were also students who did not comment, or were uncertain about their position on the matter, or whose statements were non-responses.

c) Students' evaluative statements on tests

Questionnaire Question 5 (QQ5) required students to indicate their views and feelings about the context-based tests in general. Some students' responses reflected evaluative comments about the tests. Other statements were non-committal although they appeared to suggest some ways of improving the questions. These suggestions were treated as recommendations and are presented in Section 6.2.8 below. An overview of the evaluative citations is provided in Table 6.7 below.

Table 6.7 Frequency of evaluative citations by school and unit test (QQ5)

Code	Schools and unit tests								Total	
	M1A		M2A		M3A		M4			
	E	**AL**	E	**AL**	E	**AL**	E	**AL**	E	**AL**
Fair test	0	**0**	0	**0**	2	**1**	0	**1**	2	**2**
Test okay	4	**2**	0	**0**	4	**2**	7	**0**	15	**4**
Good test	5	**4**	15	**10**	18	**10**	4	**4**	42	**28**
Nice test	0	**0**	1	**1**	3	**3**	2	**2**	6	**6**
Unfair	0	**0**	0	**0**	1	**0**	1	**0**	2	**0**

The majority of the evaluative citations were positive for both tests with a total frequency of 65 for the Electricity test and 40 for the Air and Living Things test. Data from Table 6.7 also reflect that there were students in all schools who found the tests to be good, while two schools had students who indicated they found the test on the Electricity unit unfair.

Students' perceptions of the questions and tests varied. Reasons and comments on students' choices of liked or disliked questions, acceptance or non-acceptance of context-based tests, or perceived quality of the tests provided students' perceptions and experiences of the questions and tests. The discussion of these perceptions and experiences follows the sub-categories outlined in Table 6.1 above, namely, task complexity, task importance and value, task requirements, task format, content and presentation, as well as motivation and empathy.

6.2.3 Task complexity

As mentioned in Section II of Chapter Four where students' perceptions of performance assessment are presented, task complexity was used in reference to how easy or difficult the students perceived or experienced the tasks to be when interpreting and answering them.

a) Intellectual accessibility of questions

Table 6.8 below illustrates the frequency and distribution of citations relating to aspects that students perceived to make the tests and questions to be intellectually accessible. The different ways in which students perceived test questions to be easy as indicated in different questionnaire questions (QQ), are presented.

Table 6.8 Frequency of task complexity related reasons for liking, appreciation or acceptance of questions

Perception Sub-category	QQ 2		QQ 5		QQ 6	
	E	AL	E	AL	E	AL
Understood question	47	43	0	0	0	6
Easy (question/test)	58	2	20	14	20	7
Learned (class/practical/elsewhere)	0	15	0	0	0	0
Familiar context/experiences	9	17	4	29	3	0
Successful performance	14	34	0	0	1	0

The data in Table 6.8 reflect that liking (QQ2), appreciation (QQ5) and acceptance of use of context-based tests in future (QQ6) were dominated by citations on perceptions related to how easy the questions were, and/or how familiar students were with the context of the questions. Perceived understanding of, and success in answering, the questions were also dominating reasons for students' liking of the questions. Students' liking of the questions, appreciation and acceptance of the test, thus reflected perceptions that questions were:

- easy and understandable
- focused on work done
- about familiar experiences
- successfully answered

In discussing students' perceptions, attempts were made to sample excerpts so as to represent participating schools, unit tests and the ideas presented by different students. This sometimes led to the use of several excerpts to illustrate perceptions.

i) Easy and understandable questions

Students' perceptions that questions were simple and understandable seemed to be based on the ease with which they interpreted the questions, or identified question requirements. The excerpts below reflect the perceptions that questions were easy to understand.

> I was able to understand the questions, what they wanted (EM3 2B 16 Q1, 2, 4).
>
> They are somehow simple and straight forward (EM3 2B 31 Q1, 2, 3).
>
> They are easy to understand and answer (ALM1 3A 30 Q1, 2).
>
> They are simple to understand and have clues of the answer (ALM1 3B 8 Q5).

All test questions had some students feeling that they were easy to work on. The last excerpt by ALM1 3B 8 Q5 further indicates a perception that Question 5 on the brazier (see Appendix IIIB) provided some clues for answering the sub-questions.

Students also liked questions they perceived to require less cognitive input from them, as the excerpts below illustrate.

> It was simple to do it; it never wasted my time and energy of busy thinking (EM4 27 Q2).
>
> Because I was able to draw only (EM4 30 Q2).
>
> I like these questions because there is no need for me to think a lot, its just things that are happening (ALM1 3A 12 Q5).

Perceived non-requirement of deep thinking in answering questions seemed to be linked to familiar real life occurrences to the students or the ease with which students produced answers.

ii) Successful answering of questions

Students' liking of questions seemed also to depend on their perceived ability and success in answering the questions. There were 14 citations relating to perceived success in answering questions from the Electricity test and 34 citations from the Air and Living Things test. Examples of citations relating to certainty of correctness of answers are provided in the excerpts below.

> It is because this question, I know that I will get it right (EM3 2B 43 Q2).
>
> Because it seems to be the most question I am sure I did it and if I don't get it well I'll hate it (ALM1 3A 5 Q3a (i)).
>
> Because I was able to answer it correctly without any problems (ALM3 3A 10 Q2, 4).

Students' confidence in their answers to the questions is indicated as another factor contributing to students' liking particular questions.

iii) Addressing work covered in lessons

Exposure to scientific concepts during lessons or relevant out-of-class experiences was perceived to be a prerequisite for working on context-based questions. Students liked questions they perceived to deal with work they were familiar with. Some students recognised that they had done the science content required for responding to the selected questions in previous lessons or other learning situations. The excerpts below present some examples of students' reasons for liking certain questions.

> We did the experiment and I understand it very well than the others (questions) (ALM3 3A 7 Q2, 3b (iv)).
>
> Because we did many practicals from it and it was easy to remember (ALM3 3A 8 Q2).
>
> Because we had done this experiment many times (ALM2 3B 27 Q5b, c).
>
> I understood it when the teacher taught me (ALM4 3B 2 Q2).

Lesson activities to which questions were linked, in particular practical work, seemed to have contributed significantly to students' perceived understanding of the content in the questions. Relevant experiences that made some questions easy for the students to work on, included work done in other subjects whence students could transfer relevant information to questions in the tests. The excerpts below reflect students' perceived contributions of knowledge from other learning situations, including the communication media, to answering question.

> It is because it was simple and its something we do in Agriculture (ALM1 3A 6 Q2).
>
> They talk about things we know and learn about in other subjects (ALM3 3A 40 Q2, 5).
>
> They teach us even in radios (ALM1 3B20 Q5b, c).

These statements also indicate that the context-based questions students responded to, allowed them to transfer knowledge from different sources to the science tests. This knowledge transfer was also observed by the teachers.

iv) Context familiarity

Context familiarity was used to refer to students' awareness or experiences of the real life situations used to present the questions. Real world contexts used in the questions had different degrees of familiarity amongst the students. Context familiarity seemed to contribute to the students' perception that questions were easy to understand and to answer. They felt that they were able to answer questions that focused on real life experiences by using knowledge from their experiences, without much appeal to learned science content. Thus, students liked questions that contained contexts they were familiar with, as the excerpts below demonstrate.

> Because it is not difficult when answering it because you have to use your own mind and knowledge from daily life (EM3 2B 10 Q3).
>
> Question 4 happens to our daily lives and it can be explained to a person easily (EM4 31 Q4).
>
> They are easy to answer from your mind; it's something we see almost all the time at our grand parents (ALM3 3A 22 Q1 2, 4, 5(b)).
>
> They are things we see around us and they are questions which an individual can be able to answer without any lessons being done (ALM3 3A 20 Q2, 5).

Familiarity with the contexts appeared to make students feel that they understood the questions, and that they could construct answers easily by using "informal" or general knowledge from their real life experiences. Students were of the view that answers constructed from informal knowledge were acceptable.

Some students distinguished informal and formal knowledge and recognised the need for both informal knowledge and formal knowledge from science lessons in answering the questions, as shown by the excerpts below.

> The test was not hard but you have to know things around you and have knowledge (ALM2 3A 17 QQ5).
>
> To myself (*sic*) this test is simple and straight forward but it's up to you how much knowledge you have in your mind and what do you know about the environment (ALM2 3B 16 QQ5).

These two excerpts indicate that the test on Air and Living Things was perceived to be simple, so that it was up to the students to succeed in it.

Students' perceptions that questions were easy seemed to be influenced by various experiences. Students' perceived the low difficulty of the questions to be associated with:

- understanding the concepts assessed in the question at the time they were taught;
- whether, or not, students have had previous encounters of the concepts assessed;
- students' recognition of links between issues in the context presented in the question and their real life experiences;
- students' ability to use both school and informal knowledge in generating answers;

- how related the content in the questions was to the content in other subjects or learning experiences; and
- students' confidence and certainty in the answers submitted.

b) Level of difficulty of questions

Students' perceptions of assessment through context-based questions and tests were also concerned with the level of difficulty of the questions. Table 6.9 below provides an overview of the perceptions and the frequency of citations for perceptions on task difficulty. These perceptions were derived from students' statements on reasons for their disliking particular test questions (QQ4) and other relevant statements from QQ5 and QQ6.

Table 6.9 Frequency of citations on question difficulty related reasons for disliking questions, dissatisfaction with tests and non-acceptance of tests

Perceptions	**Frequency by question and unit**					
	QQ4		**QQ 5**		**QQ 6**	
	E	**AL**	**E**	**AL**	**E**	**AL**
Did not understand question / topic	62	28	4	0	23	0
Difficult question	58	26	63	12	20	5
Difficult to answer	16	34	1	2	4	0
Tricky & Confusing questions	9	18	7	0	0	0
Need time to answer	1	3	0	0	0	0
Uncertain about response	4	7	2	0	9	4
Unfamiliar context	18	25	0	0	0	10
Unexpected questions	0	0	8	9	4	0
Do not know question/answer	9	0	0	0	0	0

It can be noted from Table 6.9 that most reasons for disliking questions were from questionnaire Question 4, which specifically sought for such reasons. Other relevant statements were also made in QQ5 and QQ6. Students' perceived difficulty of questions were concerned with the following factors:

- Problems of interpreting and understanding questions and concepts.
- Problems with constructing answers.
- Unfamiliarity with the real life contexts used in the test questions.
- Uncertainty of the nature and form of answers to be constructed.

i) Problems of understanding questions

Students disliked questions they perceived to be difficult to interpret and understand or confusing. For example, they justified their dislike for particular questions thus:

> … because it's difficult for me to see how to draw using figure not making new diagram (EM3 2B 39 Q2).
>
> … because I don't understand the question whether the seeds germinate on the oil. How can seeds germinate into the oil? Because oil is not good for germination, how does it germinate and take longer? (ALM1 3A 20 Q1).
>
> Because I have no idea on what they are asking me about (ALM2 3A 15 Q4, 5).
>
> They are very confusing and very hard to understand (ALM1 3B 8 Q1 2).

The statement made by ALM1 3A 20 Q1 illustrates that while the student seemed to understand the scientific content (conditions for seed germination) required for answering the question, s/he seemed to have had difficulty interpreting the question. The statement from EM3 2B 39 Q2 indicates that while the students needed to draw lines connecting the appropriate battery and voltmeter terminals, s/he wanted to draw a new diagram, possibly using circuit symbols. This may also indicate that this student had problems interpreting pictorial information or translating circuit symbols to real life circuit components or circuit components to circuit symbols. The disliked questions were liked by other students, who perceived them to be easy to understand and answer.

For some students difficulty of the test seemed to be associated with the different and unfamiliar format of the questions, for example:

> They need you to think, they are not questions that are formal, they are difficult to understand (ALM4 3B 18 All questions).

The importance of initial understanding of the concepts for perceived low difficulty of test questions was confirmed by perceived question difficulty that was associated with poor understanding of the concepts during lessons. There were instances where students disliked the questions because they experienced problems due to perceived low understanding of the concepts in a question, as shown in the excerpts below:

> I did not understand the experiment the time we were taught (ALM2 3B 36 Q1).
>
> Because it is difficult for me to understand what causes sparks (EM2 2B 22 Q1).

> Because I can't understand the question, I don't even know the reason why you do not have to connect many electrical appliances in one socket and in one extension cord (EM4 34 Q4).

Students' perceived level of difficulty of questions was related to their interpretation of questions, general difficulty of questions, unfamiliar format and a lack of understanding of concepts in question.

ii) Problems in constructing answers

Perceived difficulty, and therefore dislike or dissatisfaction with questions was also related to difficulties experienced in generating answers to the questions. Some students who felt they could not come up with the answers to the questions, seemed to think that those questions had no possible answers or explanations. The excerpts below illustrate these perceptions:

> They were unanswerable (difficult) (EM3 2B 13 Q3).
>
> It is not easy to understand and some of the things are just un-explainable (ALM3 3B 8 QQ5).

For other students, while they understood the questions, they seemed unable to formulate the answer or express their thoughts. For example, some students stated that:

> It is hard for me to think about it, I understand it but I don't know what I am going to write. I do not have ideas of it (ALM1 3A 24 Q1).
>
> I was not able to express the correct answers (ALM4 3B 2 Q5).
>
> Because I am unable to express my thoughts that I know about the question (ALM2 3B 4 Q5).

There might also be a problem regarding written language resulting in these students experiencing difficulty in formulating or expressing their answers, rather than a conceptual problem.

iii) Context unfamiliarity

For some students question difficulty was associated with the use of unfamiliar contexts in questions. Perceived low familiarity with contexts was presented as indicated in the excerpts below. Also indicated, were reasons for students' dislike of the questions.

> I have never seen such (especially question 1) (EM3 2B 17 Q1, 4).

> It is because we know little about it and I do not observe much of these things in my clothes (EM2 2B 29 Q1).
>
> It talks about something I have never experienced so I am not able to explain how she is going to solve her problem out (EM3 2B 48 Q4).
>
> I never heard about its answer (ALM1 3A 9 Q5b).

Perceived unfamiliarity with the contexts seemed to interfere with students' interpretation and understanding of questions, as well as the construction of answers.

Some students recognised that the degree of their familiarity with the contexts in the different questions varied according to their experiences. Thus, students acknowledged that some questions could be understood while others could not be easily understood, as the excerpts below illustrate.

> Some things are obvious since they are around us, however, some are found only in suburbs which then give us a problem. We have to imagine something you have not seen (EM3 2B 20 QQ 6).
>
> This test was very easy to pass and easy to fail. Meaning if you have never lived with old people or stay at home (rural area) instead of staying in urban areas then you would find it difficult (ALM3 3A 49 QQ5).

These students recognised the link between the test questions and their real life experiences, and the possible effect of the awareness of such a link on their achievement. Familiar or unfamiliar contexts were perceived to be strong determinants of the perceived level of difficulty of the tests. The excerpts also indicate the challenges some of the students met when working with unfamiliar contexts. Similar observations were made by students during their group interview as demonstrated in the excerpt below. The excerpt also reveals additional perceived impact of unfamiliar contexts on students' processing of questions and answers.

Victoria:	Okay. Now how do you feel about the topic tests that you were given? Writing a topic test that … uses questions about things that you see around you, context-based questions?
Norah:	It was okay. … But then we had a problem with a question in test on Air and Living Things. With this question on *immbawula* (brazier) [*Question 5 of test*], most of us live in town and we don't, are not used to *immbawula*. We just had to imagine that if that happens this happens. Most of us didn't know about this. (Mm) Fine we go and visit *gogo* (granny)...
Victoria:	So it required you to think.
Norah:	Yes, to think a lot. And just imagine that it happens like that. Even if you imagine it in a wrong way you had to write the wrong staff.
Victoria:	But if you look at the questions they were asking about the science …

Norah: Yes. But …

Ann: I think Ma'am it, in a way if I knew science, 'cause I don't know science. I fail Science. But if I knew Science it was going to make things easier for me because I get to put what I know in Science together with what they are asking, like general things. If I wore a jersey what happened if sparks are seen. So it makes me think. It makes things easier for me to imagine what really happened plus what I know in Science. I think it helps asking things like things we just see around us. It makes things easy for you to answer questions.

Nandi: But if it's like you've never seen this thing *immbawula* and then okay you know Science if you've never seen it you don't know how it works and anything else about it, then obviously you are not going to get anything because you don't know anything about this thing here.

Victoria: But if you know the situation that is being used …

Norah: Then that would be okay because you know what happens if you do this to *immbawula*. Then you are able to answer the questions, to apply like what you know about *immbawula.*

Victoria: Okay. Even though there was a picture to help you see what the situation is, do you feel it didn't help you?

Norah: No-o-o. It didn't. Okay, fine we see that it's a tin with fire inside but then you had to apply. Some questions were asking about what would happen inside if they all slept [*reads question and Victoria helps out*] (Victoria: Give two reasons why Sonile thinks taking it inside the house is a dangerous idea?) Ya the picture did help. I take it back, it did.

Victoria: But if you look at the question it says what two things are necessary for it to provide warmth?

Norah: Ya, it needs us to apply like science.

Victoria: Yes. So what does it need? It needs wood or coal and oxygen. You have learned about oxygen (Students: Yes) … and if these people take it inside the house and the house is small it has a small window

Norah: That is the more thinking. We had to do a lot of thinking, like if we didn't know it. Yes we studied the picture but if you knew it, *immbawula ah tintfo letincane leto* (brazier, ah those are small things).

Victoria: But even though there is, I mean you needed to think.

Norah: But those who knew the *immbawula* thought for 15 minutes and then *laba labangayati li*-hour (those who do not know it an hour) (M4 Interview).

Also emerging clearly from the excerpt is a perceived concern, by the students, of being subjected to testing conditions involving unfamiliar contexts that constrained their processing of information. In addition, unfamiliar contexts were perceived to place higher intellectual demands on the students, as well as delay their processing and answering of particular questions. While diagrams used for the contexts were useful to some extent, familiarity with the context was perceived to be more beneficial. When familiar contexts were used whatever students imagined or thought of could be easily linked to some known aspect of the context. Unfamiliar contexts were perceived to demand deep thinking from the students to interpret questions and produce answers.

Students' familiarity with a particular context was thus related to direct experience with the context such as practicing the activities in the context, or observing others engaging in those activities. The excerpts also indicated that unfamiliar contexts required a good deal of visualisation by the students. Pictures and diagrams seemed to facilitate the visualisation process, but were not a replacement for familiarity.

iv) Uncertainty about contexts and answers

Coupled with context familiarity were students' concerns about the authenticity of contexts and the format and content of their answers to context-based questions. Some students perceived the reality of contexts to be inconsistent between different questions, as the excerpt below indicates:

> What I do not like about these questions is that sometimes things like this one are not usually the case but some are (sic) (ALM3 3B 16 Q3, 4).

The student seemed to question the way things happen in real life and the reality of the situations described in Questions 3 and 4 from the Air and Living Things test (see Appendix IIIB). It was, however, not clear which question created a feeling of doubtfulness.

Some students felt uncertain about how to compose and present answers to the questions, and this uncertainty led to students disliking certain question(s).

> Because I do not know what to write (EM4 17 Q All questions).
>
> Because I don't know how to answer it (EM2 2B 35 Q3).

The uncertainty of how to phrase answers gave students mixed feelings about the acceptance of context-based questions and tests. Mixed feelings from one student were expressed in the following way:

> I would partly agree and partly disagree. I would agree because they may ask things I have done or see people doing then it becomes easy for me and I would disagree because I would give a general answer while the teacher would expect something from the book and specific (ALM3 3B 15 QQ6).

Statements made by the students indicated that students might have experienced a dilemma. A number of students realised that they had to distinguish teacher expected scientific answers from possible answers that they understood or knew from real life experiences.

Despite the challenges experienced by some of the students, there were those who felt that the tests were neither difficult nor easy. Other students felt partly responsible for the difficulty they experienced in answering questions. For these students the difficulty was attributed to the factors presented below. Illustrative excerpts are provided for each factor.

- A lack of concentration and attentiveness during lessons:

> I think this test is easy but I think the majority of us will fail it because it is difficult to answer and some of us did not concentrate during the practical class (ALM2 3B 11 QQ5).
>
> It is not difficult or easy, but it is moderate. This test needs more concentration in class practicals and attentive listening (ALM3 3B 33 QQ5).

- Insufficient scientific knowledge:

> It is not a hard test; you just have to know properties of i.e. oxygen, sun, etc. and it’s easy to answer any other questions (ALM4 3B 20 QQ5).
>
> All in all, this test was fine, it was not difficult and it was not also easy. It was just fine except that I don't know Question 3 (ALM2 3B 6 QQ5).

- Insufficient studying or revision of work done:

> It was an interesting test not hard not easy, but easy to those who studied and hard to those who never studied or listen during class times in Science (ALM3 3A 14 QQ5).
>
> This test was all about studying it and passing it (ALM3 3A 24 QQ5).

The level of difficulty of the Air and Living Things test was perceived in terms of the students’ responsibility and effort put into the tests. It seemed students recognised that working on their attentiveness and concentration during lessons and their commitment in revising their work (studying), was likely to equip them with sufficient knowledge to make the tests easier.

v) Weak link to class notes

Another concern for students was the extent to which tests focused on the science content covered during lessons. Some students indicated that they needed to study from their notes in order to pass the test. So, the link between the notes (provided by teacher or constructed by student) and the tests was important for these students. Their concerns are demonstrated in the following exemplifying excerpts:

> It was not good, there are a few things that we learn in class, it needed a lot of mind, we studied what was in the notes, but it was not here (ALM1 3A 4 QQ5).

> This topic test is not what we wanted and expected in this topic there is much left behind and may you please give us a test on what we have learned about in this topic, simple questions and not this test (ALM1 3B 21 QQ5).

> *Ubosikhiphela lesakubhala kuma notes hhayi kutsi ukhiphe lokutsandvwa nguwe. Kantsi lama notes ani uma singasabhali ngalokusemanotsini? Yekela kusikhiphela ema*-experiments (sic). (Tests should focus on information from our notes not what you like. What is the purpose of the notes if we do not write what is in the notes? Desist from using questions on experimental work) (ALM1 3A 17 QQ5).

The excerpts above also indicate the extent of students' convictions that the content of the tests captured less of the work they had done during lessons. The perceived absence of work covered during lessons and unexpected question type, made the questions appear difficult for the students. The third excerpt by ALM1 3A 17 QQ5 indicates that the student recognised that the tests captured aspects of practical/experimental work they had done, although the student could not identify any link to their notes.

Students' perceived omission of learned material from the tests could be an indication that some students experienced a difficulty in linking what they had learned to the contexts used in the tests and therefore assumed that there were no links.

In summary it can be noted that different students perceived the level of difficulty of the questions in different ways. For some students questions were easy, but they needed to raise their level of concentration and pay more attention during lessons and when they were studying their notes. For other students, those questions were perceived to be difficult to understand, confusing or to have omitted information from their notes. Context familiarity was an important factor in perceived level of difficulty of the questions and it affected the time students took to process answers with certainty and confidence.

6.2.4 Task importance and value

Some students perceived the context-based questions to provide some educational and social importance and value. Indicators for these perceptions and the frequency of their citations are presented in Table 6.10 below.

Table 6.10 Frequency of citations related to perceived importance and value by questionnaire question and test

	QQ 2		QQ 5		QQ 6	
Perception	**E**	**AL**	**E**	**AL**	**E**	**AL**
Provides learning experiences	44	40	28	11	23	10
Encourages thinking	0	11	0	3	5	5
Applies to real life	23	0	0	4	0	0
Contributes to Career	1	0	1	0	2	0
Improves understanding	0	0	0	3	7	0
Encourages observation	0	0	0	3	0	11
Tests different abilities	0	0	0	2	0	6

It is evident from the data in Table 6.10 that perceptions that were associated with the importance and value of the tests and questions were dominated by learning experiences. The other perceptions: promoting thinking, improving understanding, encouraging students to be observant in their surroundings, testing of different abilities, relevance to real life and awareness of career opportunities, were presented to varied extents for the two tests and the three questionnaire questions (QQ2, QQ5 and QQ6).

a) Learning experiences

Students indicated that they liked certain questions (QQ2) because these presented information they had not encountered before, or they learned new skills from the questions. So, while the real life contexts presented in questions may not have been familiar for some students, some of those contexts were perceived to provide new learning experiences. Not only did students' statements reflect that they became aware of things they had not been aware of before, some contexts provided advice on how certain things were to be done. The following excerpts reflect the perceived learning benefits from questions students liked.

> Because I can be able to connect when I am at home and help others who don't know about electricity (EM1 2A 2 Q4).
>
> For No. 2 it is something I know and I saw it. No. 5 I never made fire on a brazier and I did not know it was not suppose to be in the house and I think Vuy'sile's speech teaches me something (ALM1 3A 29 Q2, 5).
>
> I like it because they teach us about dangers that happen at home and how to grow good crops (ALM2 3B 42 Q2, 5).
>
> Because it gives us advice of what to do when using iron sheets (ALM2 3B 10 Q4).

Statements about learning experiences from the Electricity unit test were associated mostly with Question 4, and they indicated that students became aware of:

- appropriate ways of connecting electrical appliances to sockets;
- malpractices in the use of electric sockets and the possible dangers that could occur; and
- the need to change some of their behaviour towards electricity and its uses.

The excerpts below illustrate these points:

> I like it because most people in homes use their socket the same way as it was shown and this gives me a lesson to tell people that it is not safe (EM1 2A 31 Q4).
>
> Because it gives me knowledge to know that it is wrong to use one wall socket for a lot of appliances (EM3 2A 9 Q4).
>
> I felt great because it teaches me more about how to use electricity and how to use electric things and the safety of using them (EM1 2A 31 QQ5).

Further support for liking questions because of the learning experiences was provided by responses to QQ6. Students who appreciated the learning experiences from context-based questions shared the following views for accepting the use of such questions in future tests:

> Because I learnt more things about what is going on around me, if I see people not doing good I would be able to correct them (like the socket) (EM2 2A 32 QQ6).
>
> I would be happy because it helps us to see wrong things and right things that we do. It also helps us to know further more about scientific things around us not just scientific things that are in the book (ALM3 3A 26 QQ6).
>
> I would be very happy because these tests add new learning skills. It makes me enjoy Science and they make it easier (ALM3 3A 3 QQ6).

Learning experiences from students engaging with the context-based tests were thus concerned with:

- extending learning opportunities beyond the classroom;
- facilitating an awareness of inappropriate practices or more acceptable practices relating to electricity and living things;
- encouraging students to be more observant in their environment;
- increasing student enjoyment of Science;
- broadening students' learning in Science to include real life experiences; and
- encouraging students to keep pace with scientific changes in the world.

Perceived learning experiences seemed to be for personal development and empowerment, as well as for the benefit of the community in which students lived. Students felt they could share information gained from questions with other people.

Learning through context-based questions may also indicate students' perceived advantage of the use of such questions for assessing learning in Science.

b) Promoting thinking and understanding

Students also liked or accepted the use of context-based questions and tests because of a perception that they encouraged them to think and improved their understanding of the concepts assessed.

The first three excerpts, below, present perceptions that questions encouraged students to think, while the last one is on perceived promotion of understanding:

> They are making us practice thinking. And it is a good thing to think about something (ALM4 3B 30 All questions).
>
> Because such tests train our brains to think as scientists (EM1 2B 6 QQ6).
>
> Because the questions seem to teach us how to think, they require a lot of thinking (EM4 36 QQ6).
>
> Because the topic somehow easily gets into your mind, you easily understand it (EM2 2B 16 QQ6).

Teachers also made similar observations regarding the contribution of context-based questions in the training of students to think.

c) **Encouraging observation of surroundings**

Data presented in Table 6.10 above also indicated that for some students the test on Air and Living Things encourage them to become more observant of occurrences in their environment, and to think about those occurrences. Some students felt that by doing so, they would learn much from their observations, become familiar with certain occurrences and practices, as well as improve their understanding and answering of the questions on those events. The following excerpts help illustrate some of these perceptions.

> This question helps me to think about things that are around. It also helps me to go and search why things are happening in that way (ALM1 3B 13 QQ5).
>
> It is a good idea because we as students we like to concentrate on what we are taught in class not looking at the situations we see around us or the environment. At times you can find that some students don't know simple questions that need their common sense just because they don't look at the environment (ALM2 3A 3 QQ6).
>
> I would be happy because they make me think more about what I see around and learn more about it (ALM3 3A 11 QQ6).

Students perceived the questions to increase their interest and motivation to be observant. This perception seems to indicate a possible extension of students' learning from the contexts in questions to learning from similar encounters from their environment.

Not all students found the questions to be of value to them. Some students did not like certain questions because they had no perceived usefulness to the students. Such perceptions were reflected in statements such as the following:

> Because I do not know what is its usefulness (ALM1 3A 16 Q1).
>
> They talk about things which I don't gain anything from them (ALM2 3A 9 Q3, 4).
>
> Because they are not important (ALM2 3A 19 Q1, 3, 4).

These excerpts seem to suggest that as the students worked on questions they may have also looked for useful information contained in those questions.

6.2.5 Task format, content and presentation

The focus of the questions in terms of students' experiences, learned material and information from their notes, as well as how it was presented, seemed to influence how students received the questions. Students' perceptions were related to how contexts were linked to everyday experiences and school science, as well as pictorial information.

a) Links between scientific concepts from lessons and question context

Students liked and appreciated the use of questions they perceived to be practical and realistic in relation to their lives, whether visibly in their present situation or their plans and hopes for the future. The test on Air and Living Things was perceived to have applications to real life. Thus, students indicated they liked questions from this test because:

> These questions make you think about what happens in your body and look at what happens may be at our homes (ALM3 3A 35 Q3, 5).

> Question 5 is practical/something we can see and observe. Question 3 is about what I want to study about when I finish school. (i.e. doctor) (ALM3 3B 13 Q3, 5).

> Tests like this one are much better because we are asked about ordinary stuff, things that we see in our everyday lives, not only what we learn about at school, but we know about also from our homes (ALM4 3B 35 QQ5).

These students recognised and appreciated the links between the context in the questions and their real life experiences. However, there were some students who experienced difficulty in making connections between scientific information from school and information required for answering the questions. An example of a student's difficulty of identifying scientific concepts learned in school (effects of exercise on breathing and blood circulation) from a context involving a school athlete, is presented below.

Aspect of questionnaire question	***Responses***
1. Questions liked	*All the questions.*
2. Reason for liking questions	*I like that we have learnt all the questions but don't know all the answers.*
3. Questions not liked	*I hate the question where they ask you about Jabulile racing questions* (Q3).
4. Reason for not liking questions	*I don't know anything about racing, but we have learnt about breathing, the questions just need much time to answer them.* (ALM2 3A 2).

This student attributed the problem to the time available for working on the question. However, the general view of the test by other students seemed to confirm the problem of linking the physical exercise of racing to the physiological and physical processes involved in breathing and blood circulation. The excerpts provided below illustrate further the perceived absence of a link between work done in class and the content of tests.

> It wasn't fair because many of the things we learnt about them, but some questions were not fair like question 3, you were supposed to ask questions about something that we learned in class not Jabulile's stories (ALM3 3A 29 Q3).
>
> Because I am not a runner I do not know what happens when a person runs (ALM1 3B 16 Q3).

These students seemed not to recognise the scientific concepts in the context and believed that they had not done the content in class. It may also be possible that the context interfered with students' interpretation of the question and hindered students from noticing the link.

b) Pictorial representation

The use of diagrams or pictures seemed to be received with mixed feelings by the students. While some students felt that there should have been more diagrams (than the three per test) others felt there were too many diagrams. The use of diagrams was perceived useful in facilitating the interpretation and understanding of the context and therefore the questions. They helped the students to visualise or remember the situations referred to in the context. The excerpts below illustrate students' reasons for liking test Question 5 (first excerpt) and acceptance of the use of context-based questions in science tests, on account of the pictorial representation (other excerpt).

> Because when we see the pictures it is easy to think about the problem which needs to be solved (ALM1 3B 16 Q5).
>
> I can be happy as I have said above that it can help us understand what is being asked than answering something that we don't see with our naked eyes or may be we did it in the lab we sometimes forget and when the question is put with the picture then the mind remember what we were doing in the laboratory (ALM2 3B 16 QQ6).

In addition to diagrams or pictures showing what the question was about, students also seemed to find that questions with diagrams may require brief and short answers.

Other students did not appreciate the use of pictorial representations because they felt confused by them or they lacked the skill of drawing, even so in the case of simple diagrams like circuits. The excerpts below provide some indication of students' dislike of questions because of problems experienced with pictures/diagrams:

> Because I don't understand how the drawing would be like (EM1 2A 12 Q2).
>
> Because of drawing the circuits (EM3 2A 16 Q2, 3(d)).
>
> Because they want me to draw and I am not good at drawing (EM2 2B Q2, 3).

The above observations indicate that perceived abilities of interpreting pictorial representations or presenting them differed among the students, and that they affected students' interpretation and answering of the questions.

6.2.6 Task requirements

Students noted that the context-based questions assessed a different form of learning, possibly learning that required "hard" work, attentiveness during lessons, careful reading of questions and the practice of being observant in surroundings. They also identified thinking, general knowledge and studying as requirements for working on the questions successfully.

a) Thinking

It was noted in Section 6.2.4 above on task importance and value, that the questions were perceived to encourage and develop students' thinking. Students also perceived the questions to demand such thinking when interpreting the contexts, as well as integrating learned information and information obtained from out-of-class activities. The need to think was also noted in students' group interviews. Students' views regarding their perceived demand for thinking are illustrated in the excerpts below.

Students liked questions because:

> They make me think very hard about what the answers could be (ALM3 3A 6 13 Q1, 2, 4).
>
> Because it needs the mind to be active and not think about what you learnt in class only (ALM1 3A 21 Q2).

Students disliked questions because:

> The questions are difficult and need a lot of thinking (ALM4 3B 22 All questions).
>
> They were very hard and needed a lot of thinking and understanding (EM4 2B 1 Q1, 3, 4).

Students did not accept the use of context-based tests in future tests:

> Because when you use situations you see around you, that need you to think about it, rather than studying question in your exercise books (EM2 2B 2 QQ6).

Because some science topics need the brain to think so most of them are difficult (EM2 2B 36 QQ6).

On my other side I would not like it since I have to think deeply and find that what I have thought is wrong. I would also feel happy at the very same time since I would learn to be observant and have more information about the surrounding (ALM3 3A 36 QQ6).

The above excerpts reflect some concurrence in the perception that the questions demanded deep thinking from the students, irrespective of whether they liked or disliked the questions. The perceived demand for thinking seemed to be a deterrent in the acceptance of context-based questions in future tests.

Perceived demand on the students to think deeper about the question content was extended to the need for scientific thinking. Students indicated a non-acceptance of the use of context-based questions on account of their need for scientific thinking, for example:

Because it is very hard for us to think in a scientific manner (EM4 11 QQ6).

Other students seemed to appreciate the opportunity to use their scientific knowledge in answering the questions, as illustrated below:

It is a simple and straight forward test and it just needs us students to think and use our scientific knowledge to answer the questions in this test (ALM2 3B 1 QQ5).

These are interesting tests because it is where we apply our scientific knowledge on things we see in our environments (ALM3 3A 20 QQ5).

It is okay because it is based in things around us. You take the traditional knowledge, you have to combine it with the scientific knowledge then you will definitely know the reason (ALM4 3B 23 QQ5).

The last excerpt indicates a perception that merging relevant informal knowledge and scientific knowledge provides deeper understanding of the use of scientific knowledge in students' real life experiences.

b) General knowledge

Students also perceived that questions required general knowledge to answer them, which also made the questions appear understandable and simple. Some students noted that they liked some questions for the following reasons:

Because that question is a general question, I am used to it (ALM2 3B 41 Q5b).

> Because you are able to write what you know generally (ALM3 3A 15 All questions).
>
> The fact that these questions just need common sense (ALM4 3B 12 Q5).

Those who felt that they could accept the use of context-based questions in future tests stated that:

> I would be happy because it would be easy for us to answer these questions. They are more like general knowledge questions and things that we do and see almost everyday (ALM3 3A 1 QQ6).
>
> It was a simple test and only needed common sense, nothing more than that (ALM1 3B 8 QQ5).

Students had a perception that the possible use of informal knowledge made questions less intellectually demanding.

c) Studying for the test

Students presented two apparently disagreeing views regarding the need to revise or study for the tests. Some students felt that the tests needed to be studied for, as the following excerpts illustrate.

> If you did not study you find it difficult, but if you did it's not easy but less difficult (EM4 15 QQ5).
>
> I think this test is too hard and needs us to study more harder *(sic)* (EM1 2B 9 QQ5).
>
> The test is okay and I really enjoyed writing it, but I will fail because I didn't get enough time to study (ALM4 3B 16 QQ5).

Others felt that the tests were not the kind of tests that could be studied for, as the excerpts below demonstrate.

> I think you must set other tests, which are in our exercise books that we can study them not this one because you cannot study these things (EM1 2A 6 QQ5).
>
> It wasn't a test you could study for, all you need is general knowledge (EM3 2A 4 QQ5).
>
> This test is not much difficult. What we have to do is to use our own understanding, if you do not understand anything in class you can not answer this test, there is nothing to study (ALM2 3B 26 QQ5).

Students' perceived requirements for the tests were thus concerned with the demand for students to think deeply about the questions and to understand concepts taught during lessons. For some students thinking and the use of informal or general knowledge in answering questions, and no revision or studying of work covered was sufficient. For other students studying from their lesson notes was important. These different perceptions about the need to study for the tests seem to indicate that there were questions in the tests that were perceived as requiring the need to be studied for, as well as questions that needed general knowledge and less studying.

6.2.7 Empathy

Students' empathy for human characters in the contexts seemed to be a perception that was associated with perceived authenticity of the contexts used in the questions. Some students became emotionally affected by the circumstances of the human characters found in the contexts. As a result questions were liked or disliked for emotional reasons.

Some questions were liked because students appreciated the knowledge position of the human characters in the question, as the excerpts below illustrate.

> It is that Londiwe shows that she knows about Electricity (EM3 2B 36 Q2)
>
> Because there is help that you may give to Thandeka that you must not connect many things at the same time (EM2 2A 34 Q4).

Other questions were disliked because of some perceived danger of the situation described for the people or some unintelligent behaviour by other people. Thus, questions were disliked for reasons such as the following:

> I didn't like the way Thandeka connected her extension, its dangerous (EM2 2A 18 Q4).
>
> Because I don't know how a person can make a mistake of wiring a house with a series circuit to all the bulbs without thinking first (EM3 2A 28 Q3).
>
> Because sometimes when you are running fast and your heart beats faster sometimes you may collapse and collapsing is dangerous (ALM2 3B 12 Q3).
>
> Because many people died of putting a brazier inside the room to make them warm (ALM2 3B 19 Q5).

The second excerpt may also indicate the student's doubt of the context in Question 3 from the Electricity test. This student questioned the behaviour of the person in the context. The last two excerpts seem to indicate some painful experience the students or someone they knew, may have encountered.

Students' concern about the safety of characters in the questions is clearly evident in the above excerpts, particularly where humans appeared to be in danger.

6.2.8 Recommendations on the use of context-based assessment

Students' acceptance or non-acceptance of context-based questions and tests has demonstrated to be linked to their perceptions in the sub-categories of task complexity; task importance and value; task format, content and presentation; task requirements; and motivation. Table 6.11 below presents a summary of the frequency of citations for acceptance or non-acceptance of context-based tests and questions in relation to students' perceptions.

Table 6.11 Frequency of citations for acceptance and non-acceptance of context-based tests in relation to sub-category, students' perceptions and tests

Sub-category	Perceptions	Acceptance		Non-acceptance	
		E	AL	E	AL
Task complexity	Difficulty level (difficult)	8	1	20	5
	Easy	20	7	0	0
	Difficulty to answer	0	0	4	0
	Degree of Understanding	0	6	23	0
	Not learned	0	0	7	0
	Context familiarity	1	0	0	0
	Unexpected questions	0	0	4	0
Task importance and value	Develops open mindedness	1	0	0	0
	Career opportunities	2	0	0	0
	Encourage thinking	5	3	0	0
	Improves understanding	7	0	0	0
	Learning experiences	23	10	0	0
	Test different abilities	6	0	0	0
	Prepares for examination	3	0	0	0
	Encourages observation	0	10	0	0
	Provides chance to apply	0	2	0	0
Task format, content and presentation	Unusual format	0	0	1	0
	Mixed cognitive level	1	0	0	0
	Covers work done	6	5	0	0
	Pictorial presentation	3	1	0	0
	Focus on everyday experiences	1	51	0	0
	Combine context/non-context questions	0	2	0	1
Task requirements	Common-sense knowledge	0	0	1	0
	Lot of thinking	0	0	9	2
	Studying	3	3	4	0
	Encourage observation	1	0	0	0
Affective disposition-motivation	Pass/fail	2	13	7	4
	Enjoy Science	2	0	0	0
	Uncertainty	0	0	1	4

E= Electricity test; AL= Air and Living Things test

It is evident from Table 6.11 that there were varied and mixed perceptions among the students regarding their acceptance and non-acceptance of the context-based tests. The data in the table indicate that most of the perceptions were concerned with task characteristics, and that the perceptions were in favour of the acceptance of the use of context-based tests. The data also reflect that reasons for the acceptance of the tests were more related to the perceived importance and value of the task, and the task format, content and presentation. Non-acceptance of the tests was more associated with task complexity and to some extent task requirements. Motivational reasons were also cited for acceptance and non-acceptance of tests.

Relevant excerpts from QQ6 have been used in some of the sections above to illustrate students' perceptions presented in those sections. A few excerpts are presented in this section to further demonstrate the perceptions on which acceptance or non-acceptance of context-based tests and questions were based.

a) Task complexity

Favoured aspects of task complexity were associated with the ease with which students worked on the questions, with a frequency of 27 citations for both tests. There were also 9 citations of acceptance of the use of context-based tests even though they contained some difficult questions. Reasons for the acceptance of context-based tests that were related to perceived level of difficulty of the questions, were presented as follows:

> I would like it because I am able to understand them well (ALM3 3A 33 QQ6).
>
> It is because that can help me to understand them and may be I can help the society living around me (EM2 2B 8 QQ6).
>
> Most of the questions are difficult, few are easy, but I would like to know more about them (EM1 2B 8 QQ6).
>
> Because some questions are difficult and some are easy (EM2 2A 18 QQ6).

Context-based questions were accepted because they were perceived to be easy to understand or for contributing to further understanding of context-based tests. They were also perceived to provide students with knowledge they could share with other people.

Non-acceptance of the use context-based tests for reasons related to perceived task difficulty was more aligned to the test from the Electricity unit. Illustrative reasons for non-acceptance that were associated with perceived difficulty or low understanding of questions are provided in the following excerpts:

> We might fail because some questions are easy, but you don't know how to answer them in the way the teacher has told us (EM1 2A 8 QQ6).
>
> Because some of the topics are difficult and some I didn't understand (EM2 2A 36 QQ6).
>
> I would not like it because some other things are hard to answer (ALM3 3A 12 QQ6).
>
> I wouldn't agree because in some situations we do not know the scientific reasons behind what is happening (ALM4 3B 23 QQ6).

Uncertainty of how to answer the questions, difficulty of producing answers and failure to understand questions, were the main reasons for non-acceptance of the tests.

b) Task importance and value

Students who accepted the tests for reasons linked to task importance and value, gave reasons such as the following:

> It reveals your brain and your capability of thinking and imagining things. (Which of course need a lot of time to think) (EM1 2B 19 QQ6).

> Because some of them are used to train our mind to see how much we know about electricity (EM1 2B 29 QQ6).

> I would be happy because we would not only be learning science, but also applying it in a way (ALM3 3A 44 QQ6).

> I would somehow be happy because there are things I know about the environment, the life of plants and animals and I can also be in a position to help the community around me by using the knowledge I have gained in the topic test because for sure if I make a mistake I would be helped by the corrections I will do after the topic test (ALM2 3B 9 QQ6).

Thus, some of the valuable aspects of context-based assessment that emerged from students' statements including the above excerpts, were:

- assessment of different abilities of students;
- cognitive development among students;
- opportunity to use scientific knowledge in real life situations; and
- opportunities to modify knowledge and share such corrected knowledge with others.

c) Task format, content and presentation

While the students accepted the use of context-based questions in tests, they made certain recommendations regarding their construction, for example:

> Not all questions must be asked on things we see around, but some must come just from the book because we live in different places so the things that we see around are not the same (ALM3 3A 29 QQ6).

> It would be okay in some other way, but sometimes it won't be okay because we study what is in the notebook for nothing and in the test it does not come out (ALM1 3A 27 QQ6).

It is interesting to note from the first excerpt the recognition that students have different backgrounds. Students recognised that familiarity with the context by different students may vary, whereas the information presented in their books or notes is the same for all students. These excerpts seem to suggest the use of context-based and non-context-based questions in the same test to allow students of different abilities to answer some questions. They also suggest that tests should be matched to the work done in lessons. Teachers made similar recommendations of mixing non-context-based and context-based questions in tests.

d) Task requirements

Reasons for the acceptance or non-acceptance of context-based tests on the basis of a need to study for the tests varied. Some students accepted the use of context-based tests because they perceived the questions to require less studying, for example:

> I would say yes. Because they do not need to be studied and they would be very simple as this one (ALM1 3B 8 QQ6).

An opposing view, for example was:

> No. Because we then have a lot to study (EM4 3 QQ6).

There is, nonetheless, a clear message that students would not accept the tests because they demanded a lot of thinking. Data from Table 6.11 reflect 11 citations indicating non-acceptance of context-based tests and no citations for acceptance because of the thinking demands of the questions. Students were against the use of such tests for reasons such as the ones reflected in the following excerpts:

> No. Because when you use situations seen around you, that need you to think about it, rather than studying questions in your exercise books (EM2 2B 2 QQ6).

> No. Because it makes you think very hard and you may end up having a headache (EM3 2B 15 QQ6).

Students seemed uncomfortable with the intellectual demands of the questions to support their use for assessing Science.

e) Motivation

Perceptions guiding the recommendations in the sub-category of motivation were mixed regarding students' passing and failing of the tests. Some students were of the view that they would pass the tests easily, while others feared they would fail. Reasons for these views revolved around the use of real life experiences. Students who accepted context-based tests on the basis of the perceptions that they would pass the tests, presented reasons such as the following:

> I could say that it would be fine because people would pass the tests because they had experienced that thing they are asked about (ALM3 3A 19 QQ6).
>
> It will be better because everyone would know the things happening around us and won't even fail to answer the questions when asked (ALM2 3A 9 QQ6).

Students perceived the tests to be easy to pass because they could use information from real life experiences (represented by 15 citations for the two tests). This perception appears to be in agreement with the acceptance of the tests for their emphasis on everyday experiences (with 52 citations and no opposing views).

Those who feared failure articulated their reasons as the following examples illustrate:

> It is because we would fail it. It is better if we are using the information from our notebooks than using situations we see around us (EM12A 1 QQ6).
>
> Because the questions are so tricky I can find myself failing (EM1 2B 38 QQ6).
>
> I would really get less that 70% because it would be talking about daily activities or practices (ALM3 3A 40 QQ6).
>
> I think I won't like it because we study theory work than general knowledge. I think sometimes it may make us fail (ALM3 3A 46 QQ6).

The ideas presented in the excerpts indicate a different view about the use of real life experiences as a source of answers for the context-based questions. The last excerpt indicates a perceived mismatch between the content covered in class and the information obtained from out-of-class experiences.

6.2.9 Summary of Section I

This section presented students' perceptions of the use of context-based questions and tests. The perceptions were authenticated by verbatim excerpts of students' statements. Students presented a number of interesting experiences, which made selecting representative excerpts in terms of the views presented, participating classes, as well as the two unit tests, a challenge.

Students presented perceptions relating to positive and negative aspects of context-based tests and questions. According to the students' perceptions, the context-based tests used in this study:

- Comprised questions that were simple and questions that were difficult for students to interpret and answer, depending on how familiar or unfamiliar they were with the contexts.
- Used unfamiliar contexts that may have led to longer processing time for questions and answers, or made it difficult for some students to link their scientific knowledge to the context.
- Used contexts that provided learning experiences for the students.
- Did not always cover the work done in class, making them difficult.
- Allowed students to use knowledge from their experiences or general knowledge, although there may have been difficulties of producing appropriate scientific answers as required by the questions.
- Encouraged students to be observant of occurrences in their environment.
- Required students to think deeper about what they know and understand before answering the questions.
- May or may not have required studying because general knowledge could be used to answer them.
- Allowed students to integrate and transfer knowledge from different sources to answer the questions.
- Provided awareness of malpractices in students' lives or suggested appropriate practices.
- May be simplified or complicated by pictorial representation of contexts.
- Used contexts that sometimes provided clues to answering the questions.
- Used some unrealistic contexts that students had never experienced or observed others experience the situations in those contexts.

The above perceptions influenced students' recommendations for acceptance or non-acceptance of the use of context-based tests in assessing students' learning in Science.

Students also seemed to become emotionally affected by the situations human characters in the contexts were in, so that they either appreciated or became concerned about the behaviour of those characters.

Perceptions that supported the use of context-based assessment models reflect that these models could be used in assessing Science at school level, as well as that they have other beneficial outcomes for the students. Perceptions that were unfavourable for the use of context-based assessment, however, still brought some clarity regarding the assumptions that educators may make, and the precautions that may need to be taken when using such questions for assessment.

Students' perceptions whether they support the use of context-based tests for assessment in Science or not, are important to the classroom teacher. Teachers held perceptions that were similar to those of students, as well as different ones. Teachers' perceptions are presented in Section II below.

SECTION II

6.3 TEACHERS' PERCEPTIONS OF CONTEXT-BASED ASSESSMENT

Teachers' perceptions of context-based tests and questions were obtained through interview data. Their responses raised a number of views about the tests and questions comprising assessment benefits for students, effects on pedagogy, student motivation, challenges, as well as concerns for teachers.

6.3.1 Assessment of different cognitive levels

Context-based tests and questions were perceived to be good for assessment. There were views that the use of these questions in tests enabled the teacher to assess different cognitive levels and skills. The question format was also perceived to be non-threatening to the students. In response to a question on what she thought was good about context-based tests, Inna stated that:

> …different cognitive levels and skills are tested with what is not abstract to them [*students*], but with what is familiar to their everyday life. It's not something that kind of frightens them. … So if this is brought into the test they know that they have seen this in Matsapha, this is happening at home and it sort of ease their tension (Inna Interview).

The excerpt above reflects a perception that context-based questions reduced test induced anxiety. Reduced anxiety was also perceived to be achieved by allowing students to use informal knowledge from everyday life experiences in answering the questions. The use of informal knowledge was perceived to make it easier for the students to remember information from their everyday experiences than from classroom experiences. Lorraine shared the following observation on this matter during her interview.

> I noticed that it did give students a chance to voice out or to provide their own ideas. And being related to everyday life with most students it was easier to remember. Except that in some cases where the question is one that demands analysis sort of that they had a problem.

In supporting Lorraine’s assertion Jabulane noted that the questions allowed students to answer questions in ways that they understand them, particularly if they were linked to context-based teaching.

> Jabulane: [*Reading*] “They answer questions the way they understand the subject content in relation to the environment.” Okay this one was based on the teaching because if you teach the science in the environment when the question comes they are able to relate, look at the environment, then be able to come with the correct answer.
>
> Victoria: So would it also apply to the test?
>
> Jabulane: Ya it could because, let me say you have been teaching and they link that to the environment and then the question comes, it is easy for them to remember, they link that with what they know instead of recalling always. When the question comes they know you are talking about that.

Some context-based questions were perceived to be easy to work on because of their link to students’ everyday life experiences, while others posed challenges for the students. Perceived challenges were that students needed to analyse the contexts and the sub-questions to identify and relate relevant scientific concepts and principles to the situation described in the context. For example, Lorraine noted the following:

> Like the one where some plants grow under a tree and so on. They had to say why the plants under the trees did not look healthy. Some of them could not relate this to photosynthesis being responsible. Most in fact, because it was only a group that was taking Agriculture, most of them who could are the ones who take Agriculture.

Lorraine’s observations indicate that not only did the questions provide opportunities for students to link school science knowledge to relevant contexts; students also had an opportunity to transfer relevant knowledge from Agriculture to answer questions in the science test. This view was also voiced by students who mentioned Agriculture and radios as sources of some of the information they used to answer the questions. The excerpt, in addition, indicates that context-based questions were perceived to extend assessment of students' learning to their ability to connect classroom science to everyday uses of scientific knowledge from other subjects.

The performance of non-Agriculture students did not seem encouraging, as is indicated in the excerpt below.

> Lorraine: The science students gave me *eish*! Some of their answers were just out. And they couldn't even relate that to, at least these even besides relating it to photosynthesis and sunlight not being enough for these plants at least they could tell that the tree was absorbing some nutrients which could have been used by the crops.
> Victoria: Of course the question was on photosynthesis.
> Lorraine: It was …
> Victoria: That being what they had done.
> Lorraine: Yes being what they had done. But here providing them with these, I mean the structuring of the question [*use of picture*]; they couldn't come up with the ideas that they were supposed to relate this to photosynthesis (Lorraine interview).

It seems from the excerpt that even though the question was structured to include a picture that illustrated stunted growth of crops under a tree, some students appeared unable to link the poor growth of plants to the poor lighting conditions and therefore low photosynthesis, or competition for nutrients and water (See Appendix IIIB Question 2). Analysis of the context required students to engage in in-depth thinking processes.

6.3.2 Stimulating thinking in students

Teachers also perceived context-based questions to stimulate students' thinking. Indicators that the teachers perceived the questions to be useful in stimulating students' thinking, are illustrated in the excerpts below.

> The first thing I noticed is that it stimulates the students' thinking because it makes them not just to cram their notes and come and reproduce them. At least it helps them to think which is good for science learning. It also helps them to realise that Science has many applications in everyday life around their homes or different industries and so on. That makes Science a relevant subject (Josephine interview).

> Jabulane: It develops all the skills for the learners i.e. observatory and thinking skills [*written response for interview*].
> Victoria: [*following up*] These are the test questions you are not actually mixing it with the context-based teaching?
> Jabulane: It is. Also in the teaching also in the test where there are diagrams they look at the diagrams and try to link it with the content.

Not only did the questions encourage students to think, students' observation skills were also developed through studying and interpreting diagrams in the tests.

6.3.3 Demand for students' understanding

Context-based questions were also perceived to demand a demonstration of clear understanding of scientific concepts from the students. An example of a statement from the teachers illustrating this perception, is provided below.

> … when the students are answering they need to be clear about their concepts, otherwise, as I've mentioned above, they can be easily diverted. If now they are not clear about what they learned in class, if they did not do it practically or involve themselves with what was happening in class they are likely to get lost. But if they were part of the lesson in class doing what the instructions were telling them it becomes very, very simple. So as far as this is concerned it is easy for pupils to pass (Inna interview)

In addition, the contexts had a potential to divert students' interest away from the requirements of the question. Despite this problem, context-based questions were also perceived to be quite simple for students who pay attention during lessons and who understand scientific concepts. These views concur with those expressed by the students.

6.3.4 Effect on pedagogy

Perceived pedagogical effects were that context-based assessment encouraged the teachers to teach for student understanding of concepts to match the context-based model of assessment. Below is an excerpt from an interview with Josephine that illustrates this point.

> Josephine: It also helps the teacher to stress on understanding scientific facts because as you teach sometimes they have a tendency to just ignore even if they don't really understand you they will just go on because they know ... In this case you stress that they really understand what you are teaching and you also make it a point that they understand you because they need to apply what they are learning through the questions.
> Victoria: So when does this stressing of understanding take place?
> Josephine: When they see the question they see that they just need to know what they have learned to be able to answer the question and as a teacher I think it helps you see that as I teach my students must understand as I'm going to ask questions this way.

The excerpt also illustrates the perception that test questions motivated students to be attentive during lessons, to ensure that they understood the concepts. Josephine also believed that she needed to make sure that her students were properly prepared for the tests. She expressed similar views regarding the use of performance assessment (see Section 4.4.3), noting that she had to make extra effort to ensure that her students developed appropriate practical skills.

6.3.5 Motivation

Increasing interest in Science and related career opportunities, were also perceived good aspects of context-based questions used in the unit tests for this study. Indicators that students enjoyed context-based tests are shown in the excerpt from Josephine's interview presented below.

> Josephine: It makes science learning more interesting to the students especially those who are fast learners. Those who are slow tend to draw back a bit, because it's like, it's a higher level of learning. So for those who are fast it makes the subject more interesting.
>
> Victoria: Ya. I hope you are distinguishing context-based teaching and learning from context-based questions. We are talking about the assessment and you are saying … it makes Science interesting to students who are actually (Josephine: the students who are fast learners). And you see this happening through the use of context-based questions?
>
> Josephine: I think they go together, the learning and the questioning. You find … because if they are learning, applying science then the questioning also it has to do with the application of science. I think it's more or less the same thing. … I tend to think that the context-based learning and questioning have the same kind of impact if I could say in terms of generating interest.
>
> Victoria: Ok. So those were things that you felt were good?
>
> Josephine: Yes. And also it helps the student; stimulate the students into taking science careers, especially those who are fast in learning because they would see that learning Science this way, we would be answering questions this way so it makes it relevant. Somehow it makes them to have interest in science careers.

Josephine perceived the use of context-based teaching and assessment to have positive effects on motivating high performing students and improving their interest in Science. A perceived drawback was that slow learning students found the questions increasingly difficult, resulting in reduced success and interest.

For Jabulane too, the perception of an increased students' interest in Science was based on both context-based teaching and assessment. According to Jabulane, matching the questions to the teaching model was important for student motivation in working on context-based questions.

6.3.6 Perceived problems of context-based questions

While context-based tests and questions were perceived by the teachers to have the advantages enumerated above, there were also perceived problems associated with their use. These were concerned with interference from the contexts; use of informal knowledge in answers; discrimination and motivation of students; challenges for the teachers; as well as context familiarity and fairness of questions.

a) Interference from context

Contexts were perceived to present a challenge for students' focus on the question. Answers to test questions indicated that students diverted their focus from the question to the context or aspects of the contexts. Students were observed to use context information as answers rather than their scientific knowledge, as Inna notes in the interview excerpt below.

> Here learners sometimes fail to respond positively to the science we want, the scientific concepts we want them to give back. They tend to concentrate on the context story, tell you about the story and they just divert. Maybe they are looking at the diagram and tell you something about the diagram, something that we don't want. As a tester you are testing for the concept out of that, but they bring lots of things that are not needed.

Another section of the same interview confirmed Inna's view, thus:

Inna: …	as I have said that student think of getting the answer from the context. For example here the students might include the container as necessary [*for brazier to continue providing warmth*].
Victoria:	Oh looking at the context and looking for the answer there.
Inna:	They say here the tin and maybe the barrier-the house.
Victoria:	Oh. Okay wind barrier. Did you find such responses from the students' scripts or from your other experiences?
Inna:	From my other experiences not necessarily from these ones.

The excerpts above raised other perceived interferences that students may have experienced while working on the test questions. Students' answers reflected that students analysed the contexts to identify possible answers.

In agreement with Inna, although for reasons contrary to Inna's concerns, Jabulane perceived the questions to be useful in ascertaining students' understanding of the contexts used in the tests. He further accepted that contexts were a good source of information for answers, as the following excerpt demonstrates.

> Okay. Let me just take this question [*Question 5 in Air and Life*]. Now when you look at the way the questions are structured there is the diagram, so the students can actually observe the answer instead of recalling so you link that to what is happening and he or she can actually give the answer from what she has observed. And also even if you look at this one the test on electricity if the student was able to identify the battery there, the voltmeter then she's able to or he's able to connect properly and then answer the questions that would follow correctly. Why because there is a diagram instead of just reading like its English comprehension (Jabulane Interview).

Implied in the teachers' excerpts was that, while contexts may have constituted a diversion to answering questions, they could also provide answers. The use of diagrams or pictures in tests was perceived useful for answering questions.

b) Use of informal knowledge in answers

Students' presentation of incomplete answers by using informal knowledge from everyday experiences was perceived to be another drawback of context-based questions. According to the teachers, students experienced difficulty expressing their answers appropriately. This problem was associated with students' limited understanding of the scientific concepts assessed and poor command of the English language used. Josephine presented her observations in her interview thus:

> Josephine: But what came out was that some students cannot express themselves clearly. Like if they are giving an explanation they just can't come out with clear answers. So their expressions were one of the problems.
> Victoria: What could you attribute this difficulty of expressing themselves to?
> Josephine: I think it's a lot to do with their English. Some of them it's their English, some it could be that they didn't really understand so they just can't put the answer clearly.
> Victoria: They didn't understand what, the question?
> Josephine: They didn't understand the facts in class during the lessons. And so when they are trying to put across the answers they just don't express themselves clearly. Some of them were confusing facts like mentioning heat as a cause of sparks. So there is just the confusion of ideas. And I think I've already mentioned that sometimes they misunderstand the question and sometimes they did not elaborate on their answers. … some explanations are not very scientific. They just use common explanations like people that are not science students and this is a very common habit anyway.

Several factors were identified as hampering students' expression of their answers. These were students' level knowledge of the language of instruction, inadequate understanding of concepts, misunderstanding of questions, as well as the use of informal knowledge. In the process students' informal ideas and conceptual problems were also exposed.

Problems of language observed in students' answers were also cited by Lorraine. In response to how she handled feedback to the students after marking the tests, Larraine said that:

> Very few students had scored reasonably, not even high marks. So I tried to get the students who got some correct to try and explain to the others. The problem when they did that although they had the idea at times relating it to the next person is difficult, in English. The problem is the language. Sometimes you find that it doesn't make sense when they explain it to the others. So they still do not get it.

Thus, language was perceived to be responsible for the poor communication of ideas and, therefore, low scores. This was despite teachers' efforts to make some sense of the students' answers. Students' problems of explaining their conceptual understanding adequately in English and reliance on informal knowledge, were major problems perceived to constrain students' answering of the questions.

c) Discrimination between students and their motivation

As mentioned in Section 6.3.5 above, fast learning students were perceived to find it easy to access context-based questions. High-ability students were able to meet the demands of the higher cognitive level of the questions, but not slow learners. As such, context-based tests were perceived to segregate students on merit, as demonstrated in the interview excerpt presented below.

> Josephine: I thought it could discourage some students, especially the slow learners because you find that the fast ones are able to do it very fast, to answer the questions fast, and the slow ones maybe they get blank… They find it hard to operate at that level of answering question.
>
> Victoria: And that level, being?
>
> Josephine: The high level of questioning. …The fact that some students answered the questions without much difficulty. Fast learners tried to answer them at least properly and the slow ones were a bit slower and they couldn't answer them properly. I think it makes slow students to be a bit excluded and somehow the class becomes divided. The division is more than when you have the kind of question where you have recall and other lower levels of questions. Here it is like dividing the class into very fast learners and ones that are slow, these are the good ones and these are the ones that are not so good. And somehow it makes them feel a bit inferior. … the good side of this is that the fast students may have an interest in science careers and then the ones that are slow may feel like they won't make it. (Josephine Interview).

Context-based tests were perceived to be easier for above average students, but difficult for below average students. They were thus, perceived to favour high performing students and to create feelings of failure and inadequacy in low performing students.

d) Challenges for teachers

Demand for broad teacher knowledge to cope with the different experiences from students' answers, and time constraints, were other perceived unfavourable aspects of context-based tests by the teachers. Time constraints were perceived in terms of constructing questions and in marking students' answers that were not quite scientific, yet correct to some degree. Teachers felt that they may need some orientation to effectively use context-based questions in tests, as the following excerpts demonstrate.

> … I think it needs more time for a teacher to set those kinds of questions unless we gets them from exam papers to start with the setting of questions (Josephine Interview).
>
> … Maybe if teachers could be given a chance of a small training how to allocate marks, when do you give all marks, when do you think the student does not deserve all the marks. I had a problem (Lorraine Interview).

The difficulty teachers experienced in deciding on marks to allocate to a students' responses seemed to have been caused by the use of informal knowledge and problems of expression by the students. Teachers were frustrated by incomplete answers that lacked scientific concepts, and the task of deciding how to grade the varied responses presented by the students. Teachers felt that they needed to be meticulous when going through students' answers and in deciding on the acceptability of answers that were not in the marking guide. Here are Lorraine's experiences, as she communicated them during her interview:

> Lorraine: … at times you are bound to think about the answer, of course you are challenged. You might expect one or two possible answers and you find that students being provided the chance of giving their own ideas give so many answers that you just have to find out if it's really out of the question or something could be there. As a result of that the mark allocation becomes a problem again to the teacher, because if it's the students' idea you have your own answers, suppose there are three marks for that possible answer, you find other answers which you have got to find out if they are really wrong. If you think they are possible answers then how much are we going to award the students?
>
> Victoria: So you don't just cross them out and say they are wrong.
>
> Lorraine: No! No! I really don't most of the time. I don't just cross it out as being wrong. At times you do find some ideas when you give yourself some time to check what might be in that student's answer and you do find some portions which have a meaning. The only problem is the awarding

of marks as to how much now should one give them. … You don't simply cross. You have to find out if that idea has anything to do with the question.

The above excerpt reflects an effort by the teacher to mark students' answers fairly. The use of context-based questions seemed to invite context-based answers. Thus, teachers were faced with the task of finding acceptable answers in students' responses. Teachers felt that it would be unprofessional to reject responses as non-answers without first establishing their correctness or lack of it. Inna agreed with Lorraine that she checked if students' answers were indeed not acceptable or if students experienced problems of expression.

Josephine indicated that she dealt with non-scientific answers in the following way:

Victoria: How did you deal with those responses you thought were not quite scientific? You look for the scientific information it is not there, but it's the everyday language that they use.

Josephine: Eish! I don't mark such answers correct. I think they are wrong because I think anyone in the street if you ask someone about the brazier they'll just talk about the smoke, it will affect them, but anyone can say that. I believe that the students have to have a scientific explanation to it. So I was marking them wrong.

Evident in the above excerpt is that the teacher expected answers with a higher degree of scientific accuracy.

Jabulane indicated that he also awarded some credit for answers he considered to be of low scientific level and which made use of vernacular (siSwati) terms. His views are illustrated by the interview excerpt presented below:

> Now you, basically as a teacher you end up giving the mark. Then you correct them along to give them the science terminology. Because if you would actually mark them wrong yet they are right, the only thing is that they don't know it the way you know it. So the giving of the mark is just an encouragement. They will remember next time that this is what we were referring to. So if they again put the same siSwati words then that is when you can mark them wrong, later on.

Jabulane also indicated that he assisted students to deal with language problems, as reflected in the excerpt below:

Jabulane: … So you need to be correcting a number of things, instead of focusing on the science, scientific parts on what you are teaching. So you have to try to help on the development of the language in a number of learners. …

Victoria: So you actually make corrections?

Jabulane: Yes corrections on the scripts, yes even on the sentences they are trying to construct. If it doesn't make sense so you try to make sense in the process.
Victoria: Do you write any corrections? I mean the corrected statements?
Jabulane: Yes. You have to so that when you return the scripts they must see where you marked it wrong.

Writing out detailed feedback and corrections on individual student's scripts brought on time constraints and delays in returning scripts.

The teachers' practices indicate that they used different ways in handling responses from context-based questions, namely: establishing correctness or acceptability of answers that appeared non-scientific; disregard for non-scientific answers; and making detailed corrections on students' scripts. The approach used seemed to be determined by whether teachers wanted to be fair in scoring students' answers, motivate students, or whether they wanted students to demonstrate higher scientific understanding through their answers. It seemed essential for some teachers to be satisfied that they were not discarding responses that students believed were correct answers to the questions.

e) Context familiarity and fairness of questions

The use of context-based questions in tests was perceived to be unfair, at times, for certain students. Such unfairness of the questions was perceived in terms of familiarity with the contexts to the students. Teachers acknowledged that familiarity with contexts differed for different students so that contexts that were not familiar to certain students were perceived to present unfair assessment conditions for those students. In her interview, Lorraine explained her concerns regarding a number of issues pertaining to context familiarity and fairness of questions, as follows:

> …. Then when you assess students, instead of … the practical that we take in class are different from real life activities, because here you think of, take a trolley move it on gravel, pull it on grass and so on. When you give a context-based question like the oxen pulling a sledge and the forces responsible for that students find it hard to relate, yet those who are using the sledges as transport at home it becomes easier for them to relate. So giving topic tests using these types of questions, sometimes you note that it is somehow not fair to some students. But then a full paper where there would be these different topics involved. I think it would balance up somehow because you note that a full paper will have a choice. So you notice that one student who lives in urban areas is able to select those questions and it gives a chance to a student who lives in rural areas too, as well.

Two issues contributing to the perceived unfairness of questions emerge from the above excerpt. One relates to the relationship between teaching activities during lessons and the real life contexts used in the test questions. The other relates to the extent to which context-based questions should be used in tests and examinations.

Regarding lesson activities and how they related to real life contexts used in questions, Lorraine noted that context-based or non-context-based learning activities tended to make use of models (equipment) to demonstrate concepts, while contexts in questions were real world experiences. Students who were familiar with the real world contexts were thought to be more likely to link the context to the models used in class, and therefore the relevant science content. These students were therefore at an advantage of understanding the questions when compared to students who were not familiar with the real life context. Due to the possible variation in links between models used in lessons and contexts in test questions, topic tests that comprised only context-based questions were perceived to be unfair. Teachers proposed the mixing of context-based and non-context-based questions in topic or unit tests, to improve the degree of fairness of the tests for students with different backgrounds.

Regarding examinations, the excerpt reflects a perception that the use of context-based questions is more appropriate, where students have a choice concerning the questions they have to respond to. Another proposal was that rural and urban contextual practices need to be mixed when constructing context-based tests or examinations, in order to improve the fairness of tests for students from both rural and urban backgrounds, as is the case for students in Swaziland.

The need to combine context-based and non-context-based questions in topic or unit tests, was also noted by Josephine in an interview, as follows:

> Most of the students found them challenging … actually they were complaining, some of them, that they found it very difficult, because our students are so used to recall kind of questions. If it is applications, it has to be very few questions, so at least the recalling questions would boost them…
>
> When they meet them [*context-based questions in tests and examinations*] they are not as much as this, what you are doing and the way you are doing them now [*in tests used in this study*]. You find them maybe one or two. You, normally in an exam setting they avoid those questions or if you can mix them, I mean have at least a good number of recall and some of the application questions.

Josephine's views for having only a few context-based questions in a test seemed to be based on her perceived ability of her students to work on questions of high cognitive demand, as the context-based questions were perceived to be.

Inna was also of the view that questions with different formats ought to be combined in tests. Her reasons were influenced by her perception that information from the contexts sometimes interfered with the students' choice of relevant information for answering questions. She presented her views in an interview as follows:

> I think that they should be used, but they must be mixed with the other types of questions, one or two of them not necessarily everything being contextualised. That's what I think. … We have mentioned that the students sometimes concentrate on getting the answer from the context. So just because of that they fail the test not because they do not know, but maybe they think in their mind the answer is here. But if you can bring in some other forms of questions which are not contextualised maybe there is something they can gain out of that. So that is why I think, though we should have contextualised tests, a few of the questions should be of the other form of questions.

According to Inna, having both context-based questions and non-context-based questions in a single test was likely to reduce the loss of marks that may arise as a result of interference of contextual information. That is, students would have better chances of getting a pass grade in those tests.

6.3.7 Future use of context-based questions

All the teachers accepted the use of context-based questions, but suggested some modifications on the structure of the tests. Proportions and diversity of context-based questions in tests were to vary to address issues of fairness due to differentiated experiences with context, discrimination of students based on achievement and the possible interference from contexts in answering questions. These suggestions focused on adhering to the current structure of tests and examinations used in Swaziland, that of having a few context-based questions among non-context-based questions.

The acceptance of the use of context-based questions and suggested modifications seemed to have been also influenced by the teachers' concerns about their efficiency in fulfilling their professional responsibilities. Teachers were asked about their main concerns regarding the use of context-based tests in their science classes. Different views were presented on issues relating to subject content, classroom management, teacher professionalism, as well as student motivation. The excerpts below demonstrate these views.

Lorraine expressed her concerns thus:

> The only problem would be when you have got so many classes to attend to. When you do the marking, as I have said, at times you may think of one or two possible answers you find there are seven of them. That means stopping marking and trying to find out if these other possible answers being given now still relate to the same answers. So if you have several classes to teach with large numbers of students like it is common in our schools, every month you have to give some tests and you find that you move to mid-month still trying to mark. So if you take three Form I classes of about 44 or 47 students per class and you are to mark all these books, remember you still have to teach at the same time. In that case one would find himself or herself just taking the teachers' ideas as being the correct ones and not considering the students'. But if one would have a manageable number of classes it would be possible to apply this all the time.

Lorraine's concern revolved around judging students' answers to context-based questions in a fair way. The effect of large classes and teaching loads on the execution of her duties in marking students' work was another concern. Jabulane's concerns were also about the time constraints imposed on teachers by marking a large number of students' scripts. Especially so, as he spent much time making sense of answers and giving detailed feedback on those answers.

Inna's concerns were content related. She was of the view that the science content taught and assessed would be reduced, if it was not carefully monitored during lessons and in students' answers to the tests. Her comments were as follows:

> Inna: The disappearance maybe of science concepts, science in the contextualised testing. The science should be there.
>
> Victoria: Ok. That's one concern. The science (Inna: is getting lost). Do you see it disappearing?
>
> Inna: It might not disappear provided as teachers we are alert about this. We need to be alert. Otherwise the science might disappear and, if we concentrate mainly, again - I keep going back to the teaching - with the teaching we may find ourselves bringing a lot of things that do not have the science - we talk about this and that and then time may be consumed in class if we keep on talking about this and that. Though I have mentioned that we have realised they can do some of the things at home. We still need to be aware that we should keep time so that we do things thoroughly with understanding, we should not forget that there is the syllabus that we should finish.

Inna's concern was also linked to the amount of teaching time spent discussing a variety of contexts to illustrate the use of scientific knowledge in those contexts and to accommodate students' experiences. She felt that this reduced the amount of time that was available for discussing science concepts during lessons. Another concern was the possible reduced use of science concepts in the answering of questions when informal knowledge is accepted as answers to context-based questions.

Josephine's concerns were motivation related, and were presented briefly as:

> Dividing the class into fast learners and the slow learners and making the slow learners feel inferior to those who are doing better. And what I said that the slow learners may tend to shun careers, science based careers. Those are my concerns.

Josephine's concerns had to do with the motivational effect of segregating good performing students from the not so good performing students, and how to boost the morale of students who experienced difficulties with the questions. She perceived a lack of equal opportunities for all students to perform well in the tests used in this study, because they had very few non-context-based questions. The suggested combine context-based and non-context-based questions noted above, was expected to reduce this division.

6.3.8 Summary of Section II

This section of the chapter described perceptions of teachers regarding the use of context-based tests and questions in assessing Science at secondary school. These perceptions covered positive aspects of context-based tests and questions, some drawbacks, as well as thoughts about possible modifications in their construction.

In summary one can say that teachers perceived the context-based tests and questions to:

- assess different cognitive levels and skills among students;
- allow students to use knowledge from their experiences in a particular context to answer questions with reduced anxiety;
- stimulate students' thinking and understanding;
- provide opportunities for students to transfer relevant knowledge from other subjects to answer the questions;
- encourage teachers to teach for understanding;

- encourage students to be more attentive in class in order to understand concepts;
- demonstrate the relevance of Science through the use of everyday contexts; and
- increase students' interest in Science and related careers.

Perceived drawbacks of using context-based tests and questions were of two kinds, those that affected students and those that affected teachers.

Aspects of context-based tests that affected students were:

- that students tended to interpret the contexts as though they contained answers and therefore used terms from the contexts in constructing answers;
- students' use of informal knowledge to provide answers, resulting in incomplete answers and answers that lacked scientific understanding;
- the difficulty for the students to relate science learned in school to science applied in the contexts, a difficulty that was perceived to be promoted by the use of teaching models that were not matched to the real world contexts used in questions;
- the discrimination between students in the same class leading to increased motivation of high performing students and reduced motivation of low performing students; and
- the unfamiliarity of certain contexts for some students, resulting in unfair testing and assessment processes.

Teachers' concerns regarding the use of context-based tests included:

- Challenges to teachers' knowledge regarding the use of scientific information in real life contexts and the need to do research regarding the varied and "informal" answers often presented by students.
- A need for more time to mark and score the large number of scripts from the large classes, due to reading non-scientific answers and deciding on their acceptability and amount of credit to give, compared to accepted answers in the marking guide.
- A threat to the reduction of science content resulting from accepting informal knowledge in the interest of students' motivation and recognition of students' real life experiences in science tests.

- The perceived increase in motivation of high performing students, lowered motivation of low performing students, and the segregation of students into high performing and low performing students.

In light of these perceived advantages and disadvantages of using context-based tests, teachers also perceived the tests used for this study to be over-contextualised. They felt that the tests contained very few of the standard format questions or recall questions. Therefore, while accepting the continued use of context-based tests, they recommended a practice of combining context-based and non-context-based questions. Another recommendation was combining the use of urban and rural scientific practices in the test contexts. These suggestions were perceived to be a way of:

- aligning the tests to the intellectual needs of low-ability students;
- reducing the magnitude of some of the problems students experienced as a result of their unfamiliarity with the contexts used in test questions; and
- reducing the amount of interference from contexts in the questions with students' selection of appropriate information for answering the questions.

6.4 SUMMARY OF CHAPTER

Chapter Six describes the perceptions of students in Section I, and the perceptions of teachers in Section II. It provides the data and results for answering Research Question Number 2, namely: *How do students and teachers view the use of context-based assessment in assessing learning in Science?* Both students and teachers perceived positive and negative aspects of the context-based assessment model.

Perceptions from students were summarised in Section 6.2.9. They are only briefly mentioned here. These perceptions were largely in relation to the characteristics of the questions such as the:

- level of difficulty of interpreting, understanding and answering of context-based questions due to familiar or unfamiliar contexts;
- requirements and demands of the questions on students' thinking and other abilities;
- extent of the coverage of work from lessons and from out-of-class activities;

- unfamiliar format of questions and uncertainty concerning the format and content of answers;
- learning opportunities provided by unfamiliar contexts as opposed to those consolidated by familiar contexts; and
- authenticity of contexts used in the context-based questions.

Other findings were that contexts may induce some emotional concerns in students regarding the safety or experiences of the human characters used in question contexts.

Teachers' perceptions (also summarised in Section 6.3.8) dealt with pedagogical, conceptual and motivational effects of context-based questions. Teachers perceived context-based tests and questions to:

- encourage teachers, and students, to improve students' understanding of concepts during lessons;
- reduce anxiety of tests by allowing students to use their informal knowledge as part of answers while assessing different cognitive levels;
- be prone to interference of context and informal knowledge in the interpretation and answering of questions, and therefore lower scores;
- discriminate students with regard to their performance in the tests, causing the de-motivation of low performing students while motivating high performing students;
- introduce excessive work for teachers in terms of marking and searching for the acceptability of answers containing informal knowledge; and
- have a potential of reducing scientific knowledge in answers to the tests.

In the next chapter students' and teachers' perceptions are discussed with regard to the sub-categories of task characteristics: task complexity, task importance and value, task requirements; task format, content and presentation; as well students' empathy and motivation. Relevant literature is also used to illustrate the link between the findings of this study to other relevant studies on context-based assessment models and perceptions of students and teachers regarding contextualised assessment formats.

7. CHAPTER 7

DISCUSSION OF RESULTS II PERCEPTIONS OF CONTEXT-BASED ASSESSMENT

7.1 INTRODUCTION

This chapter focuses on an in-depth discussion of the findings presented in Chapter Six concerning the perceptions of secondary school students and teachers regarding the use of context-based tests and questions in assessing learning in Science. The discussion derives support mainly from work on context-based assessment reported by Ahmed and Pollitt (2001, 2000) and Pollitt and Ahmed (2000). Other studies and reports on context-based teaching, learning and assessment are also used.

Students' views on general aspects of the context-based tests and questions are discussed first to provide a general overview of the perceptions. These general aspects are with regard to questions which students "liked" and "disliked", as well as their feelings about the quality of the tests and acceptance or non-acceptance of context-based tests, generally. Reasons for student preferences are discussed below.

Teachers' perceptions that complement those of students, particularly in instances that affected students, are discussed simultaneously with relevant student perceptions. Additional perceptions from teachers that are associated with perceived challenges in marking students' scripts; effects of context-based questions on pedagogy; and discrimination of students on merit, are also discussed and merged with relevant sections of students' perceptions.

7.2 GENERAL VIEWS

The results of the analysis of students' questionnaire responses indicated that different students reacted to different questions differently. Different students liked or disliked different questions, such that each question had students who liked it, as well as students who disliked it. Some students indicated that they liked several questions or even all of the questions, while others disliked several questions or disliked all the questions in the tests. There were questions that were particularly liked by more students than those who disliked them, as well as questions that were disliked by more students than those who like them.

Whilst this study may have benefited from a closer analysis of the relationship between liked or disliked questions and their contexts and content, it was beyond the scope of the study to do so. Nonetheless, when viewed at a glance the data reflected that students seemed to favour questions that involved human characters in situations of perceived danger. In particular, they liked Question 4 from the Electricity test (see Appendix IIIA for the question) and Question 5 from the Air and Living Things test (see Appendix IIIB for the question). Questions 1 and 2 from the Air and Living Things test had elements of some socio-economic benefits and were also generally liked (see Appendix IIIB for the questions).

Students' acceptance or non-acceptance of the tests or questions also varied. There were students who agreed with the suggestion that such tests or questions should be used for students' assessment and those who disagreed.

The variations in the questions that were liked or disliked, or the acceptance or non-acceptance of the tests by students, indicated the differences in individual students' preferences and interests in certain content and contextual issues in the tests. Perceptions from students and teachers also indicated that the students participating in this study had varied backgrounds and experiences of the contextual issues presented in the questions. The tests were perceived to unfairly discriminate against some students in terms of their intellectual abilities, as well as their familiarity with regard to the context.

Students who ventured an opinion about the perceived quality of each test, expressed similar views. Most of the statements that reflected perceptions concerning the quality of each test, were positive. There were some students who perceived the tests to be unfair or who were dissatisfied with them. Some students did not commit themselves to any evaluation of the tests, but indicated some displeasure regarding certain aspects of the tests. Students complained about the extent to which tests adequately covered work from their notes, the unexpectedness of question type or format and that they could not study for this kind of tests. In short, students presented varied perceptions of the quality of the tests.

Reasons students provided for "liked" questions or acceptance of the continued use of context-based tests, are highlighted and discussed below. Also discussed are students' reasons for disliking questions and their concerns about context-based questions and tests. The discussion focuses on the following issues, which unavoidably overlap here and there.

- Task characteristics:
    - task complexity
    - task importance and value
    - task requirements
    - task format and presentation
- Motivation
- Empathy

7.3 TASK CHARACTERISTICS

7.3.1 Task complexity

The perceived level of difficulty varied for different questions and for different students. Students perceived the level of difficulty of questions in terms of how easy or difficult the questions were to interpret and to understand, how easy or difficult it was for the students to formulate answers, as well as how confident students were about the answers they produced. Thus, some students perceived certain questions to be easy while others found the same questions to be difficult. Questions which students perceived to be easy were those that had a direct link to students' real life experiences or classroom activities and notes. Some students felt that it was difficult for them to construct answers to questions, even when they understood the questions. Yet others still, felt uncertain about how to present their thoughts as answers to certain questions, that is, whether to use scientific knowledge or knowledge from their everyday experience.

Students' perceived level of difficulty of the questions seemed to depend on:

- content familiarity (prior learning experiences);
- contexts as source of answers;
- context familiarity;
- students' ability to answer the questions.

a) Content familiarity and prior learning

Students' perceptions of how easy questions were to understand and answer, were related to their prior learning and familiarity with the science content assessed. Prior lessons in Science or other subjects contributed to students' familiarity with the concepts required to answer the questions. Some students were able to identify links between what they had learned and the content of the questions. Those questions were perceived to be easy. Other students were unable to make similar matches of learned and assessed content. These students attributed their perceived difficulty of the questions to the exclusion of work covered in lessons or information from their textbooks.

The perceptions noted above, indicate that students interact with, and respond to, assessment questions and contexts in different, unexpected and individualistic ways, as stated by Boaler (1993). Ahmed and Pollitt (2000) and Linn and Gronlund (2000) assert that every real world context brings out characteristic reactions from individual students, who tend to build their own mental representation of a question. They further note that as students read test questions, different schemata are activated which the students use to construct meaning of the situation presented in the question contexts. Students may construct idiosyncratic models of what the question demands in the same way they construct individual knowledge schemes during the learning process. While reading questions in a science test, students also determine the science content in the questions. The activation of relevant and irrelevant concepts in the students' minds affects their interpretation of the questions. The students then need to select the relevant science concept that matches the answer to the question and then construct answers using words that best represent their understanding of the question. This process demands that the student understands both the context and the science content to be used in it (Ahmed and Pollitt 2001; Pollitt, Marriott and Ahmed. 2000). Thus, the relevant content from prior learning needs to be invoked when students read and analyse questions. If not, they feel that they have not covered the work.

The use of context-based questions, not only tested students' understanding of science concepts, it also tested their ability to interpret contextualised questions and express themselves in writing. The students' competence in the language of instruction was perceived to affect their understanding of the questions, as well as their communication of their answers. Language related problems in teaching, learning and assessment have long been recognised and that they do become more pronounced when contexts are used in questions. Real life contexts may bring in non-scientific meanings to words that may have been intended to have a scientific meaning when used in the questions. If students interpret these words in terms of their everyday understanding of the contexts, they may be encouraged to answer in informal ways that lead to incomplete answers (Pollitt and Ahmed 2000).

Perceived non-alignment of the test questions to class activities or notes seem to have contributed to students' perceived difficulty of tests and questions. From these observations, it is clear that when the students worked on the questions they sought links to class notes, which they had studied. Those who could not find links, experienced difficulties with interpreting and answering the questions. These concerns seem to indicate a need for test designers to ensure that context-based tests are aligned to class activities in obvious ways. The need to align assessment content to curriculum and instruction, was also noted by teachers in a study by Aschbacher (1993).

b) Context as a source of answers

Some questions were perceived to be easy because students used the contexts in the questions as sources of information for their answers. Thus, students indicated that some contexts provided clues for answering questions. Teachers observed students using information and terms from the contexts in their answers. Teachers' views on the use of contexts for answers varied, with some accepting this use while others saw it as misleading to students. Furthermore, generating answers directly from information obtained from the context, may have diverted students from thinking further about possible and appropriate answers to the questions.

Students have been found to use terms from the context as answers, while other students described contexts, including graphics to reduce reading load, as part of their answers (Ahmed and Pollitt 2000). The temptation to use information from the contexts has been explained in terms of the large amounts of irrelevant information contained in descriptions of contexts and in pictures that attract the attention of students (Ahmed and Pollitt 2001).

An alternative, and speculative, explanation for students' use of information and terms directly from contexts, could be the possible transfer of skills of answering interpretive structured questions from standard non-contextualised questions, to answering context-based questions. In standard non-contextualised questions, the introductory statement, or other relevant description in a question, provides information that is used directly to answer the question. Students were very familiar with this format of test questions. Whereas, in context-based questions, the context may or may not provide information for directly answering sub-questions, and the students may not have recognised that.

c) Context familiarity

Students' degree of familiarity with the contexts in the different test questions varied according to their individual experiences. Questions for which students recognised familiar experiences in the contexts were perceived to be easy. Familiar contexts stimulated a sense of reduced anxiety and improved self-efficacy in students. Teachers' perceptions concurred with these views to some extent.

Familiarity with contexts was, however, not a guarantee that students would be able to formulate answers. There were students who recognised familiar contexts, but who indicated having difficulty formulating answers to those questions. Students' difficulty of constructing answers to context-based questions was attributed to unfamiliar question format. In addition, perceived confusion about how to answer context-based questions may be explained in terms of Ahmed and Pollitt's (2000) observation that some contexts may be very familiar to some students, so that it may become difficult for these students to select those aspects of the context that are relevant to answering the questions.

Unfamiliar contexts were perceived to interfere with students' interpretation and understanding of the questions. For some students, questions that used unfamiliar contexts, took longer to interpret and process, than those in which the contexts were familiar. Making sense of the context to identify and select the relevant concepts to formulate answers, delayed their progress on the questions. Students spent some time first de-contextualising the question into a meaningful science question as Ahmed and Pollitt (2001) expressed it. Unfamiliar contexts were also perceived to make it difficult to link classroom science to science in the context.

Teachers' perceptions with regard to context familiarity and its effects on students' understanding and answering of questions indicated that teachers recognised that students had different backgrounds and experiences. The use of contexts in test questions was perceived to be unfair for those students who were not familiar with the contexts. These teachers' concerns concur with Ahmed and Pollitt's (2001) view that, students whose context familiarities vary, end up writing different tests if context-based questions are used.

Problems of unfamiliarity with the contexts were also perceived to be due to the use of different equipment and models for teaching, from the real life contexts used in assessment questions. Science lessons conducted by the teachers, whether context-based or non-context-based, tended to use small scale models or equipment designed for classroom use to demonstrate scientific concepts or principles, or show their use in problem-solving. Students were expected to use scientific knowledge, which may have been developed using the models, to explain or describe aspects of real life contexts or solve real life problems presented in the contexts. Teachers felt that students who were not familiar with a particular context would find it difficult to link the real life context in a question to the models used during lessons, or to adequately understand and answer the questions. Students who were familiar with contexts were perceived to find it easy to recognise the links between the models and real life contexts, and therefore, more likely to understand and answer the questions. The perceived lack of concurrence between teaching materials and the real life contexts in test questions may also explain, albeit only in part, the perceived difficulties experienced by some students in identifying links between the scientific concepts in a the context to science learned in school. Such difficulties could also be observed even in instances where the questions were aligned to work students had done during lessons.

As a result of the perceived effects of unfamiliar contexts on students' understanding of questions, or their construction of answers, teachers suggested strategies to minimise these effects. One way of improving the fairness of the tests, with regard to context familiarity, was the use of both context and non-context-based questions in unit or topic tests. This suggestion was perceived to reduce the number of contexts used, and to cater for students with different abilities. It was seen as a way of giving all students a chance to answer some of the questions. Another perceived way of dealing with the problem of unfairness of context-based questions, was to give students several questions to choose from in a contextualised test or an examination paper. This would also provide opportunities for students to find contexts they may be familiar with. A third suggestion was that of combining contexts from rural and urban situations to cater for students from the different social backgrounds in Swaziland.

Combining contextualised and non-contextualised questions in a single test paper may need to be strategically implemented to comply with the teachers' suggestions, as well as to avoid the need for students to alternate their thinking in and out of the question's context. Ahmed and Pollitt (2000) note that students need to recognise questions that require information from the textbook or class notes, even if they are embedded in a contextualised question. This would require them to think of such questions outside of the context. These authors observed students' failing to answer simple recall context-free questions, which were embedded within a contextualised question, because they attempted to answer the question using the context of the question.

Interference of unfamiliar contexts in students' understanding of questions seemed, in addition, to be associated with students' perceived authenticity of the contexts used in the questions. There were indicators that students sometimes doubted the plausibility of some of the contexts presented in the questions, having not seen or heard about the situations described therein. Some students even felt they had not been exposed adequately to the content that was assessed because they could not identify links to the work they had covered in lessons. These perceptions iterate Ahmed and Pollitt's (2001) observation that student may want to associate the context in a question with their personal experience and, if they have not had that experience, may view it as far fetched and therefore, feel helpless in answering the question. Furthermore, unrealistic or fictional contexts, may lead students to think they cannot answer the questions because they have not encountered information related to the relevant science content in the question (Ahmed and Pollitt 2000). Therefore, contexts used in test questions may need to be closely aligned to the lesson content, be realistic and meaningful for students, if they are to find working on the questions worthwhile (Ahmed and Pollitt 2001; Aschbacher 1993).

Wistedt (1994) notes that students make sense of, and understand questions that are grounded in practical experience, or deal with well-known situations and objects. She also notes, in addition, that context-based tasks invoke everyday images that may help or hinder students in solving a given problem. Realistic and meaningful questions are more likely to stimulate similar schemata from students as intended by the questions, than fictitious contexts. Students were also less likely to spend time extracting the questions on science content from the contexts (de-contextualising the contexts) to make sense and meaning of the question if authentic contexts were used (Ahmed and Pollitt 2001).

d) Quality and acceptability of answers

Students in this study were also more confident about their understanding and answering of the questions that involved familiar contexts. However, for some students such perceived confidence seemed to be based on the use of informal knowledge derived from experiences related to the contexts in the questions. Other students, however, did recognise the importance of scientific knowledge in constructing answers. Thus, from the students' perspectives, there were two possible ways of answering the questions: the scientific way and the everyday commonsense way. The use of a mixture of scientific explanations and informal knowledge in answers was perceived to be acceptable by the students. It seemed easier for the students to use commonsense, non-scientific answers, although some students were concerned about the degree of the correctness of such answers. Students' doubts regarding the use of common sense answers could be an indication that these students recognised the different ways of answering the questions and "debated" in their minds, which way to use. Such "debates" may have also prolonged their processing of the answers. Ahmed and Pollitt (2000) have also made observations concerning students' reactions to questions involving real life contexts. These students were observed to experience uncertainties about the acceptability of their answers and the amount of credit they would get.

Teachers' perceptions were in line with students' perceptions. Teachers perceived context-based questions to allow students to expresmps their answers in ways they understood them in relation to both school and out of school experiences. Teachers noted that students' use of general or informal knowledge to answer the questions sometimes resulted in incomplete statements or answers without proof of the necessary scientific understanding of the question. The limitations of such answers were perceived to be promoted by the students' poor command of English, the language of instruction. This observation further indicates that the perception that questions were easy, was based on answers of varied levels of acceptability.

Teachers also perceived the use of contextualised questions for assessment in Science to invite a plethora of responses that were formulated in informal ways and with varied degrees of acceptability. Some answers were highly descriptive and sometimes poorly constructed. Dealing with such answers was perceived to be a administrative and conceptual challenge by the teachers. Administrative challenges were concerned with the large class sizes prevailing in the schools. Conceptual challenges were related to the marking of students' scripts. Students' informal or common sense knowledge answers were sometimes not matched to the teachers' marking guide. In such instances, teachers felt confronted by having to decide how to allocate marks for non-scientific, but "acceptable" answers. Teachers indicated that, because of such responses, they spent a considerable amount of time establishing the acceptability of the answers and what grade scores to assign. This, however, was time well spent according to Pollitt and Ahmed (2000), as these responses were part of the students' thinking processes that needed to be considered for acceptance. Unfortunately for the teachers, the validating process increased their workload. This was particularly so, for those teachers who had several classes to teach and had to grade their scripts at the same time. Teachers also had certain reservations about the implications of their endorsement of responses on the students' scripts and the reliability of the scores they awarded as a measure of students' knowledge in Science.

Context-based questions also exposed differences between students' knowledge and teachers' knowledge on contextual issues that had implications for the grading of students' work. Teachers felt it professionally unacceptable to reject students' answers without first ascertaining whether they were simple guesswork or authentic responses, which the teachers may not have been familiar with. Teachers also felt that rejecting answers before checking their acceptability would discourage the students, particularly at a time when they appreciated opportunities provided by context-based tests to use their real life experiences in science tests. Such practices by the teachers seem to address Boaler's (1994) caution to assessors that discrediting students for what they may believe to be true, could cause extreme concerns amongst the students, especially so, if it could cause them to fail their subject. The practice may also allay possible doubts of teachers' knowledge and interpretation of natural phenomena that may be developed in students if teachers discredited work in which students believed (Koosimile 2004). Teachers acknowledged learning from the students through the process of validating their answers to the test questions.

Two perceived unfavourable effects of students' use of informal knowledge in answers and their acceptance for award of credit by the teachers could be identified. One effect already noted above, was that informal answers increased the amount of marking for the teachers. This was coupled with the perceived interference with teachers' professional competence due to delayed return of students' scripts and feedback. Further delays were experienced by teachers who had several classes to teach. The other unfavourable effect was the perceived possibility of reduced emphasis of science content in students' answers, particularly when informal knowledge was accepted in answers, although not fully credited. Whilst the use of informal knowledge from everyday experiences may have reduced anxiety and increased students' confidence in these answers, it may have inhibited students investing effort in searching for ways of linking and integrating scientific knowledge from the classroom with relevant knowledge from their experiences or the contexts used in the questions. There were indications that teachers were concerned that the science content specified in the syllabi may be “lost” through the use of contextualised teaching and assessment. While acknowledging the need for recognising the students’ informal knowledge in answers to contextualised questions, the teachers also perceived a need to be cautious not to overlook the importance of emphasising science content. Teachers were convinced that students still needed to demonstrate their scientific knowledge by explaining and answering questions using scientific language and concepts in tests, whether contextualised or not. Similar concerns were also raised by Ahmed and Pollitt (2000) who warned of the possible limitations of teaching and assessing students in contextualised formats. Their concerns were about the non-development of understanding of abstract science concepts by students. According to these authors, such conditions would lead to students’ failing to recognise concepts they understood in one context as applicable in another context, for example a test or examination. They thought students were likely to treat each context they encountered as a new context.

These perceptions present possible challenges for assessment through context-based questions, of ensuring that students’ scientific knowledge and understanding were effectively assessed and encouraged through the context-based assessment model.

7.3.2 Task importance and value

Students perceived several aspects of the context-based questions and tests used in this study as important. The perceptions that the questions:

- provided some learning opportunities for the students;
- encouraged students’ thinking;
- improved their understanding of concepts in the topics;
- encouraged students to be observant in their surroundings;
- tested different abilities;
- made Science relevant to real life; and
- developed students’ awareness of career opportunities;

were presented to varied extents for the two tests. Teachers also recognised the contribution of context-based questions in improving the relevance of scientific knowledge acquired in school to real life activities.

a) Learning from questions

Unfamiliarity with the contexts used in the questions was not always perceived negatively by the students. Even though unfamiliar contexts may have caused students to experience the various problems discussed above, they also perceived context-based questions to provide learning experiences. Learning opportunities were perceived to be offered by both familiar and unfamiliar contexts in questions. Perceptions of learning were concerned with developing awareness of information students were previously not aware of.

Familiar contexts appear to have helped students recognise some of the inappropriate practices they or other people engaged in, such as the use of electrical appliances or improper use of the brazier. This observation suggests that the questions may have stimulated students to reflect on their real life practices to which the contexts in the questions had some similarity. Some students even indicated that they would share their new experiences with other people. These perceptions were not supported by perceptions from the teachers.

b) Increased interest in observations of surroundings

Perceptions of learning from contexts in test questions seemed to be extended to other students' activities. Students perceived the context-based questions and tests to promote their interest in taking closer observation of, and thinking scientifically about, certain occurrences in their surroundings. Some students stated that they would be more observant of their situations or be more on the look out for inappropriate practices in the use of electrical appliances or growing of crops under trees, as alerted to by the questions in the tests.

Students' perceptions of learning from contexts in questions seemed to concur with Malcolm's (2004) assertion that contexts shape learning in unique ways. He suggested that contexts should be selected to ensure that they provide opportunities for deep learning of and about scientific concepts, as well as learning about the context.

Perceived learning from question contexts, which may be familiar, but had never been thought of in scientific ways, or unfamiliar contexts encountered by the students, seemed to support the perceived need to use authentic rather than hypothetical contexts. Students' perceptions of learning from contexts seemed to be influenced by their beliefs in the authenticity of the contexts in the questions. Students perceived the situations in the contexts to represent real life occurrences, as intended. Wistedt (1994) advises that contexts in tasks that are to be used for teaching must be meaningful to the students. This advice seems to have been extended by the students' perceptions of the need for authentic contexts to be used in assessment questions. Authentic contexts seemed to attract more commitment from students who accepted the reality thereof.

The perception of learning through contexts used in questions might contribute to the move towards using assessment for learning (Gipps and Stobart 2003) as advocated by proponents of the shift in assessment focus from a measurement to an assessment model. Whether this learning involved the learning of scientific concepts as expected in the classroom or learning of socially related information, is a matter for another study. Nonetheless, there were indications that both science concepts and socially related information were learned, and that the authenticity of contexts was salient for this to be achieved. Perceptions of task importance and value were also extended to the value of the questions in assessment activities.

Context-based questions were also perceived to have some assessment benefits. Perceived assessment benefits were concerned with assessing students' cognitive competencies such as thinking, understanding of concepts, and use of their knowledge in familiar or unfamiliar situations. These perceptions are discussed below under task requirements as there was some overlap in the case of these two sub-categories.

7.3.3 Task requirements

Teachers and students, generally, perceived the context-based questions used in this study to allow for the assessment of different abilities. Teachers perceived the tests to help in assessing students' skills and abilities at different cognitive levels with less abstract means. These perceptions concur with assertions that putting a question in a contextualised format makes it less abstract and more concrete, and relevant for the students (Ahmed and Pollitt 2000). These authors further note that the questions also allow for the assessment of students' abilities to apply learned material rather than only recall such knowledge from the content in notebooks and textbooks. Nonetheless, some students expressed displeasure at the higher cognitive demands expected of them, as these increased question difficulty. These students expected the tests to focus on what they had studied from their notes.

The context-based tests were also perceived to demand, assess and promote the development of students' thinking abilities and their understanding of the scientific concepts used in the context.

a) Need for thinking and understanding

Students in this study perceived the tests and questions to demand a deep level of thinking about what they know with regard to the scientific concepts embedded in the contexts, as well as the concepts required to answer the questions. The tests were also perceived to assess students deeper understanding of concepts taught during lessons. Students indicated that they needed to understand the concepts at the lesson stage to be able to answer the questions. This was perceived to encourage students to learn for understanding and therefore be more attentive during lessons to benefit from their learning. When presented with contextualised questions, students need to extract the scientific task from the context. Once they have de-contextualised the question they have to identify what they are expected to do in answering the question (Ahmed and Pollitt 2001). This processing of questions necessitates a prior understanding of the relevant concepts and deep thinking about the relationship of the concepts to the context and the questions to be answered (Malcolm 2004). Students who generally perform well in Science were observed by the teachers to perform well in the context-based tests. Teachers attributed this success to these students' understanding of the concepts prior to the task.

The perception that students needed to learn concepts with understanding seemed to be supported by the teachers' perception that context-based questions encouraged them to teach in such a way that students understood scientific concepts, as the tests would require such understanding.

b) Need for revising school work for test

Students' perceptions that questions required their understanding of concepts and deep thinking about how these concepts related to the contexts seemed to be linked to the perceived amount and effort of revising or studying their school work. However, students' perceptions of the amount of revision work required prior to writing the tests varied from a lot of revision of school work to less revision and to no revision. Some students seemed to be of the view that certain questions required more thinking and understanding than revision of work, with some questions requiring no revision at all because they required general knowledge answers. Other students indicated they needed to study a lot for the test and blamed their perceived difficulty of the questions and anticipated failure in the tests on a lack of revision for the tests. There were also those students who felt that they wasted their time on pre-test revision.

Students identified four reasons for wasted effort and time on pre-test revision. One reason was linked to the perceived requirement for students' deeper understanding of concepts in order to explain certain phenomena. Thus, whether students revised their work or not, it seemed to make no difference to their understanding and answering of the questions, if they had no adequate understanding of the concepts required in the questions. The other reason was the perceived low coverage of learned material and information from notebooks in the questions, and therefore a weak link to what they had studied. Some students felt that what they had studied, was not asked in the way they had expected it to be asked in the test.

A third view was that contexts in some of the questions provided clues for answering the questions. Students felt they could obtain answers to some of the questions from the contexts. This perception was also confirmed by the teachers who observed that students tended to look for clues or answers from the question contexts. Some students were observed to use the same terms used in the context to answer the questions.

In addition, students perceived that questions could be answered using general knowledge from their daily life experiences and with less input from school acquired science knowledge. So, studying for the tests was perceived to be of less important in context-based tests.

The perceptions on the importance of understanding concepts prior to writing tests seem to have encouraged some students to work towards such understanding, with the help of teachers. However, the perceptions giving rise to the four reasons justifying the perceived non-requirement of revision of work for the tests, may inhibit some students spending effort and time working on their preparation for the tests. These findings are not very different from observations by other researchers. Researchers such as Maclellan (2001) and Brookhart and Bronowicz (2003) share findings that students' perceptions of assessment tasks; their experiences of assessment tasks; perceived self-efficacy; as well as their perceived importance and value of the assessment tasks, determine the extent to which the students commit themselves to spend effort on assessment tasks or their learning.

c) Being observant of occurrences in the environment

Context-based questions were also perceived to take a dual role concerning students' interest in observing events in their environment. While the questions were perceived to promote learning from surroundings by encouraging students' observations, such observations were also perceived to be beneficial in answering context-based questions. When students read the instructions of a question, relevant schemata in students' frame of reference are stimulated and in the process students recall previous encounters that match elements of the questions (Ahmed and Pollitt 2000). Students in this study perceived it advantageous to be observant of their surroundings as a general practice. Such observations were perceived to increase students' awareness and understanding of different aspects of familiar contexts and, therefore, made it easy to process and answer questions that were linked to those contexts. These two perceptions seemed to suggest a 'pre-test-post-test' complementary function of context-based tests in the development of students' interest in activities occurring in their surroundings.

d) Knowledge transfer

Teachers and students acknowledged the contributions of context-based questions to opportunities for students to transfer relevant knowledge from Agriculture or other sources of information to the science tests. This observation seems to indicate that the test questions enabled teachers to assess students' ability to integrate knowledge from different sources. However, students were also observed to have problems transferring science knowledge acquired in school to answer the contextualised questions. Teachers perceived this problem to be a result of a weak understanding of the concepts, poor understanding of the questions, or students' failure to communicate their ideas in the answers. Thus, while the contextualised questions used in this study may have encouraged the students to transfer knowledge from other sources, such transfer seemed to depend on the topic content and individual student. Barker and Miller (2000) advise that transferring knowledge and skills across topics in one single subject, can be a difficult process for students. Students who are assessed through contextualised questions need to transfer scientific knowledge to real life contexts. In the same exercise they use personal experiences, knowledge and value systems that may add complexity to the knowledge transfer process (Yang 2004).

Thus, as far as students' interpretation and answering of context-based questions, Ahmed and Pollitt (2000) observe that studies on the nature of students' responses to contextualised questions show that students react to such assessment questions in different, individual and unpredictable ways. Boaler (1993) also made similar assertions. Contextualised questions, in addition, invoke everyday images that help or hinder students' attempts at answering the questions (Wistedt 1994). Students in Mathematics were observed by Wistedt *(ibid)* to solve problems either without considering details of the tasks, or by taking into consideration all relevant aspects of the task before formulating their responses. Making decisions on what to consider as relevant, may be a difficult process for some students.

According to Pollitt *et al.* (2000), the art of answering questions is dependent on the process of building mental models of the questions. This process needs to be guided by an appropriate choice of contexts for questions, to improve fairness to students. Cue words in questions provoke schemata that are responsible for the student's plan with regard to an answer. The role of cue words in the final answer needs to be acknowledged as they influence the schemata evoked and mental representations of the demands of the question. The use of familiar contexts seems to support the mental model making process by students, who may use their own understanding of contextual situations in the test questions as a basis for answering the questions. Wistedt (1994) warns that context familiarity may not readily assist students to reach the anticipated responses to a task. Neither does the use of everyday situations in science teaching readily enable students to use school science in dealing with everyday life situations (Campbell *et al*. 2000). The transfer of school knowledge to contexts in tests, or real life situations presented during lessons, is not an easy process for students.

7.3.4 Task format, content and presentation

Context-based questions take the format of structured questions except that the opening statement or introduction to the question describes a context in which scientific concepts are used and which may or may not provide direct answers to the sub-questions. There were indicators that students perceived contextualised questions to be informal and difficult to understand and answer. Students also perceived the unfamiliar format of the questions to affect their mode of responding to the questions. They felt uncertain of how to present their response, even in instances where they understood the questions. Since perceived effects of question format, focus and presentation seemed to have some impact on the interpretation, understanding and answering of the questions, it was discussed under task complexity.

Questions that used pictures or diagrams in their contexts were perceived by some students to be beneficial and useful in interpreting questions. For these students the use of graphical representations in contexts helped them to visualise the contexts or remember the situations described in the contexts. Other students perceived graphical representations of contexts to be confusing. These students needed to develop special skills to interpret such representations to access their messages.

For a third group of students the pictures and diagrams may have interfered with their answering of the questions by tempting these students to search for answers from the pictures or diagrams. Graphical representations of contexts were, therefore, beneficial to those students who could access them, but not to those who had trouble in interpreting and using diagrams to present their answers.

7.4 MOTIVATION

The use of a test that comprised exclusively context-based questions was perceived to affect students' motivation. Teachers observed the tests to separate students into groups of poorly performing students and high performing students. The context-based questions were perceived to be of a higher level for the students who were generally used to answering recall questions with less or no application of scientific concepts. The very low performance of some students in the tests, was perceived to affect the motivation of those students, while the motivation for the higher performing students would be maintained. The observed differentiated performance among high performing students and low performing students, concurs with Ahmed and Pollitt's (2000) observation that context-based questions tend to favour higher performing students. Less able students tend to answer the questions using information from the context because they are unable to select what is relevant information for answers from their evoked schemata. Even though context-based questions may increase the interest and motivation of students, as noted by Boaler (1994), teachers in this study perceived the tests to contribute to a loss of interest by low performing students in Science and careers requiring a science background due to a perceived incapability of learning Science.

7.5 EMPATHY

Whilst this study did not seek to explore the extent and effect of students' emotional involvement in the contexts of the questions, there were indicators that students were sensitive to some of the contexts. Students' emotional involvement seemed to be linked to contexts that concerned human characters used in the context's story. The nature of this involvement included appreciation of the behaviour exhibited by the human characters in the question contexts, fear for the lives of these characters, or being surprised by their behaviour.

Some students appreciated the behaviour of the human characters in the context's story that were perceived to be intelligent and knew what they were doing. Contexts that involved the exposure of people to some form of danger seemed to stimulate students' compassion for those people and some dislike for those questions. That is, questions were disliked because the people in the context were in danger. Disbelief emerged in situations where the characters acted in ways students perceived to lack intelligent thought. These observations indicate three possible ways that may have resulted in students' diversion of their focus from the scientific knowledge in the contexts as a result of affective reactions to the behaviour of the characters in the contexts or the threat to their safety. Students may have concentrated on the human characters in the context story and their problems, a form of behaviour that may interfere with their interpretation and processing of the questions and delay their progress in working on the questions.

Students' reactions seem to confirm their perceived need for authenticity of the contexts used. The reactions may also imply the need for caution when selecting contexts for constructing assessment questions to avoid contexts that may stimulate excess empathy from the students.

Students have been found to take into consideration real world variables of the questions, particularly when the contexts are familiar. In such instances they tend to experience problems extracting relevant issues from the contexts to answer the questions. This was observed by Boaler (1994) in a study involving mathematics students. Ahmed and Pollitt (2000) also acknowledge the emotional involvement of students with subject matter or contexts. They note that the way contexts are phrased, may provoke emotional reactions from the students that may compromise the assessor's control over what is intended to be measured by those questions, as well as to reduce the competencies measured.

7.6 SUMMARY OF CHAPTER

Perceptions of the students and teachers regarding context-based questions and tests indicated positive and negative aspects. Some perceived good aspects also had perceived drawbacks. The discussion of the perceptions is summarised as follows:

- Context-based questions were perceived to be easy to understand and answer, although this seemed to depend on students' familiarity with the concepts dealt with in the questions, as well as familiarity with contexts used.
- Familiar contexts in questions were perceived to reduce students' anxiety as they were able to answer in ways they understood the questions and answers.
- Perceived success and confidence in answering context-based questions was related to the use of informal knowledge from experiences, although there was

a counter perception that the answers tended to be incomplete or superficial, with reduced scientific knowledge. Such answers:

  - failed to demonstrate students' level of understanding of scientific concepts.
  - increased teachers marking load and time and effort spent in validating the answers before accepting or rejecting them.
  - were contributing to a playing down of the importance of maintaining acceptable standards of scientific concepts as required by the syllabi, particularly if they were reinforced by the awarding of credit.

- Context-based questions could contribute to uncertainty of procedures of answering the questions or to a possible failure to communicate answers by students.
- Perceived difficulty of questions was linked to unfamiliar content and contexts in questions, which demanded a higher level of cognitive engagement from students.
- Use of context in questions was perceived to make tests unfair for students who were not familiar with the contexts. Teacher perceived strategies to improve students' motivation and achievement were:
  - combining contextualised and non-contextualised questions in tests;
  - combining questions containing urban and rural contexts; or
  - allowing a wide selection and choice of contextualised questions in a test for students to answer.
- Some questions were perceived to allow the use relevant knowledge acquired through other subjects to answer the questions, but this "transfer" depended on the topic or content and the individual students.
- Some questions were perceived to omit work presented in lessons as students were unable to find links between scientific concepts learned from school and scientific concepts used in the contexts, possibly due to unfamiliar contexts, or as perceived by teachers, a lack of understanding of the relevant content and the question.
- Context-based questions were perceived by students to provide learning opportunities from both familiar and unfamiliar contexts. Students became

aware of information they had not been aware of before. Learning was also perceived to be encouraged by:

    - improved attentiveness of students during lessons; and
    - motivation for being observant of activities in surroundings by students.

- Unfamiliarity with the context was perceived to prolong question processing and de-contextualisation.
- Perceived requirements for studying for context-based tests varied. Some students:
    - found it difficult to study for the tests due to a view that questions demanded an appropriate level of students' understanding and thinking;
    - found it unnecessary to study because commonsense or informal knowledge from daily experiences was considered sufficient to answer questions;
    - used the contexts in the questions as a source for answers, therefore viewing prior revision as a waste of time.
- Context-based tests were also perceived to discriminate between high performing students and low performing students. This discrimination was also perceived to lower the level of motivation and interest of low performing students in Science.
- Perceived authenticity of contexts was important in motivating students to work on the questions; however, sometimes emotional reactions were provoked in students, particularly with regard to contexts involving human characters.

The next chapter presents a concluding discussion of the findings on the perceptions and experiences of students and teachers concerning performance assessment and context-based assessment models. Recommendations and limitations of this study are also noted in Chapter Eight.

8. CHAPTER 8

CONCLUSIONS AND RECOMMENDATIONS

8.1 INTRODUCTION

This final chapter of the thesis begins with a recapitulation of the purpose of the study, the conceptual framework, the rationale, as well as the context of the study. A summary and a brief discussion of the main findings of the study are presented for each research question. The chapter also provides the conclusions drawn from the study and the limitations of the study. Some recommendations for classroom assessment practices and suggestions for further research on the assessment models used in this study are also presented.

This case study set out to explore and describe the experiences and perceptions of students and teachers regarding the use of two alternative assessment models, namely: performance assessment and context-based assessment, for assessing Science at junior secondary school level. The context of the study was Form II and later Form III Science classes where units on Electricity and Air and Living Things were taught following a contextualised approach. Performance assessment tasks (comprising practical hands-on tasks and context-based sub-questions) and context-based unit tests (comprising questions phrased to include real world situations) were used to provide participants the experiences of which perceptions were sought. Swaziland has adopted a contextualised science curriculum where teaching includes the utilisation of real world experiences of students to develop concrete and abstract scientific concepts. There is still a dearth of information about the assessment of students' learning following this teaching approach or how students and teachers perceive assessment through context-based questions.

Teaching, learning and assessment are classroom processes that are linked in complex ways (Gipps 1994). Their link may be viewed as culminating in the concerted effort of developing students' cognitive, affective, physical and social competencies. A re-conceptualisation of the interrelatedness of the three classroom processes has given rise to a new conceptualisation of the assessment process. In the new assessment paradigm advocated by Shepard (2000) and Gipps (1994), assessment is aligned to current cognitive and constructivist theories of learning and a reformed vision of the curriculum. Formative purposes of assessment are encouraged by using assessment activities to promote learning by students. The assessment process emphasises student learning by integrating assessment and instructional activities rather than emphasising the grades awarded to the students. In this study alternative assessment approaches, such as performance assessment and context-based assessment, were viewed as instrumental in achieving the goals of the new view of assessment.

Since the three classroom processes involve individuals who have different backgrounds, interests, motivational levels, expectations and competencies, their reception and perceptions of assessment in the new paradigm were of interest in this study. Aschbacher (1993) and Sharikzadeh (2003) note that the success of implementation of these classroom processes tend to be influenced by the teachers' and students' perceptions of the approach. Thus, the exploration of perceptions was influenced by the premise that perceptions of assessment held by students and teachers tend to drive their reactions, interpretations and efforts in their preparation for and carrying out of assessment activities, as noted by Maclellan (2001) and Brookhart and Bronowicz (2003).

Performance assessment tasks were used in similar ways as learning activities where students worked on the tasks collaboratively. Context-based tests focused on higher order thinking skills and were administered at the end of each unit. Thus, assessment was used in this study to sample and check how much learning had taken place (assessment of learning) in addition to providing opportunities for students to consolidate what they had learned and learning new information (assessment for learning and learning through assessment). Students interacted with the assessment tasks, with each other and with teachers for enhanced intellectual development.

Alternative assessment models such as performance assessment are advocated to assess what students know and can do (Gipps and Stobart 2003). Performance assessment was considered relevant for assessment in Swaziland in that it can be aligned to, and support assessment and learning through the context-based curriculum used in the country. Performance assessment and context-based assessment models have not been commonly used in Swaziland, although the use of context-based assessment is likely to increase, following the introduction of a contextualised curriculum and the assessment trend in the International General Certificate of Secondary Education (IGCSE) used at the senior secondary school level. Students and teachers are the final users of curriculum innovations; their perceptions may have an important role in the implementation, acceptance and success of these innovations in Swaziland.

The summary of the main findings of the study are outlined below together with the research questions.

8.2 SUMMARY OF MAIN FINDINGS

The findings of this case study reflect perceptions of the participants that encompass a variety of aspects that indicate possible strengths and drawbacks in the use of performance and context-based assessment models. These finding also indicated that when the students were assigned the performance assessment tasks or context-based questions and tests, several thought and perceptual processes pertaining to the tasks, questions and contexts were stimulated in the students' schemata. There were indications that the students reflected on the following factors while working on the tasks/questions:

- Their self-efficacy in the tasks/questions. (Can they do the task or answer the questions?; Is the task/question easy or difficult?; What is difficult?)
- The requirements of the tasks/questions. (What do they need to perform the tasks or answer the questions?; Do they need to think and/or understand the concepts?; Did they need to study (revise) their school work for the test or pay attention during lessons?; Were they adequately taught?).
- The importance of the task/question. (How important, and for what, is the task/question?; Is it teaching anything?; are they learning anything?; What do they gain from the task?).

- Their motivation for performing the task or working on the question. (Are they doing the task/test for the grades?; Is it interesting, fun and enjoyable?; Are the assessment conditions fair?)
- The content validity of the task. (Does the task/question cover the material learned?; how much of the work covered in class is in the test?; Is the context used in the task/question authentic or fictitious)?

Students' reflections on group assessment focused on the following:

- Are all students in the group contributing in a balanced way?
- Are there students who are withholding information or depending on others, or hiding behind the enthusiasm of others, or dominating students?
- Why should the marks be the same if contribution levels are different?

The above questions show that the students interacted with the assessment tasks/questions and contexts in the questions in different ways, as observed by Boaler (1993). Students constructed their own interpretations of the assessment tasks/questions, the context of the administration of the assessment models, as well as the meanings and understandings of the different situations presented in the questions, to determine the direction to take in answering the questions. Students also applied their background knowledge in interpreting assessment tasks and constructing meanings of the task and their demands, as noted by Ahmed and Pollitt (2000) and Linn and Gronlund (2000) noted in Sections 2.4.3 and 7.3.1a above.

Responses to combinations of the questions noted above pointed towards the different perceptions of the students and teachers regarding the assessment models. They also resulted in varied levels of motivation of the students. Some students were highly motivated and worked enthusiastically on the assessment tasks/questions. Other students were disgruntled and were reluctant to put a lot of effort into preparing for, and working on the assessment tasks/questions. Yet others were frustrated and in despair as a result of perceived low self-efficacy.

The summary of the main findings of this study are presented according to the sub-categories of perceptions and experiences of students and teachers for the two assessment models. They are also presented according to the research questions. The answers to the two research questions investigated for this study are reported briefly in Sections 8.2.1 and 8.2.2 below. It is worth noting that this study could have benefited from a close analysis of the relationship between the questions liked or disliked by students and their content and context, but this was beyond the scope of the study.

8.2.1 Perceptions of performance assessment model

Research Question 1:

> **What perceptions and experiences do students and teachers have about the use of performance assessment as alternative strategies for determining the attainment of learning outcomes in Science?**

Data for Research Question 1 were obtained through a qualitative questionnaire and group interviews of students and teachers. Data was generated from questionnaire responses using the ATLAS.ti 4.1 (Scientific Software Development, 2004) programme and reading of the transcripts. The perceptions were classified into the following categories and sub-categories, which are further elaborated upon below.

- Task characteristics, which dealt with task complexity, task importance and value, and task requirements
- Metacognition
- Motivation
- Group assessment: peer collaboration and support, and group assessment problems
- Use of multiple tasks.

a) Task complexity

Perceptions relating to perceived level of difficulty of the performance assessment tasks varied for students as a result of different backgrounds and cognitive abilities, as well as for different tasks. Perceived levels of difficulty of tasks depended on the level of difficulty of the topic or content assessed by the task, the nature and format of the tasks, prior learning experiences, expectations of the tasks' demands of the students, as well as the mode of administering the tasks, including their scoring, using rubrics.

Topic and task content: The perceived level of difficulty of performing the tasks was influenced by the level of abstractness of the content assessed through the tasks, the complexity of the arrangement of the equipment and the procedure of testing the material, as well as the manipulation of the experimental data. There were students who perceived the task on testing the electrical conductivity of material easier than that on the electrical resistance of conducting materials. Difficulty was associated with the abstractness of the concept of electrical resistance and the calculation of the resistance of the wires. Sequencing of tasks to begin with simpler tasks and the use of such simple tasks may be useful in building up students' confidence in working with more challenging assessment tasks. In addition, tasks from the Electricity unit were perceived to be more difficult than the Air and Living Things tasks. Many students found the task stage of data collection through testing materials easier than planning the investigation of the task and answering questions embedded in real world contexts. Difficulties of grasping concepts in Electricity were also observed by Chang and Chui (2005) in a study involving students at the same class level, as discussed in Section 5.2.1b above.

Teachers attributed the difficulties students experienced with the assessment tasks to inadequate pre-task preparation, unfamiliarity with the performance assessment model, and low confidence of students in their abilities concerning practical work. These perceptions indicate that it is possible to overcome the perceived difficulties associated with pre-task instruction and familiarity with the assessment format and confidence with continued use and experience in performance assessment tasks. Difficulties associated with content difficulty may be more challenging to address. Careful selection of content when designing tasks for performance assessment may be necessary to reduce cognitive load on students and build up confidence in each topic.

Nature and format of the tasks: The hands-on practical format of the performance assessment tasks was perceived to contribute to making the tasks easier. Students associated the low level of difficulty of the tasks to the use of data they generated through the tasks and data they believed to be relevant to answer the sub-questions in the task.

Prior learning experiences: Tasks that students recognised to be similar to previously encountered in-class or out-of-class experiences, were perceived to be simple and boosted their confidence in working on the tasks. Perceived level of difficulty of the tasks was thus linked to pre-task instruction. This finding does not imply that activities in which students engaged in class were repeated, but rather that such activities could be extended to include other unfamiliar aspects.

*Expectations of the task demands***:** For some students the performance assessment tasks were challenging and confusing. Students' confusion was observed in the difficulties they experienced in using the equipment provided or in selecting materials required for the tasks. Confusion was attributed to unfamiliarity with the task format, as well as students' low confidence in their ability to do the tasks.

Mode of administering the task: The use of group assessment also contributed to making the tasks appear simple to the students. Students shared ideas, learned from each other and collaborated to produce the best outcome concerning a particular task.

Teacher mentorship reduced the difficulty of tasks by redirecting students towards generating meaningful data that led to their success in the task. Some students used this opportunity to the fullest by constantly consulting with the teachers. Teachers perceived this behaviour by students as a means of ensuring that they acquired a good mark by submitting work that had been approved by the teacher. Thus, the support students received from each other and from the teachers while they worked on the performance assessment tasks, influenced the perception that the tasks were easy.

Teachers perceived the use of the rubrics for performance assessment useful in guiding and focusing their observation on specific students' actions and outputs and the scoring of these. However, the amount of detail in the criteria made the rubrics used in this study difficult for the teachers to adhere to. There were many actions to be scored, and students omitted some of them. Difficulties of scoring were enhanced by students' dependence on the teachers for assistance. Further challenges for teachers were their dual and simultaneous role of mentor to the students and judge of their performance on the assessment tasks.

b) Task importance and value

Another motivation for students' efforts in the tasks was their perceived importance, value and future functionality of the tasks. Brookhart and DeVoge (1999) assert that if students identify some value in a task, they are likely to devote more time and effort towards working on the tasks. Students also hold strong views about different formats of assessment, which are important driving forces for the effort students invest in executing assessment tasks (Struyven *et al*. 2003; Maclellan 2001; Brookhart and Bronowicz 2003). Thus, tasks considered important and of value to students stimulate a need to do the task.

Performance assessment tasks were perceived to hold some importance and value for the students and, to some extent, for the teachers. The students and the teachers perceived the performance assessment tasks to provide learning opportunities for the students and assessment opportunities for the teachers.

Learning opportunities were perceived to emanate from the following:

The performance assessment tasks: Students felt that they improved their understanding of the scientific concepts embedded in the tasks through the direct demonstration of scientific content as they carried out the practical procedures of the task. Effective knowledge construction was perceived to occur through the observation of task processes. Students also developed or consolidated their procedural skills at the same time.

Peer tutoring and teacher consultations and mentorship: The use of group assessment and the allowing of students to consult teachers, or teachers to guide students while they performed the assessment tasks, were perceived by both students and teachers to be valuable opportunities for learning through peer tutoring and teacher mentorship. There was, however, a perception that students' consultation with the teachers were, at times excessive and bordering on student dependence on the teachers.

Pre-task lesson activities: Students perceived an increase in their seriousness and attentiveness during lessons, while teachers indicated that they made special effort to help students develop procedural skills.

Post-task observation: Students indicated that the assessment tasks encouraged them to be more observant of occurrences in their surroundings so that their awareness of the relevance and uses of science in their surroundings, increased.

All the above learning opportunities contributed to the perceived improved retention and recall of the knowledge constructed and skills acquired through performing the task.

Perceived improvement of students' cognitive abilities can be expected. Struyven *et al.* (2003) note that, while the students work on the performance assessment tasks they think about the science content in creative ways that enables them to internalize the content for better retention and recall. The practical nature of the performance assessment tasks also contributed to the perceived learning benefit from the tasks. The fact that students were assessed may have induced greater motivation and a good frame of mind for learning.

Assessment opportunities: The performance assessment tasks were also perceived important for assessing skills not normally assessed through non-practical assessment items. This study used hands-on practical activities that contained context-based sub-questions. They were perceived useful in the measurement of students' practical and procedural skills as they allowed students to demonstrate their manipulative and cognitive competencies. They also assessed students' ability to relate science learned in school (experimental data from the tasks) to science used in their environment.

Students' recommendations for the continued use of performance assessment were largely on the grounds of their perceived importance and value to their learning in Science.

c) Task requirements

For students to embark successfully in a performance assessment task, they have to be equipped with the necessary intellectual and physical tools to tackle the task. These tools have to be developed prior to or during the task (Linn and Gronlund 2000). Accordingly, students in this study perceived the tasks to require their prior development of:

- procedural skills and knowledge through exposure to, and experience of practical activities;
- an understanding of science content, which required their attentiveness during lessons; and

- an awareness of the uses of science in the environment in order to link scientific knowledge and the environmental uses of such knowledge.

These were pre-task requirements, by which performance assessment was perceived to induce students' motivation for active participation, attentiveness, seriousness and commitment to lesson activities and the assessment tasks.

Perceived in-task requirements were that students needed to:

- interpret and understand the task to be able to use what they know;
- think deeply about the task before responding; and
- have balanced competencies with regard to theoretical and procedural knowledge.

Students' preparation for the tasks was influenced by their perceived task requirements. Students' perceived their readiness for the tasks in terms of pre-task preparation, level of difficulty of topic content, teaching pace, availability of time to prepare (revise) for the tasks, familiarity with the assessment model and expectations of the assessment task format. In this study students' unfamiliarity with the assessment tasks led to their uncertainties about how to work on the tasks and therefore, a feeling of a lack of readiness.

Task requirements also included the resources for pre-task preparation and in-task performance of the assessment tasks. Time, equipment and personnel were important requirements.

The teachers perceived the impact of time constraints from four perspectives, namely: school administration and organisation, instructional matters, students and task related factors.

School administration and organisation: Time available for teaching science classes in general in the schools was perceived to be insufficient to enable students to complete the given tasks.

Instructional factors: Teachers were uncertain about whether the administration of the tasks took a large amount of time or not, but were convinced that the use of performance assessment tasks increased the amount of teaching time and that it slowed down the teaching pace in the process. Instructional delays arose from pre-task activities as teachers attempted to ensure that students developed the competencies which were likely to be assessed by the performance assessment tasks. The teachers used the tasks more for measurement purposes, that is, they accorded the performance assessment tasks the same test status as they did for traditional tests. Being a new innovation and unfamiliar to the teachers and students, and as an initial experience, it was to be expected that the approach would take more time than would be the case in familiar test models.

Student factors: Student factors that contributed to perceived time constraints in implementing performance assessment tasks, were:

- delays by students to start working on tasks;
- teachers attending to students' problems;
- excess consultation of teachers by the students; and
- unproductive group discussions.

Tasks factors: three factors inherent in the administration of the tasks, through which the tasks served as learning and assessment activities, perceived and observed to take up time, were:

- the approval of students' plans before they proceeded to the testing stage of the tasks. Sometimes students had to wait for some time while the teacher was busy with another group;
- mentoring students who had difficulties understanding the questions or content and procedures, or students who checked the correctness of procedures continuously and slowed down teacher visits to other groups; and
- adherence to the assessment and scoring criteria and insisting on students' performance of required steps before proceeding with further task related steps, some of which slowed down the progress made.

Equipment was perceived to affect pre-task preparation of students and the format of administration of the performance assessment tasks. A lack of equipment resulted in large assessment groups and therefore student participation problems. However, equipment demand was found to be task dependant with some tasks requiring more specialised equipment than others. Tasks that used delicate and expensive equipment were perceived more likely to be affected by equipment shortages than tasks that used basic laboratory equipment that was usually available.

The working condition of the equipment was important in the performance of the tasks, but so was the students' careful handling of the equipment to avoid malfunctioning or breakages. As such, malfunctioning or broken equipment required the services of additional personnel, who assumed the role of a technician. It would be very difficult for a single teacher to effectively administer performance assessment tasks without additional help, until students were more confident in working on the tasks and needed to consult the teacher less.

d) Motivation

Performance assessment tasks were perceived to provide a rationale for practical activities students carried out during lessons. They extended the use of practical activities as learning activities to assessment activities for a combination of abilities and skills. They were also perceived to have a strong motivating effect on students for learning. Such motivation manifested itself in students' increased level and quality of participation, attentiveness, seriousness and commitment during the performance of the tasks and during lesson activities.

Performance assessment tasks were also perceived to improve the assessment scores of students. Students and teachers attributed the improvement of scores to the following:

- Allocating marks for each of the different stages of the tasks, that is, planning, testing/investigating, analysis and interpretation of results and answering of questions. The marks were perceived to show a recognition of students' efforts in the different parts of a task, which was appreciated by the students.
- Group collaborative work on the tasks that produced quality answers.
- Teacher mentorship that led to acceptable and useable data.
- Reduced students' anxiety as a result of group assessment and teacher guidance, as well as the use of familiar tasks that had some similarities with class activities.

The use of group assessment to administer and score performance assessment tasks was perceived to be beneficial for improved scores for some students and problems of unfairness for others. Teachers and students perceived the marks awarded for the assessment tasks to be an important factor in motivating students' serious and committed engagement with tasks, as well as in pre-task learning. The importance students attached to the scores and grades emerged clearly in their complaints about marks being given to students who did not participate fully during the performance of the tasks, however, any consideration of non-scoring of students' attempts at the tasks was perceived by the teachers to be likely to reduce students' seriousness with regard to the tasks. Teachers' expectations concurred with Boud *et al*.'s (1999) acknowledgement that the non-assignment of scores to assessment tasks leads both students and teachers to see those tasks as less important and to divert their effort to other learning goals.

The findings on the motivational effects of the tasks further show that the teachers perceived the performance assessment model to influence their teaching approach towards emphasising the development of procedural skills to equip students for acquiring good scores. The development of students' procedural skills was also promoted by the assessment tasks. The influence of other external forces, such as participation in this study, on teachers' motivation concerning working towards ensuring students' success in the performance assessment tasks, cannot be ruled out.

e) Metacognition

Another finding in favour of the use of performance assessment tasks was the students' perception that the assessment tasks used in this study promoted elements of metacognitive learning. Students felt that the tasks were useful in their self-assessment and monitoring of what they knew. Students also noted that, during the tasks, they questioned what they knew, or had read, or had been told by their teachers, and they sought corroborative evidence from the results of the tasks and their peers in the groups.

f) Group assessment

Performance assessment tasks were administered through a group assessment approach because of an inadequacy of some of the equipment, such a voltmeters and ammeters.

The use of group assessment was perceived to have certain advantages and disadvantages, as alluded to in Sections 8.2.1a and 8.2.1d above.

Advantages were associated with peer collaboration and support, which was perceived to lead to improved understanding of the tasks and subject content. Students interpreted the tasks collaboratively, which helped to make the tasks and its requirements easier to understand. Students felt that through peer collaboration, they were able to construct quality answers and obtain a combined higher performance output. In the groups students tutored each other to acquire new skills, ideas and knowledge to improve their understanding of scientific concepts. Such benefits from group assessment have also been observed by Webb *et al.* (1997) and Fawcett and Garton (2005).

Students and teachers also perceived group assessment as promoting the development of important life-long social skills. These skills were associated with collaboration, teamwork, communication and learning to learn from listening and watching others during the performance of the tasks.

While group assessment was perceived beneficial, as noted above, students and teachers also perceived certain problems associated with group assessment. These problems were:

- differentiated participation such as dominance (due to students' confidence and enthusiasm); dependence (due to students' unwillingness to participate and a lack of knowledge); and
- a lack of opportunity to participate (due to large group size and a lack of equipment).

Other group assessment problems were the perceived unfairness of allocating marks to group members irrespective of the weight of their contribution during group work. Another problem was the differential acceptance of ideas contributed during discussions. There were perceptions among students that some ideas were accepted, because they were considered superior than ideas from other students, or that the acceptance of ideas depended on who advanced the idea.

Students and teachers also perceived discussions in groups that reached no consensus to be non-beneficial in terms of the concepts assessed or learned through the task. Inconclusive discussions produced no answer or crystallisation of concepts for understanding.

The perceived group assessment benefits and problems are inherent in group assessment. This thesis views this apparent conflict as an indication of the challenges the teachers and students in this study experienced in using group performance assessment tasks for assessment and learning. The interest in marks hindered any perception of the benefits of debating different views in developing personal and group ideas, due to the fact that such debates were viewed to be hampering progress made on the assessment task.

g) Multiple tasks

The use of multiple assessment tasks was with regard to the use of more than one assessment model or activity in any one unit. Students and teachers held similar perceptions regarding the use of multiple assessment tasks. This study used two performance assessment tasks for Electricity and one context-based unit test, while one performance assessment task and one unit test were used for the Air and Living Things unit. The findings indicated perceptions that the use of multiple assessment tasks in a single unit:

- provided opportunities for a balanced assessment of students' abilities, while accommodating different assessment preferences among the students;
- could lead to an over-assessment of certain competencies in the different assessment items; and
- would need more teaching time to allow teachers time to ensure students' adequate development of the necessary skills and for revising work covered in lessons in preparation for students' writing of the unit tests.

Pre-task time constraints may, however, be reduced by viewing performance assessment tasks more as an extension of learning activities and less as tasks for generating grades.

The use of performance assessment tasks raised perceptions on exciting aspects of the assessment model among students and teachers, as well as perceptions on challenging aspects. These perceptions need further interrogation to establish their impact on the use of assessment for learning in Science in Swaziland. They are also important because of the effect they may have on students' motivation to invest effort in working on group performance tasks.

8.2.2 Perceptions of context-based assessment model

Research Question 2 focused on students' and teachers' perceptions regarding the use of context-based assessment and was phrased thus:

> **How do students and teachers view the use of context-based assessment in assessing learning in Science?**

Students' and teachers' views regarding the use of the context-based assessment model were obtained from data extracted from responses to a questionnaire that was administered to students soon after writing each of the two context-based unit tests and interview transcripts. The same procedures used for generating data for a performance assessment model were used. Perceptions were predominantly in the categories of task characteristics and affective disposition. Thus perceptions relating to task complexity, task importance and value, task format, content and presentation and task requirements, as well as motivation, constituted the main findings of the study. Emotional involvement was also an important finding with implications for the selection of contexts.

a) Task complexity

Students perceived questions to be easy or difficult depending on their ability to recognise the link of content focus in the question to content done during lessons in Science or encountered in other subjects. Questions for which no links could be identified were perceived to be difficult. Difficulty of questions was also associated with students' perceived ability to construct answers to the questions. Thus, perceived level of difficulty of the context-based questions was dependent on students' understanding of relevant content dealt with during lessons, which enabled them to interpret and answer the questions. These were perceived to be dependant on prior learning and extent of familiarity with the contexts used in the questions.

Content familiarity: Students preferred questions which they could easily relate to work they had covered during lessons, or to their real life experiences. Familiarity with the content and context was an important factor to perceived low difficulty levels of questions. Questions students perceived to be unrelated to work presented in class were disliked and sometimes considered difficult. Work done implied a clear inclusion of content information from notes or text books in the questions. The design of the tests was such that they covered work expected to have been taught, so that the perceived lack of coverage of work done in the questions, may imply that there were students who:

- had not really dealt with the content and skills assessed by the questions;
- did not use their workbooks where information from practical activities was recorded, but only relied on the notes provided by their teacher; and/or
- had difficulty associating the science content in the question context with classroom science content.

The different views regarding the perceived level of difficulty of the questions among the students, indicate that the students interacted with, and reacted to, the questions and context in different and in individual ways, as noted by Boaler (1993).

Variation in students' perceptions regarding the level of difficulty of the tasks concurred with the teachers' perception that the tasks allow the assessment of students' skills and abilities at different cognitive levels in less abstract ways. The use of contexts may have made the questions less abstract and more relevant, and in line with familiar experiences to the students. However, students and teachers perceived some questions to make higher cognitive demands on the students, such as analysis, thinking and demonstrating understanding, by showing how scientific concepts work in the contexts presented. These findings concur with views expressed by Pollitt and Ahmed (2000) and Boaler (1994).

Context familiarity: The degree of familiarity with the contexts used in questions affected students' perceived level of difficulty of questions. When the context was familiar, students could relate the context to their own experiences and therefore perceived the questions to be "easy".

Differing perceptions of the effects of context familiarity were presented by the teachers. Whilst there were perceptions that the students should have been familiar with the contexts used in the test questions, there were also perceptions that it was not fair for students to be asked questions that contain unfamiliar real world contexts. The teachers' concerns were that unfamiliar contexts made it difficult for students to recognise links between the teaching models used during lessons and the real life context presented in the questions. For students who were familiar with the context, it was perceived to be easy for them to recognise such links. This finding also provides a reason for students' observed inability to see links between science activities, which they participated in, in the classroom, and scientific concepts encountered in the real life contexts. Other explanations were the lack of understanding of concepts and interference from the context, causing diversion.

Clues from contexts: Questions were perceived to be easy because the students, and to some extent the teachers, believed that contexts in questions provided clues for answers. The perception that contexts provide answers to the questions corroborates observations made by the teachers that students selected information from a context and presented it as answers to questions. Whilst searching for clues from the experimental data worked well in the case of performance assessment tasks, or conventional tests, it may not work as easily for context-based questions. In the performance assessment tasks, the testing stage generated data that were interpreted and used to answer context-based sub-questions embedded in the task. According to Pollitt *et al.* (2000), the diversion of students' focus to the context as a source of answers, is driven by cue words in the contexts, which provoke schemata that are responsible for students' plan and choice of words for the answer.

This finding also indicate a possible transfer of the experience of answering context-based questions embedded in performance assessment tasks by using experimental results, to the answering of context-based questions in unit tests by using information from the contexts. Another way of explaining the search for answers from contexts, could be that students may be transferring, to context-based questions, their experience and practice of answering conventional interpretive structured questions. Structured questions provide information in the form of data tables or diagrams for use in answering those questions. Context-based questions require a different method of interpretation from the conventional interpretive type questions with which the students were familiar.

Quality of answers: Context-based questions also invited context-based answers. Students presented perceptions on how they could frame their answers, as follows:

- Informal knowledge or commonsense answers.
- Both informal knowledge and scientific knowledge could be used in answers.
- Scientific knowledge was required.

These views also influenced students' perceived level of difficulty of test questions, where the use of informal knowledge was associated with a perceived low difficulty level of the questions. The need for scientific knowledge was perceived to make the answering of the questions difficult. Students' familiarity with the contexts encouraged a perceived adequacy of context-based answers rather than the investing of more energy in their search for scientific answers.

The perception from students that informal knowledge was adequate in answering questions, presented a challenge for teachers in making decisions concerning the acceptability of answers framed in informal terms. Teachers presented different perceptions about the handling and/or grading of students' answers that were framed in everyday language. Teachers felt that:

- they could not disregard the scoring of answers that were undeniably true in respect of experiences of both students and teachers or in respect of teacher unfamiliarity with students' answers, despite the exclusion of such answers from the marking guide;
- they needed to spend time ascertaining the acceptability of the answers and deciding on the scores to assign, because they were uncomfortable about awarding marks if they were not certain of the correctness of the answers; and
- there was a possibility that scientific knowledge could be lost due to their "loss" of control concerning the direction of students' thinking processes in answering context-based questions, as noted by (Pollitt and Ahmed 2000).

Time invested by teachers ascertaining the acceptability of students' responses to context-based questions, was perceived to delay the scoring process and the returning of students' scripts, particularly where one teacher taught several classes.

Authenticity of contexts: Unfamiliar contexts induced doubts in students about the authenticity of the context and concerns about being asked questions on things that did not happen in the contexts from which they came. Wistedt (1994) notes that practical experience is an important element in students' making sense of, and understanding, a context-based task, as many images are invoked that may help or hinder a student in the processing of the task. Thus, according to Ahmed and Pollitt (2001; 2000) it is important to align contexts as closely to the content as possible to make them realistic and probable. Unrealistic contexts make students think they have not come across the work in the classroom situation.

b) Task importance and value

Context-based questions were also perceived to possess the following useful aspects to the students:

- Meaningful learning opportunities.
- Encouragement of students to be observant in their surroundings.
- Encouragement of students' to think beyond the classroom.
- Improved understanding of concepts.
- Assessment of different levels of cognitive competences.

Learning opportunities: Contexts were also perceived by students to provide opportunities for meaningful learning. Familiar contexts were perceived to help students become aware of the scientific implications of certain practices in those contexts, while unfamiliar contexts made students aware of events and practices they had not encountered previously.

Observant in surroundings: As was reported for learning through performance assessment tasks, context-based questions were also perceived by students to extend learning opportunities beyond the test, and to become more observant of the surroundings.

The perceptions that students learn from contexts corroborate Malcolm's (2004) assertion that contexts shape learning. He thus advises that contexts used to promote learning, should be selected carefully to ensure that students engaged in deep learning of and about science and the context.

Promoting students' thinking: Students and teachers perceived the context-based questions to stimulate students to think deeply about the science concepts as they attempt to relate the questions to the contexts

Improving students' understanding: The perception that the questions improved students' understanding of the concepts assessed was associated with the thinking demands questions made on students. As students thought about the content and contexts, they seemed to feel that they improved their understanding.

Assessment of different abilities: Teachers perceived the questions to allow them to assess students' skills and abilities at different cognitive levels in less abstract ways and in a non-threatening way. Students had to recall, show understanding, analyse contexts and use scientific knowledge learned in class to explain contextual occurrences. That is, they had to link school science to science in the contexts.

c) Task requirements

Students perceived the requirements for working successfully on the questions to comprise pre-task preparation and place in-task demands on them. Perceived requirements were related to:

Attentiveness to activities in lessons and surroundings: Students generally felt that they needed to be more attentive during lessons and to events of scientific relevance in their surroundings and also study for the tests. Some students felt they only needed to be attentive in lessons to understand the subject matter and be aware of what goes on in their surroundings because the questions needed them to think and understand what to do.

Degree of studying: The perceived amount of studying required for the tests, varied. Some students had a perception that they needed to study or revise their work, while others felt they did not need to study. Others still felt that they found it difficult to study for such questions. The perception that they do not need to study for the questions seems to support the perception that the questions:

- could be adequately answered from students' knowledge of their everyday experiences or from clues provided in the context; and
- did not cover work they had engaged in during lessons.

Thinking and understanding: Students and teachers perceived the questions to demand that students think deeply about and understand the assessed content and procedures in order to answer the questions meaningfully. Pollitt and Ahmed (2000) note that the process of answering context-based questions involves the de-contextualisation of the question by students as they extract the scientific task from the question. In order to do that they need prior understanding of the relevant concepts and deep thinking about the relationship between the concepts, context and questions to be answered (Malcolm 2004).

Some students appreciated the perceived demand for thinking deeper about the questions and about the context. Other students did not appreciate the demand for thinking at a deeper level, feeling that it contributed to the difficulty of the questions.

d) Task format, content and presentation

Perceived task difficulty was also associated with the unfamiliar presentation and format of the questions. For some students the questions were perceived not to be formal questions and difficult to answer, because they were uncertain of how to respond to the questions. Thus, some students presented their answers using the familiar everyday knowledge and language, obtaining answers from the contexts, while other students used scientific knowledge as required by the questions. This finding and other reports from the literature indicate the challenge the students faced when working with context-based questions, being tempted by cue words in questions and the familiarity with the contexts. Reports from the literature also indicate that contexts in questions invoke everyday knowledge schemata that students find appealing as answers to the questions (Ahmed and Pollitt 2001, and Pollitt *et al.* 2000).

e) Motivation

The extent to which different students perceived questions to be easy or difficult, was observed by teachers in the extreme discrimination of students on merit. Students who performed well in the tests were on the upper level of the score while those who were at the lower end of the scores really performed poorly. These observations agree with Ahmed and Pollitt's (2000) observation that context-based questions favour high performing students. The findings of this study indicated that the motivational level of students who performed poorly declined in relation to their interest in Science and careers requiring a scientific background

f) Empathy

Students stated reasons for liking or disliking questions that indicated that they became emotionally involved with the context and in particular empathised with human characters perceived to be in some danger. Students in this study who indicated possible emotional attachments to a particular context were concerned about the effects of the behaviour of the human characters described in the contexts on their lives. Contexts have been observed to invoke certain emotions in students, particularly familiar contexts. These emotions may divert students' focus from the scientifically relevant aspects of the context to emotionally touching aspects and may, therefore, result in students missing the expected scientific interpretation.

8.3 CONCLUSIONS

Perceptions students and teachers hold about assessment affect the motivational levels they may have, as well as the effort they may invest in preparing for and working on assessment tasks/questions. The awareness of these perceptions by teachers may be important in the construction of assessment tasks and their use in classrooms. While the study was a case study and its findings were not intended for generalisation beyond the participants, it did provide some useful indicators about the nature and types of perceptions that students and teachers may develop regarding the use of performance and context-based assessment models. Applicability of the findings to situations beyond the parameters of the case studied is not ruled out.

The findings show that participants in this study perceived positive, as well as negative aspects of the two alternative assessment models used. The positive perceptions concerning certain aspects highlight the promise for the meaningful use of the assessment models in science classrooms. Perceptions on the drawbacks imply that caution may need to be employed if, and when, these assessment models are used. Nonetheless, further research is required to establish the extent of the perceptions exposed by this initial and exploratory study.

The main findings of this study provide information on how teachers and students perceived performance assessment and context-based assessment models from their experience in two units of the secondary school science curriculum in Swaziland. The findings on the perceptions about the various aspects of the assessment models show possible important roles these models could play in promoting students' learning and the assessment of their learning in Science. The assessment models as used in this study, were perceived to have complementary properties. For example, real life contexts were used in phrasing the sub-questions of the performance assessment tasks; the two models had similar effects (pre-task, in-task and post-task) on students' learning, teacher pedagogy, motivation (pre-task, in-task and post-task) and requirements for answering the questions. The importance of performance assessment tasks in formative assessment is evident in these perceptions.

The findings also indicate that the use of the two assessment models can support students' learning in various ways. Assessment was perceived to support learning:

- directly, through student engagement with the content and procedures of the assessment tasks/questions,
- indirectly, through its influence on
  - teacher pedagogy, by encouraging teachers to assist students to learn with more understanding;
  - its mode of administration (group assessment/teacher mentorship);
  - students' level and quality of participation, seriousness and commitment during pre-task lessons and in carrying out the assessment tasks and in answering questions; and

- students’ increased post-task observations of occurrences in their surroundings.

The practical nature of the performance assessment tasks, the use of groups and teacher mentorship, the use of familiar contexts and informal knowledge answers, were also perceived to make the tasks/questions easy and to reduce students’ anxiety about the tasks. There were, as indicated above, also perceived problems regarding these aspects.

Teachers were challenged by the scoring of informal knowledge answers that were at variance with the marking guide, and varied levels of acceptability. Ascertaining their acceptability in order not to reduce students’ motivation, or accounting for students’ learning, enabled the teachers to update their contextual science knowledge. It also made teachers aware of the possibility of ideas among the students that may need to be dealt with when using context-based teaching and assessment approaches.

Unit or topic tests which comprised context-based questions at the exclusion of other familiar non-contextualised and simpler questions, were perceived to increase difficulty. Unfamiliarity with the question format and contexts were also associated with perceived difficulty of the questions. These factors were perceived to separate students into two extreme groups of high performing and low performing students.

The perceptions and experiences of the students and the teachers regarding the use of the context-based assessment model present indicators for possible implications for their implementation in science classrooms in Swaziland. Test development may benefit from taking cognisance of these perceptions. Context familiarity plays an important role in students’ confidence and their perceived level of difficulty of questions. Students have varied contextual experiences, preferences and interests in content and contextual issues, which mean that familiarity with the contexts used in a test may differ from one student to another. Differences between teaching models and the real life situations in the contexts used in questions did not simplify problems related to unfamiliar contexts. These findings imply that test designers need to make an effort to provide variation in the contexts used in test questions to increase chances of context familiarity for students. Teachers proposed that the variation in the contexts could be achieved by constructing tests that combine contexts from different societies (urban and rural), as well as combining contextualised and non-contextualised questions in a test, and by allowing students to choose questions in such tests.

The use of performance assessment tasks, together with unit or topic tests was found to provide a balance in the assessment of different competencies in a complementary way. Strategic emphasis of certain areas of content in the different assessment models can be used to assist students to reach a better understanding of those concepts.

Students became emotionally involved with some contexts indicating that contexts may interfere with students' interpretation and understanding of the relevant aspects of the questions. Users of context-based questions may need to be cautious in their selection of contexts for questions to reduce their negative effects on students. Alternatively, students may need assistance in the development of strategies they use in the answering of context-based questions.

8.4 LIMITATIONS OF STUDY

This study was an exploratory case study, which, although confined to the conditions and settings of the participating schools, students and teachers in Swaziland, shares certain findings with other studies reported in the literature.

A few variables which could not be controlled within the limits of this study, may have influenced the outcomes of the investigation. These limitations are outlined below.

1. The study set out to explore perceptions about performance assessment among students and teachers. School conditions regarding time as a resource, equipment and personnel defined the group mode used to administer the performance assessment tasks. The use of group assessment introduced variables that students and teachers perceived to be characteristics of the performance assessment model. Nonetheless, group assessment had several positive effects for the implementation of the performance assessment model.

2. The perceptions of students and teachers reported in this study were based on experiences from two science content units. Some of the perceptions reflected problems that were associated with limited experience of the participants in the use

of the performance assessment and context-based assessment models. The experiences with these assessment approaches were possibly isolated to the two units used in the study in a large pool of conventional assessment experiences and practices. This study had hoped to provide the participants with experience from four curriculum units in an attempt to prolong their experiences in the use of the alternative assessment models, but was constrained by school variables such as time management and other school programmes.

3. As alluded to in the second limitation above, this study was a short term study involving practices that have not been used by the participants. Thus change regarding the use of performance and context-based assessment may not have had sufficient time to take effect. Hord (1987) advises that change is a process that needs time to bring about any effect. On the same note, students and teachers may have needed time to develop familiarity with the assessment models.

4. To explore students' views and experiences of context-based assessment assumptions were made that students would be able to interpret the questions considering that they had been exposed to the scientific concepts embedded in the contexts during lessons. However, students' background in electricity was not established sufficiently to ascertain whether their difficulties were related to context familiarity (as there are still un-electrified homes) or content difficulty.

8.5 RECOMMENDATIONS

This was a case study conducted in Swaziland and its findings cannot be generalised as students' and teachers' perceptions of performance assessment and context-based assessment models, even though some may be similar to findings from other studies. So, the following recommendations are made with a view of further broad scale and/or in-depth exploration to ascertain their possible extension to general practice.

From the findings the following recommendations could be made:

1. Some of the problems experienced in administering the performance assessment were induced by the design of the tasks. Some tasks required more assistance to be provided to the students. It may be necessary to design tasks

which do not need a large amount of additional assistance to students as they perform the task. Alternatively, team administration of tasks could be used where the one teacher assumes the role of mentor and scorer, and the other teacher takes the role of technician. Care needs to be exercised to ensure that students do not proceed to gather data in the absence of a mentor or scorer, as this may result in unfair grading.

2. The scoring of students' performance of assessment tasks was perceived to be problematic, partly because of the number of criteria to be followed in the rubrics, as well as the use of group scores for individual student achievement. It is recommended from this study that the use of performance assessment should emphasise learning through the tasks and put less emphasis on the use of scores for achievement of individual students, if group performance assessment is to be used. Criteria in rubrics could therefore serve as a guide with regard to the important competencies that students are to develop or consolidate. The use of group performance scores could, in addition, be limited to group scores rather than being translated into individual scores. Group assessment would thus be used mainly for its learning benefits to individual students in the hope that such benefits would be transferred to other assessment tasks to improve individual student achievement.

3. There were indications that some students became emotionally touched by contexts in which the characters were perceived to be in danger, or if the context related to a dangerous experience the student might have experienced or heard of. Caution needs to be exercised when selecting or constructing contexts, so that, while the contexts maintain their authenticity, they avoid life threatening situations. Or, at least until such time that students have had sufficient experience of handling context-based questions.

4. Context-based questions invited a range of answers of varied levels of acceptability. Teachers were challenged by the scoring of answers that were focused more on informal knowledge and less on scientific knowledge. Furthermore, teachers' research on the acceptability of the answers delayed the marking process and the return of students' scripts. These challenges need to be

explored and attended to in Swaziland, particularly because of the adoption of a contextualised science curriculum.

5. Students and teachers were not familiar with the use of performance assessment tasks and only familiar to small extent with context-based questions. They needed training and time to get used to the intricacies of working with these assessment models. Interpreting questions and selecting appropriate scientific facts and arguments learned during lessons, responding to context-based questions, as well as the setting and scoring of assessment tasks and questions by teachers, were the main aspects for which training would be beneficial.

6. Context-based tests were perceived to discriminate between high performing students and low performing students to such an extent that low performing students might have felt discouraged by their poor performance. Students' motivational level may be increased if unit tests are structured to include both context-based and non-context-based questions. However, these two types of questions would need to be placed in different sections of the test, as mixing them has been observed to have problems of confusing students, by having to think in-and-out context, as noted in Section 7.3.1 c) above.

8.6 SUGGESTIONS FOR FURTHER RESEARCH

This case study was intended as an exploratory study that could serve as a precursor for enquiry into the use of performance assessment and context-based assessment models for learning in Science in Swaziland. The recommendations listed above thus require further investigation.

For the performance assessment tasks, perceptions were influenced by the method of administration of the tasks. The strengths of using group performance assessment in supporting learning, as well its limitations due to group dynamics, were demonstrated. Group assessment was motivated by the low availability of equipment in the schools. The perceptions that the assessment tasks/questions provided learning opportunities for the students supported the principle of the new assessment paradigm that emphasise assessment for learning and learning through assessment. Thus, as a start, further research is suggested in the following areas:

1. The impact of using the hands-on practical performance assessment model for classroom based assessment on the students' acquisition of conceptual and procedural knowledge and skills. Such a study could also explore whether students acquired scientific knowledge and/or socially related information while they engaged with these assessment models.

2. Students' perspectives on factors that influence their motivation and level of participation in group assessment activities.

3. The relationship between perceived task complexity and the nature and quality of students' answers to context-based questions.

REFERENCES

Ahmed, A. and Pollitt, A. (2001). *Improving the validity of contextualised questions*. Paper presented at the BERA Conference, Leeds, September 2001.

Ahmed, A. and Pollitt, A. (2000). *Observing context in action*. Paper presented at the International Association of Educational Assessment Conference, Jerusalem, May 2000.

Ames, C. and Archer, J. (1988). Achievement Goals in the Classroom: Students' Learning Strategies and Motivation Processes. *Journal of Educational Psychology*, 80 (3): 260-267

Aschbacher, P. R. (1993). *Issues in innovative assessment for classroom practice: barriers and facilitators.* CSE Technical Report 359 National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education University of California, Los Angeles.

Aschbacher, P. R. (1991). Performance assessment: state activity, interest and concerns. *Applied Measurement in Education*, 4(4): 275-288.

Baker, E. L., O'Neil, H. F. Jr. and Linn, R. l. 91993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48(12):1210-1218.

Bachor, D. G. and Anderson, J. O. (1994). Elementary teachers' assessment practices ad observed in the Province of British Columbia. *Assessment in Education: Principles, Policy and Practice*, 1(1): 63-94.

Barker, V. and Miller, R. (2000). Students' reasoning about basic chemical thermodynamics and chemical bonding: what changes occur during a context-based post-16 chemistry course? *International Journal of Science Education,* 22(11): 1171-1200.

Belt, S. T., Leisvik, M. J., Hyde, A. J. and. Overton, T. L (2005). Using a context-based approach to undergraduate chemistry teaching – a case study for introductory physical chemistry. *Chemistry Education Research and Practice,* 6(3): 166-179. [Online]. Available

Bennett, J. (2003). Context-based approaches to the teaching of science. In *Teaching and learning science: a guide to recent research and its applications*. London: Continuum.

Biehler, R. T. and Snowman, J. (1990). *Psychology applied to teaching.* 6th ed. Boston: Houghton Mifflin.

Black, P. (1998). *Testing: friend or foe? Theory and practice of assessment and testing.* London: Falmer Press.

Black, P. and Wiliam, D. (1998a). *Inside the black box: raising standards through classroom assessment.* [Online]. Available: http://www.pdkintl.org/kappan/kbla9810.htm [2004, October 17].

Black, P. and Wiliam, D. (1998b). Assessment and classroom learning *Assessment in Education: Principles, Policy and Practice*, 5(1): 7-74.

Bloom, B.S., Hastings, J.T. and Madaus, G.F. (Eds.) (1971). *Handbook on the formative and summative evaluation of student learning.* New York: McGraw-Hill.

Boaler, J. (1994). When do girls prefer football to fashion? An analysis of female under-achievement in relation to 'realistic' mathematic contexts. *British Educational Research Journal*, 20(5): 551-564.

Boaler, J. (1993). The role of contexts in the mathematics classrooms; do they make mathematics more "real"? *For the Learning of Mathematics*, 13(2): 12-17.

Borich, G. D. and Tombari, M. L. (1997). *Educational psychology: a contemporary approach.* 2nd ed. New York: Longman.

Boud D., Cohen, R. and Sampson, J. (1999). Peer learning and assessment. *Assessment and Evaluation in Higher Education*, 24(4): 413-426.

Broadfoot, P. (1996). Assessment and learning: power or partnership? Goldstein, H. and Lewis, T. (Eds.), *Assessment: problems, developments and statistical issues.* Chichester: John Wiley and Sons: 21-40.

Broadfoot, P. and Black, P. (2004). Redefining assessment? The first ten years of Assessment in Education. *Assessment in Education: Principles, Policy and Practice,* 11(1): 7-27.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education,* 10(2): 161-180.

Brookhart, S.M. and Bronowicz, D.L. (2003). "I don't like writing. It makes my fingers hurt": students talk about their classroom assessment. *Assessment in Education: Principles, Policy and Practice*, 10(2): 221-242.

Brookhart, S.M. and DeVoge, J.G. (1999). Testing classroom theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(4): 409-425.

Brown, G.T.L. (2004). Teachers' conceptions of assessment: implications for policy development. *Assessment in Education: Principles, Policy and Practice*, 11(3): 301-318.

Brualdi, A. (1998). *Implementing performance assessment in the classroom*. [Online]. Available: http://ericae.net/pare/getvn.asp?v=6andn=2 [2003, September 22].

Buckland, M.K. and Florian, D. (1991). Expertise, task complexity, and the role of intelligent information systems. *Journal of the American Society for Information Science,* 42(9): 635-643.

Campbell, B., Lubben, F. and Dlamini, Z. (2000). Learning science through contexts: helping pupils make sense of everyday situations. *International Journal of Science Education,* 22(3): 239-252.

Chang, S.-N. and Chiu, M.-H. (2005). The development of authentic assessments to investigate ninth graders' scientific literacy: in the case of scientific cognition concerning the concepts of chemistry and physics. *International Journal of Science and Mathematics Education,* 3: 117-140.

Cohen, L. and Manion, L. (2000). *Research methods in education*. London: Routledge.

Deemer, S. A. (2004). Classroom goal orientation in high school classrooms: revealing links between teacher beliefs and classroom environments. *Educational Research,* 46(1):73-90.

Dietel, R.J., Herman, J.L. and Knuth, R.A. (1991). *What does research say about assessment?* [Online] Available: http:/www.ncrel.org/sdrs/areas/stw_esys/4assess.htm [2002, October 14].

Dlamini, B. (2003). Looking at teaching science using contextualised vs non contextualised activities: some views from students in Swaziland. *Proceedings of the Eleventh Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education*, Volume I, 11-15 January. Waterford Kamhlaba UWC, Swaziland: 550-556.

Dlamini, B. (2000). *The impact of promoting an STS approach to teaching Science on the classroom practice of primary teachers in Swaziland*. Unpublished doctoral thesis. Faculty of Science, University of the Witwatersrand. Johannesburg, South Africa.

Dlamini, B., Lubben, F. and Campbell, B. (1996). Liked and disliked learning activities: responses of Swazi students to science materials with a technological approach. *Research in Science and Technological Education,* 14(2): 221-235.

Dochy, F.J.C. and McDowell, L. (1997). Introduction: Assessment as a tool for learning. *Studies in Higher Education,* 23(4): 279-298.

Donovan, M.S. and Bransford, J.D. (Eds.) (2005). *How students learn: science in the classroom*. [Online]. Available: http://www.nap.edu/catalog/11102.html [2006, November 06].

Downing, C. (not dated). *Ruminating on rubrics: an abstract of two papers*. [Online]. Available: http://www.accessexcellence.org/LC/SER/JA/rubrics.html [2004, October 14].

Elliot, S.N. (1995). *Creating meaningful performance assessments*. [Online]. Available: http://www.ed.gov/databases/ERIC_Digests/ed381985.html [2002, September 4].

Elliot, S.N., Kratochwill, T.R., Cook, J.L. and Travers, J.F. (2000). *Educational psychology: effective teaching, effective learning.* Boston: McGraw-Hill.

Fawcett, L.M., and Garton, A.F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology*, 75 (2): 157-169.

Gay, L.R. and Airasian, P. (2003). *Educational research: Competencies for analysis and applications*. 7th ed. Upper Saddle River, New Jersey: Merrill Prentice Hall.

George, J.M. (1999). Contextualised science teaching in developing countries: Possibilities and Dilemmas. *Proceedings of the 7th Annual Conference of the Southern African Association for Research in Mathematics and Science Education.* Harare Zimbabwe. 144-151.

Gijbels, D., van de Watering, G. and Dochy, F. (2005). Integration assessment tasks in a problem-based learning environment. *Assessment and Evaluation in Higher Education*, 30(1): 73-86.

Gibbs, G. and Simpson, C. (not dated). *Does your assessment support your students' learning?* [Online]. Available: http://www.open.ac.uk/science/fdtl/document/lit-review.pdf [2005, September 10].

Gipps, C. (1994). *Beyond testing: towards a theory of educational assessment.* London: Falmer Press.

Gipps, C. and Stobart, G. (2003). Alternative assessment. Kellaghan, T and Stufflebeam, D.L. (Eds.), *International Handbook of Educational Evaluation,* Vol. 9. Dordrecht: Kluwer Academic Publishers.

Glaser, B. and Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine.

Goldstein, H. and Lewis, T. (Eds.) (1996). *Assessment: problems, developments and statistical issues*. Chichester: John Wiley and Sons.

Goldstein, H. (1994). Recontextualising mental measurement. *Educational Measurement: Issues and Practice*, 13(1): 16-19.

Good, T.L. and Brophy, J. (1995). *Contemporary Educational Psychology*. 5th ed. New York: Longman.

Gott, R. and Duggan, S. (2002). Problems with the assessment of performance in practical science: which way now? *Cambridge Journal of Education,* 32(2): 183-201.

Griffin, M.M., Griffin, B.W., Warkentin, R.W., Quinn, G.P., Driscoll M.P. and McCown, R.R. (1995). *The effects of cooperative assessment on goals, perceived ability, self-regulation and achievement*. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA, April 18-22. (Eric Document Reproduction Service, No. ED 383 757).

Hall, S. (1993). *Task as a unit of teaching analysis*. Paper presented at the International Conference on Language in Education, Hong Kong, December 1993. (Eric Document Reproduction Service, No. ED 372 600).

Herman, J.L (1992). What research tells us about good assessment. *Educational Leadership*, 49(8): 74-78.

Herman, J.L., Klein, D.C.D. and Wakai, S. (1997). American students' perspectives on alternative assessment: do they know it's different? *Assessment in Education: Principles, Policy and Practice,* 4(3): 339-352.

Honeyfield, J. (1993). Responding to task difficulty: what is involved in adjusting the relationship between learners and learning experiences? Tickoo, M. L. (Ed.), *Simplification: Theory and application*. Anthology Series 31. (Eric Document Reproduction Service, No. 371 582).

Hord, S. (1987). *Evaluating educational innovation*. Pub. Croom Helm. ISBN 0-7099-4703-8.

James, A.R., Griffin, L.L. and France, T. (2005). Perceptions of assessment in physical education: a case study. *Physical Educator,* 62 (2): 85-95.

Kampfer, S.H., Horvath, L.S., Kleinert, H.L. and Kearns J.F. (2001). Teachers' perceptions of one state's alternative assessment: implications for practice and preparation. *Exceptional Children,* 67(3): 361-374.

Koosimile, A.T. (2004). Out-of-school experiences in science classes: problems, issues and challenges in Botswana. *International Journal of Science Education*, 26(4): 483-496.

Kotzé, G.S. (2003). The duality of coding assessment information. *South African Journal of Education,* 23(4): 338-344.

Krathwohl, D.R. (1998). *Methods of educational and social science research: an integrated approach*. 2nd ed. New York: Longman.

Lichtenstein, J. M. (1980). *A study of the implementation of the Swaziland Integrated Science Programme.* Unpublished doctoral thesis. Graduate College of Education, University of Illinois. Urbana-Champaign.

Linn, R.L. and Gronlund, N.E. (2000). *Measurement and Assessment in Teaching.* 8th ed. Upper Saddle River, New Jersey: Merrill (Prentice Hall).

Lubben, F., Campbell, B. and Dlamini, B. (1997). Achievement of Swazi students learning science through everyday technology. *Journal of the Southern African Association for Research in Mathematics and Science Education,* 1(1): 26-40.

Lubben, F. Campbell, B. and Dlamini, B. (1995). *In-service support for a technological approach to science education,* Serial no. 16 ISBN 0 902500 71 6 (ODA).

Maclellan, E. (2001). Assessment for learning: the differing perceptions of tutors and students. *Assessment and Evaluation in Higher Education,* 26(4): 307-318.

Madaus, G. F. (1985). Public policy and the testing profession-you've never had it so good. *Educational Measurement: Issues and Practice,* 4(4): 5-11

Madaus, G.F. and Kellaghan, T. (1993). Testing as a mechanism of public policy: a brief history and description. *Measurement and Evaluation in Counselling and Development*, 26(1): 6-10. [Online]. Available: http://web.ebscohost.com/ehost/detail?vid=1&hid=18&sid=30a8db21-5956-4829-b849-eee062025dcf%40sessionmgr8 [2006, September 24].

Malcolm, C. (2004). Context, relevance and interest. *Labtalk,* 48(1): 31-36.

Manyatsi, S.E. (1996). The interface between junior secondary and senior secondary school Chemistry in Swaziland. Unpublished doctorial thesis. Faculty of Science, University of the Witwatersrand. Johannesburg, South Africa.

McNair, S., Bhargava, A. Adams, L., Edgerton, S. and Kypros, B. (2003). Teachers Speak Out on Assessment Practices. *Early Childhood Education Journal*, 31(1): 23-31.

Merriam, S. (1998). *Qualitative research and case study applications in education.* San Francisco, CA: Jossey Bass Publishers.

Min, H. and Xiuwen, Y. (2001). Educational Assessment in China: lessons from history and future prospects. *Assessment in Education: Principles, Policy and Practice*, 8(1): 5-10.

Moni, K.B., van Kraayenoord, C.E. and Baker, C.D. (2002). Students' perceptions of literacy assessment. *Assessment in Education*, 9(3): 319-342.

Moorcroft, T., Desmrais, K.H., Hogan, K. and Berkowitz A.R. (2000). Authentic assessment in the informal setting: how it can work for you. *Journal of Environmental Education*, 31(3): 20-24.

Neesom, A. (2000). *Report on teachers' perceptions of formative assessment*. London: QCA.

Nitko, A.J. (2004). *Educational assessment of students.* 4th Ed. Pearson Merrill Upper Saddle River, New Jersey: Pearson (Prentice Hall).

Nuttall, D.L. (1992). Performance assessment: the message from England. *Educational Leadership*, 49(8): 54-57.

OECD/PISA (1999). *Measuring student knowledge and skills: a new framework for assessment.* Paris: OECD Publication Service.

Pollitt, A. and Ahmed, A. (2000). *Comprehension failures in educational assessment*. Paper presented at the European Conference on Educational Research, Edinburgh, September 2000.

Pollitt, A., Marriott, C. and Ahmed, A. (2000). *Language, contextual and cultural constraints on examination performance.* Paper presented at the International Association for Educational Assessment, Jerusalem, Israel. May 2000.

Popham, W.J. (1999). *Classroom assessment: what teachers need to know.* 2nd ed. Boston: Allyn and Bacon.

Putsoa, B. (1999). Linking School Science and Technology with that of the Community and Workplace. *Proceedings of the 9th Symposium of the International Organisation for Science and Technology Education*, Durban South Africa: 502-510.

Putsoa, B., Dlamini, C., Dlamini, E., Dlamini, N., Dube, T., Khumalo, E., Masango, T., Ndlela, F., Nhlengethwa, L. and Tsabedze, S. (2003). Comparing learners' performance in contextualised and non-contextualised science tasks. *Proceedings of the Eleventh Annual Conference of the Southern African Association for Research in Mathematics, Science and Technology Education,* Volume II, 11-15 January. Waterford Kamhlaba UWC, Swaziland: 25-30.

Putsoa, B., Manyatsi, S. and Dlamini, M. (2000). Quality and Equity in Mathematics, Science and Technology Education for Secondary Education. Proceedings of the *Conference on Mathematics, Science and Technology (MST) Education in the Next Millennium*. Department of Science Education, National University of Lesotho, Roma, Lesotho: 89-103.

Putsoa, B. and Maphalala, T. (1999). *Linking School Science with Technology and Industry: Acids, alkalis and gases teachers' guide*. University of Swaziland.

Ramsden, J. (1997). How does a context-based approach influence understanding of key chemical ideas at 16+? *International Journal of Science Education,* 19(6): 697-710.

Rhoton, J. and Bowers, P. (Eds.) (1997). *Issues in Science Education*. Arlington, VA: National Science Teachers Association and National Science Education Leadership Association.

Roberts, L. and Wilson, M. (1998). *Evaluating the effects of an integrated assessment system on science teachers' assessment perceptions and practice.* [Online]. Available: www.bearcentre.berkeley.edu/publication/teacher98.pdf [2006, November 5].

Roberts, R. and Gott, R. (2004). Assessment of Sc1: alternatives to coursework. *School Science Review,* 85(313):103-108.

Robson, C. (2002). *Real world research: A resource for social scientists and practitioner-researchers.* Oxford: Blackwell.

Roeber, E. D. (1996). Guidelines for the development and management of performance assessments. *Practical Assessment, Research and Evaluation,* 5(7) [Online]. Available: http://pareonline.net/getvn.asp?v=5and=7 [2002, October 14].

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science,* 18(2):119-144.

Sambell, K. and McDowell, L. (1998). The construction of the hidden curriculum: messages and meanings in the assessment of student learning. *Assessment and Evaluation in Higher Education,* 23(4): 391-402.

Sambell, K., McDowell, L and Brown, S. (1997). "But is it fair?" An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation,* 23(4): 349-379.

Sanders, W.L. and Horn, S.P. (1995). Educational assessment reassessed: the usefulness of standardised and alternative measures of student achievement as indicators for the assessment of educational outcomes. *Education Policy Analysis Archive,* 6(6). [Online]. Available: http://epaa.asu.edu/epaa/v3n6.html [2004, November 25].

Schäffner, M. Burry-Stock, J. A., Cho G-P., Boney, T. and Hamilton, G. (2000). *What do kids think when their teachers grade?* Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA, April 24-28.

Schumacher, S. and McMillan, J.H. (1993). *Research in education: a conceptual introduction.* 3rd ed. New York: Harper Collins College Publishers.

Sebatane, E.M. (1998). Assessment and classroom learning: a response to Black and Wiliam. *Assessment in Education: Principles, Policy and* Practice, 5(1): 123-130.

Segers, M. and Dochy, F. (2001). New assessment forms in problem-based learning: the value-added of the students' perspective. *Studies in Higher Education*, 26(3): 327-343.

Sharikzadeh, N. (2003). Assessment in constructive learning environments. *Proceedings of the International Conference on Emerging Technologies*, Minneapolis, Minnesota, August 25-26. [Online]. Available: http://citeseer.ist.psu.edu/691767.html [2004, October 09].

Shavelson, R. J. (not dated). On the integration of formative assessment in teaching and learning with implications for teacher education. [Online]. Available: http://www.stanford.edu/dept/SUSE/SEAL/Reports_Papers/On%20the%20Integration%20of%20Formative%20Assessment_Teacher%20Ed_Final.doc [2007, February 16].

Shavelson, R.J. and Baxter, G.P. (1992). What we've learned about assessing hands-on science. *Educational Leadership,* 49(8):20-25.

Shavelson, R.J. Baxter, G.P. and Pine, J. (1991). Performance assessment in science. *Applied Measurement in Education,* 4(4):347-362.

Shepard, L.A. (2000). The role of classroom assessment in teaching and learning. Richardson, V. (Ed.), *Handbook of Research on Teaching.* 4th ed. Washington DC: American Educational Research Association.

Shepard, L.A. (1995). Using assessment to improve learning. *Educational Leadership,* 52(5): 38-43.

Silbermann, A. (not dated). *A closer look at alternative assessment: Mathematics teachers discuss their implementation of new assessment techniques*. [Online]. Available: http://www.learner.org/theguide/assess.html [2004, October 14].

Slater, T.F. (not dated). *Performance assessment*. [Online]. Available:

http://www.flaguide.org/extra/download/cat/perfass/perfass.pdf [2005, May 4].

Solano-Flores, G., Jovanovic, J., Shavelson, R.J. and Bachman, M. (1999). On the development and evaluation of a shell for generating science performance assessments. *International Journal of Science Education*, 21(3): 293-315.

Southwest Educational Development Laboratory (1996). Assessment: a window of learning. *Classroom Compass*, 2(2) [Online]. Available: http://www.sedl.org/scimath/compass/v02n02/assessment.html [2005, September 9].

Stepanek, J. (2002). Classroom assessment and the pursuit of illuminating feedback. *Northwest Teacher*, 3(2): 2-3, 13. [Online]. Available: http//www.nwrel.org/msec/images/nwteacher/winter2002/classroom.pdf [2003, October 17].

Stray, C. (2001). The shift from oral to written examinations: Cambridge and Oxford 1700-1900. *Assessment in Education: Principles, Policy and Practice*, 8(1): 33-50.

Struyven, K., Dochy, F., and Janssens, S. (2003). Students' perceptions about new models of assessment in higher education: A review. Segers, M., Dochy, F. and Cascallar, E. (Eds.), *Optimising new modes of assessment: in search of qualities and standards*. Dordrecht; Kluwer Academic Publishers: 171-223.

Taber, K.S. (2003). Examining structure and context – questioning the nature and purpose of summative assessment. *School Science Review*, 85(311): 35-41.

Tao, P-K. (1999). Peer collaboration in solving qualitative problems: the role of collaborative talk. *Research in Science Education*, 29(3): 365-383.

The Government of Swaziland (2005). *International General Certificate of Secondary Education: consultative document*. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (2000a). *National Assessment Framework*. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (2000b). *Proposal for the introduction of (H)IGCSE into Swaziland School*. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (1999). *National policy statement on education*. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (1997a). *Science Grade 1 to Grade 12: A statement of policy*. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (1997b). *Broad Aims of Science Education in Primary/Secondary Schools in Swaziland.* Paper presented at the Science Syllabus Development orientation workshop, William Pitcher Teacher Training College. Ministry of Education. Mbabane, Swaziland.

The Government of Swaziland (1995). *National Development Strategy*. Ministry of Economic Planning and Development. Mbabane, Swaziland.

The Government of Swaziland (1985). *Reform Through Dialogue: Report of the National Education Review Commission*. Mbabane, Swaziland.

The Government of Swaziland (not dated). *Swaziland Intergrated Science Programme Volume 1*. Ministry of Education. Mbabane, Swaziland.

Tittle, C.K. (1994). Toward an educational psychology of assessment for teaching and learning: theories, contexts, and validation arguments. *Educational Psychologist,* 29(3): 149-162.

van Merriënboer, J.J.G., Kirschner, P.A. and Kester, L. (2003). Taking the load off a learner's mind: instructional design for complex learning. *Educational Psychologist*, 38(1): 5-13.

Veronesi, P.D. (1997). *A case study of alternative assessment: student, teacher, and observer perceptions in a ninth grade biology classroom*. [Online]. Available: http://www.ed.psu.edu/ci/Journals/97pap39.htm [2006, February 13].

Webb, N., Nemer, K., Chizhik, A. and Sugrue, B. (1997). *Equity issues in collaborative group assessment: group composition and performance.* CSE Technical Report 457. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education, University of California, Los Angeles.

Wiggins, G. (1996/1997). Practicing what we preach in designing authentic assessments. *Educational Leadership*, 54(4): 18-25.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership*, 46(7): 41-47.

Wilbrink, B. (1997). Assessment in historical perspective. *Studies in Educational Evaluation*, 23: 31-48.

Wiliam, D. (2001). What is wrong with our educational assessments and what can be done about it? *Education Review,* 15(1): 57-62.

Wiliam, D. and Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of assessment, *British Educational Research Journal,* 22(5): 537-548.

Winegardner, K. (not dated). *The case study method of scholarly research*. [Online]. Available: http://www.tgsa.edu/online/cybrary/case1.html [2001, July 30].

Wistedt, I. (1994). Everyday common sense and school mathematics. *European Journal of Psychology in Education,* 11(1): 59-67.

Yang, F-G. (2004). Exploring high school students' use of theory and evidence in an everyday context: the role of scientific thinking in environmental science decision-making. *International Journal of Science Education,* 26(11): 1345-1364.

Yung, B.H.W. (2001). Three views of fairness in a school-based assessment scheme of practical work in biology. *International Journal of Science Education,* 23(10): 985-1005.

APPENDICES

APPENDICES

APPENDIX IA

Syllabus objectives, content and learning outcomes

Integrated Science Syllabus Objectives

The general objectives presented below have been taken from a section of the Integrated Science syllabus. Objectives that have been identified as the most relevant for the learning outcomes explored in the study are presented, rather than all the syllabus objectives.

As arranged in the integrated Science syllabus the objectives are categorized into three sections namely: knowledge and understanding, attitudes, and practical skills. The syllabus objectives are the same for the units have been generated for all the four units to be used in the study: Electricity; Air and Life; Force, Support and Movement; and Acids, Alkalis and Gases. (the last two units were not covered due to time availability).

A Knowledge and Understanding:

Pupils should acquire knowledge and understanding of

a) scientific facts and concepts concerning the environment
b) the use of scientific instruments in scientific experiments
c) scientific vocabulary and use of the vocabulary in communication
d) scientific knowledge to be able to select appropriate knowledge and apply it to new situations
e) scientific knowledge to be able to analyse data and draw conclusions
f) scientific knowledge to be able to solve a scientific problem.

B Attitudes: Pupils should acquire

a) an awareness of the contribution of science to the social and economic life of the community
b) respect for the dangerous nature of certain substances and phenomena
c) objectivity in observation

C Practical skills: Pupils should acquire

a) science-based skills (observation, classifying, measuring, communicating, predicting, inferring, interpreting data, controlling variables, hypothesizing, experimentation)
b) an ability to demonstrate the use of some experimental techniques including several skills (Swaziland Examinations Council, 1999/2000:30).

Unit 6. Electricity: Syllabus content:

Electrostatics; acquaintance with simple phenomena; attraction and repulsion.

Conductors and insulators; electric circuits (use of circuits boards assumed and simple conventional circuit diagrams included).

Cells, batteries and generators as sources of electrical power; voltage as `electric pressure' determining rate at which energy is available; voltmeters (working of cells, generators and voltmeters NOT expected).

Series connections in cells (*sic*); structure and rating of light bulbs (volts and watts). Measuring current (use but not theory of ammeters); nature of current electricity; the electron and an elementary idea of the atom built up of smaller particles, some of which can be removed in certain circumstances. Resistance as opposition to flow of electrons; dependence of resistance and current in series and parallel circuits (qualitative treatment only). Short circuits; heating effects of current, fuses, domestic applications of heating effects (Swaziland National Examinations Council, 1999:32).

Learning Outcomes

(Generated by the researcher from the syllabus content and the general objectives)

Learning Outcomes	
It is expected that pupils will	
i)	Explain, using correct terminology how objects become electrically charged
ii)	Explain the electrostatic behaviour of certain material in terms atomic structure and charges differences
iii)	Classify given material into conductors of electricity and electrical insulators
iv)	Give the properties of conductors and insulators of electricity
v)	Use appropriate equipment to construct electric circuits to show a simple circuit, series circuit and a parallel circuit
vi)	Use appropriate symbols when representing circuits as conventional circuit diagrams
vii)	Connect ammeter and voltmeter correctly in a circuit
viii)	List cells, batteries and generators as sources of electrical energy
ix)	Measure electric current at different positions in a circuit
x)	Identify a suitable light bulb for a given situation on the basis of its ratings
xi)	Measure the voltage across different components of a circuit
xii)	Give operational definition of voltage as “electric pressure” determining the rate at which energy is available?
xiii)	Give operational definition current as a flow of charge/electrons?
xiv)	Use appropriate apparatus to determine the relationship between current voltage and resistance in different types of circuits (qualitative only)
xv)	Calculate the electrical resistance of a wire using current and voltage measurements
xvi)	Give operational definition electrical resistance as opposition to flow or electrons
xvii)	Identify the wire with the least electrical resistance given data about the wires
xviii)	Explain the variation of current in parallel and series circuits in terms of electrical resistance
xix)	Explain how the light bulb works to emit light in terms of the heating effects of electricity
xx)	Explain the use of fuses in protecting electrical appliances from short circuits
xxi)	Conduct an investigation on the domestic use of electricity
xxii)	Apply knowledge of the electrical conductivity and resistance of materials to solve a given related problem

APPENDIX IB

Syllabus content and learning outcomes: Air and Living Things

Syllabus objectives are the same as presented above

Unit 7: Air and Living Things:Syllabus content

Simple preparation of oxygen and carbon dioxide; identification and simple properties of oxygen, carbon dioxide and nitrogen (both lime water and bicarbonate indicator tests for carbon dioxide expected) – word equations introduced as supplementary means of recording some of the above changes. Comparative combustion in air and oxygen; air as ‘diluted’ oxygen; approximate composition of air, including noble gases; nitrogen as a relatively inert gas. Combustion of carbohydrates and other organic materials in oxygen.

Differences between inhaled and exhaled air (including temperature) comparisons of respiration and combustion; breathing – the lungs, diaphragm, etc.; respiration as gaseous interchange in blood, tissues etc; maintaining body temperature, introduction to circulatory systems; effects of exercise on breathing and pulse rates and (partial) explanation.

Respiration in plants and germinating seeds; photosynthesis; carbon dioxide/oxygen cycle (Swaziland National Examinations Council, 1999/2000:32-34**).**

Learning Outcomes

	It is expected that pupils will
i)	List the components of air
ii)	Write word equations for oxygen preparation reactions.
iii)	Write word equations for carbon dioxide preparation reactions.
iv)	Identify oxygen gas by its effect on a glowing splint and bicarbonate indicator
v)	Identify carbon dioxide gas by its effects on a lime water and bicarbonate indicator
vi)	Explain the variation in combustion in air and in oxygen in term of oxygen concentration levels in air.
vii)	State carbon dioxide and water as products of combustion of carbohydrates
viii)	Explain inhaling and exhaling of air in terms of pressure changes in the chest cavity
ix)	Compare and contrast inhaled and exhaled air
x)	Describe the role of oxygen in burning, respiration and rusting
xi	Define respiration in terms of gaseous exchange in blood and tissues
xii)	Describe the effects of exercise on the breathing rate.
xiii	Explain the changes in breathing rate and pulse rate occurring during exercise
xiv)	Link ability of plants to photosynthesis to the presence of light
xv)	Infer the conditions of plant growth from leaf starch test results.
xvi)	State water, oxygen and warmth as conditions necessary for seed germination
xvii)	Explain failure of seeds to germinate in terms if absence of oxygen and water and low temperature.
xviii)	Carry out starch tests given samples to be tested
xix	Infer the presence and amount of starch from starch test results
xx	State that oxygen is important for respiration in organisms

APPENDIX IIA

Performance Assessment Tasks: Electricity

Performance Assessment Task 1: Electricity

NB: Some adjustments have been made to economise on space

Group Number ……… School ………………………………… Date ……………

Names of Group Members…………………………………………………………

Instructions:
You will work on this practical task in groups. You are allowed to discuss among yourselves in your group and you may check with the teacher for some information that you may need.

In this activity you will plan and carry out an investigation on the electrical conductivity of a set of material. You will use an electric circuit to test the materials.

You will then classify the materials into conductors and non-conductors of electricity.

You will also choose the most suitable material to use in a given situation.

Material provided:

circuit board connecting wires, 2 x 1.5V cells,	switch lamp and holder various materials labelled P, Q, R, S, T, U.

Procedure:

Use the space below to write out your plan of how you will test the given materials and indicate how you will record your results. Your plan should show the necessary diagrams. Show your plan to the teacher before you begin testing the material.

Connect your electric circuit using the material provided and test that it is working.

Use the circuit to test for the electrical conductivity of the given material.

Questions:

a) Classify the materials you tested into conductors and non-conductors. …..… (2)

Why have you classified them in this way? …………………….…………..… (2)

b) The picture below shows the top part of a power-cable pole.

Part A

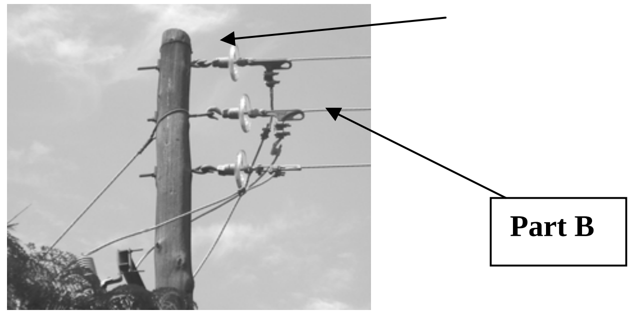


Figure 1 Top part of power-cable pole

i) Which of the materials you tested would be most suitable for replacing **Part A,** labelled in the picture? ……………………………….…………….….. (1)

ii) Explain your answer. ………………………………………………….... (2)

c) i) Which of the materials you tested would be most suitable for replacing **Part B,** labelled in the picture?
……………………………………….………………………..……….. (1)

ii) Explain your answer. ……………………………………….…………… . (2)

Performance Assessment Task 1: Electricity Scoring Guide

School ………. **Date……………….**

Rating scales used to guide grading of performances. They will be used to guide grading of each group of pupils working on a given task; each member gets the same mark. Pupils discuss their ideas and can consult the teacher or workbook books should they need to refer to a diagram.

NB: Please outline to them what you assess.

Activity 1: Please rate the performance of the pupils / group according to marks given in [**bold].**

Assessment Aspects	**Assessed ability**	**Mark / Ratings**
		Group Number
		1
A: Planning of testing of materials	Correctness of diagram of circuit to be used to test materials **[0-4]**	
Plans show	Data recording table with appropriate headings e.g. (material (P,Q….); light bulb, brightness of light, conduction … **[0-4]**	
B: Manipulation of equipment	Connects circuit that matches diagram **[0-2]**	
	Tests materials correctly **[0-2]**	
C: Capturing data	Makes accurate observations **[0-2]**	
	Records data correctly **[0-2]**	
D: Analysis and Interpretation of data	Question in paper **[12 marks]**	
E: Co-operativeness	Group works co-operatively with one another (no one pupil dominates) **[0-2]**	
Total	**30 marks**	

Comments:

Performance Assessment Task 2: Electricity

Group Number ……… School …………………………… Date ……………………….

Names of Group Members ………………………………………………………………..

Instructions:

You will work on this practical task in groups. You are allowed to discuss among yourselves in your group and you may check with the teacher for some information you may need.

In this activity you will work in your group.

i) You will assemble a circuit that includes a voltmeter and ammeter.

You will use the circuit to

ii) measure the voltage across different pieces of wires, and

iii) measure the current through the different pieces of wires

You will then

iv) calculate the resistance of each the given wires.

Material provided:

3 x 1.5V cells (when you are ready) 1 circuit board 1m long pieces of wires marked H, I, J, K	1 voltmeter 1 ammeter switch

Procedure:

1. Use a separate sheet to write out a plan of how you are going to connect the circuit and how you will **record your results**. Show your plan to the teacher.
2. Set up your equipment and use it to measure
   i) the current passing through the wire and
   ii) the voltage across each wires
3. Record your results in an appropriate table
4. Calculate the resistance of each piece of wire.

Questions

a) List the wires in order of increasing resistance beginning with the wire with lowest resistance ……………………..……………………………………………….….(2)

b) Thembi wants to design a torch. She asked her father for some wires for her torch project.
   i) Which of the wires you tested could Thembi's father give her? …………….. (2)
   ii) Why do you think this wire would work best for Thembi's project? ……….. (2)

c) Suppose you want to work on a project to show that a "current carrying wire" can be used to boil water.
   i) Which wire would be most suitable for your project? ………….……….…. (2)
   ii) Why would the chosen wire be most suitable? ……………………………… (2)

Performance Assessment Task 2: Electricity Scoring Guide

School ………………………………. Date……………………………….

Rating scales used to guide grading of performances. They will be used to guide grading of each group of pupils working on a given task; each member of the group gets the same mark. Pupils discuss their ideas and can consult the teacher, but marks need to be deducted for specific help on working out the task.

NB: Please outline to them what you assess.

Task 2: Please rate the performance of the pupils / group according to the marks stated in bold

Assessment Aspects	**Assessed ability**	**Mark / Ratings**	
		Group # 1	**Group # 2**
A: Planning of testing of materials	Diagram of circuit for showing correct set up for (resistor, cells, switch, ammeter and voltmeter connection) **[0-5 marks]**		
Plans show	Data recording table with appropriate headings (wire (H, I, J, K), readings on ammeter & voltmeter; resistance or (V/I) **[0 – 4 marks]**		
B: Manipulation of equipment	Connects circuit that matches diagram **[0 - 2 marks]**		
C: Capturing data	Measure the currents correctly (incl units) **[0 - 2 marks]**		
	Measure voltage correctly (incl units) **[0 - 2 marks]**		
	Records data correctly **[0 - 2 marks]**		
D: Analysis and Interpretation of data	Calculates resistance correctly **[0 - 2 marks]**		
	Questions on question paper : Identifies wires: with high/ low resistance, suitable for given situation **[10 marks]**		
F: Co-operativeness	Group members work co-operatively with one another **[0-1 mark]**		
Total	**30 marks**		

Comments:

Materials needed for Performance Task 2 on Resistance

6-8 circuit boards, 1 switch per group = 6-8 switches
3 cells per group = 18 - 24 max 6 connecting wires per group = 42-48 connectors
4 Resistance wires 1m long each /group (shared if necessary) = 24-32
Sample readings accept reading that ranges around these data

Resistor	**Ammeter reading (A)**	**Voltmeter reading (V)**	**Resistance (ohms) (Voltage/current)**
H = thick nichrome wire	0.6 0.6	2.6 2.0	4.3 3.3
I = copper wire	2.2 1.5	0.1 0.5	0.05 0.33
J = constantan	0.2 0.2	3.0 3.5	15 17.5
K = thin nichrome wire	0.25 0.25	3.5 3.0	14 12

APPENDIX IIB

Performance Assessment Task: Air and Living Things

Group Number ……… School …………… Date…………………….

Names of Group Members ……………………………………..………………….

Instructions:
You will work on this practical task in groups. You are allowed to discuss among yourselves in your group and you may check with the teacher for some information you may need.

Assessment Task 1:
In this activity you will **plan** and **carry** out an investigation to find out the identity of the given colourless gases, Gas A, Gas B and Gas C. The three gases are thought to be oxygen, carbon dioxide and air.

Assessment objective:

In this activity you will be assessed on the following:

i) Your plan of how you will carry out the test on the given samples of gases including what you expect from the tests
ii) Correct handling of the equipment when testing the gases.
iii) Observations you make.
iv) Proper recording of all your observations.
v) Decision on what each sample of gas is.
vi) Use of the information obtained from the investigation and other knowledge to answer questions.

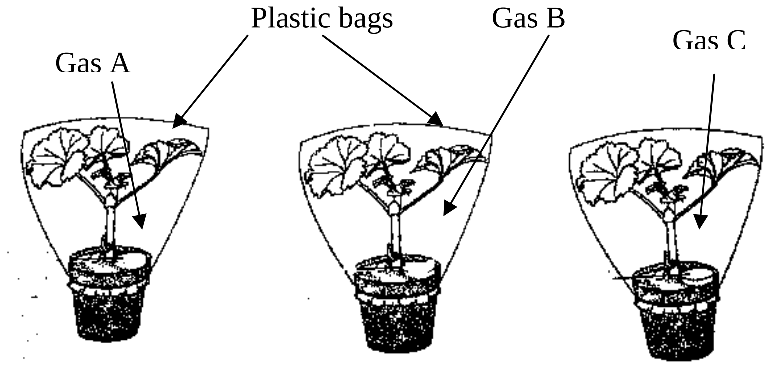
Material provided
Six test tubes of gases labelled Gas A, Gas B and Gas C in a test tube rack
Other materials are available on request (ask for them from your teacher when you have decided on them).

Procedure

i) In your group discuss and plan how you will test for the gases using the given material. Write out your plan giving a brief description of what you will do when testing the gases and how you will write down your results.
ii) Show your plan to the teacher before proceeding with testing the gases.
iii) Carry out the tests on the gas samples.
iv) Record your observation.

Questions

1. What do you think Gas A is? …………………………………………….………(1)
   What do you think Gas B is? ……………………………………………….….……(1)
   What do you think Gas C is? ……………………………………………………..(1)
2. Welile's father is a welder. He normally uses electricity for welding. One day he decided to use a hot flame to weld two metal sheets together.
   i) Which of the gases you tested would you suggest he uses when doing this job?...(1)
   ii) Explain why you think the gas you suggested is suitable. ....(2)
3. Three potted plants were covered with plastic bags that had been filled with the gases you tested, as shown in the diagram below. The plants were then placed in a sunny place. After 48 hours leaves from each plant were taken and tested for starch.



i) Which plant would produce the most starch? ……………….….……………….. (1)
ii) Explain your choice of plant. …………………………………………….… (2)

4. In which of the three gases would iron objects be damaged quickly by rusting? … (1)

Performance Assessment: Task: Air and Living Things Scoring Guide

School ………………………………. Date……………….
Rating scales used to guide grading of performances. They will be used to guide grading of each group of pupils working on a given task; each member gets the same mark. Pupils discuss their ideas and can consult the teacher or workbook books should they need to refer to a diagram.
NB: Please alert groups on what they will be assessed on as outlined in their assessment tasks.
Activity 1: Please rate the performance of the pupils / group according to marks given in [**bold].**

Assessment Aspects	**Assessed ability**	**Mark / Ratings**
		Group Number
		1
A: Planning for testing of gases **[0-8]**	Description to what to do to each gas: e. g. open test tube carefully **[1]**, insert glowing splint **[1½]** or add lime water to each gas**[1½]**, observe**[1]**; expected observations **[3]**	
B: Manipulation of equipment **[0-4]**	Carry out test correctly (taking care to minimise loss of gas **[2],** shake test for lime water test**[2]**	
C: Capturing data Results table shows **[0-6]**	Data recording table with appropriate headings e.g. test, observation for each test and gas (i.e. two tests for each gas] **[3]** observations [relight/puts out/no effect to glowing splint **[1½]**; milky lime water, no change **[1½]**	
D: Analysis and Interpretation of data and application	Questions in paper **[10 marks]**	
E: Co-operativeness	Group works co-operatively with one another (no one pupil dominates) **[0-2]**	
Total	**30 marks**	

Comments:

APPENDIX IIIA

Context-Based Test: Electricity

Name ……………………… Class …………… Date ……………

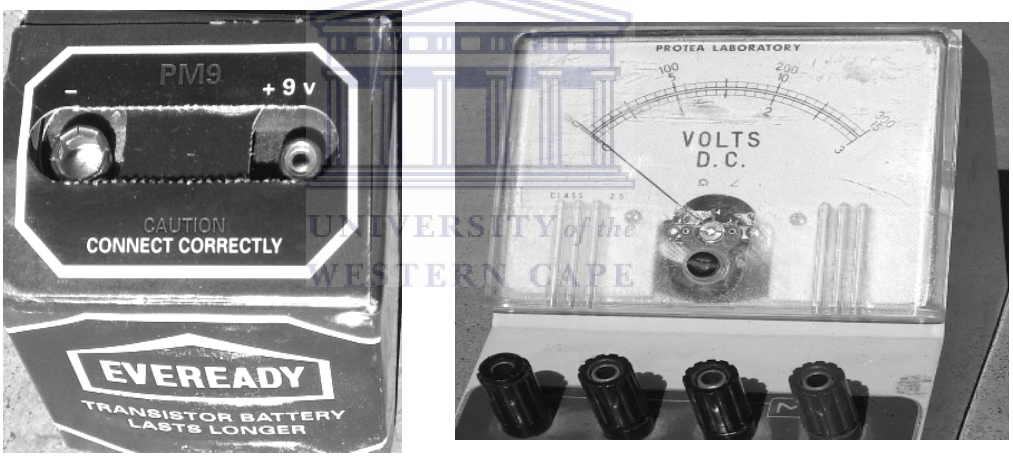
Instruction: Answer all the questions. Use the spaces provided to answer your questions

1. As Nosipho was walking home one evening she felt hot and decided to take off her jersey. It was already getting dark and as she took it off, she noticed some sparks and heard a crackling sound.

   a) What name would you give to the "source" of the sparks and sound?
   ………………………………………………………………………… (1)

   b) Explain how the sparks coming from Nosipho's jersey are produced.
   ……………………………………………………………………...…. …… (2)

2. Londiwe went to the shop to buy a PM 9 battery (Figure 1 below) for her radio. She asked the shopkeeper to check whether the battery was "full" or not.
   a) Draw on Figure 1 below how the shopkeeper would test the battery (2)



3V 15V 300V +

PM 9 Battery Voltmeter

Figure 1: A voltmeter and a PM 9 battery

3. Mandla thinks that bulbs connected in series use less current than the same number of similar bulbs connected in parallel. He decided that all the bulbs in **his house** must be wired in series to "save money". Though he saved some money, he was disappointed because the lighting in his house did not work the way he had expected.

   a) Name **two** possible sources of electrical energy Mandla could use to light his house.

   ……………………………………………………………………………..(2)

b) What **two possible observations** by Mandla could result from wiring his house light using a series circuit?
…………………………………………….……………………………….. (2)

c) What do you think causes the observations Mandla made (as given in (b) above)?
……………………………………………………………………………………. (2)

d) Draw a circuit diagram to represent a circuit that Mandla could use to wire his house to solve the problems he observed with his new circuit. Use four bulbs (lamps) and three cells in your circuit. (3)

e) i) Select from the bulbs, A and B, shown below the type of bulb that would be most suitable for lighting Mandla's house? ………………... (1)

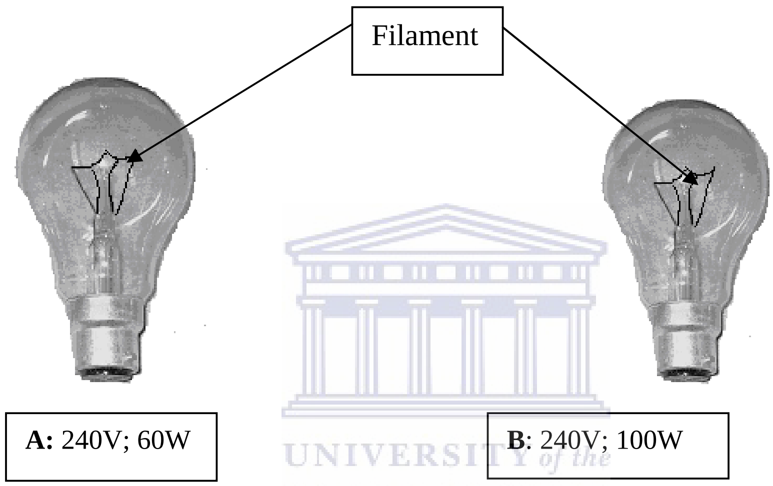
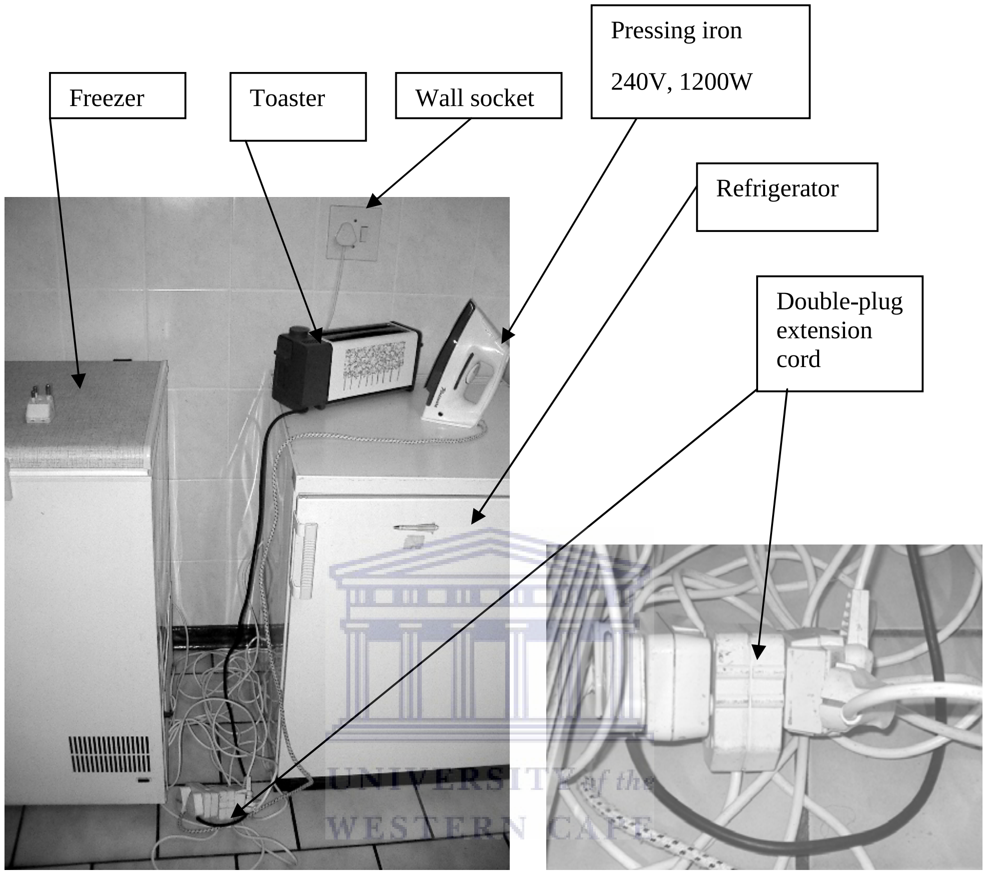


Figure 2: Light bulbs with different ratings

ii) Explain your choice of light bulb. …………………………………..……….. (2)

iii) Describe how a light bulb, like the ones shown in the picture, works to give out light when switched on. …………………..……………………………………………… (3)

4. The picture below shows how Thandeka connected her electrical appliances, a freezer, a pressing iron, a toaster and refrigerator.



Enlargement of connection of plugs

Figure 3: Several electrical appliances connected to one socket using an extension cord

a) i) Give **two** reasons why it is not good to connect many electrical appliances to one socket and extension cord as Thandeka did.
…………………………………………………………………………………………............ (2)

b) What **two** things might Thandeka do in order to use electrical appliances safely?.. (2)

c) What energy changes are taking place in the pressing iron when it is in operation during the ironing of clothes? ……………………………..……………………. (2)

d) The pressing iron in Figure 2 above is marked 240V and 1200W.

i) What does the V in 240 V stand for? ……………………………………….…… (1)
ii) What does the W in 1200W stand for? ……………………………………….. (1)
iii) What does the 240V tell you about the pressing iron? ……………................... (2)
iv) What does the 1200W tell you about the pressing iron? ……………………..….. (2)

v) Calculate the current in the pressing iron circuit when the iron is drawing maximum current. (Use the formula **Power = Voltage (V) x Current (I)**
…………………………………………………………………………….…… . (2)

vi) What does the term “electrical resistance” mean? ……………………….... … (2)

vii) Complete the following sentence by writing the word **high** or the word **low.**

A good conductor of electricity has a …………… resistance. (1)

viii) Calculate the electrical resistance of the pressing iron? …… ……………… (3)

Total Marks 40

>>

Pupil Questionnaire
(Writing space was provided)

Please answer the following questions to express your feelings about the test.

1. Which questions of the test do you like?
2. What makes you like these questions?
3. Which questions do you **not** like?
4. Why do you **not** like these questions?
5. Describe other feelings or ideas you have about this test.
6. Would you like your other science topic tests to ask questions that are similar to the questions asked in this test (i.e. use situations you see around you)?
7. Explain your answer.

APPENDIX IIIB

Context-Based Test: Air and Living Things

Name ………………… Class…………… Date …………… School……………

Instruction: Answer all the questions. Use the spaces provided to answer your questions

1. Mrs. Mlipha planted maize seeds that had fallen into oil by mistake. The seeds germinated later than Mrs. Mlipha had expected. She thought the oil caused the late germination of the seeds.

   Explain how the oil caused the seeds to take longer to germinate?

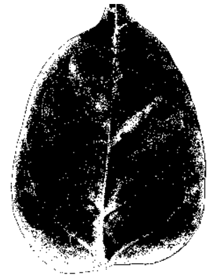
…………………………..…………………………………………………………...… (2)

2. The Ministry of Agriculture and Cooperatives in Swaziland employs Extension Officers (*Balimisi*) to encourage farmers to practice good ways of farming. The Extension Officers tell farmers to grow fewer large trees in their fields because crops do not grow well under trees, as the picture below shows.

a) What causes the poor growth of crops that are planted under trees?

……………………………………………………………………………………….…… (3)

Leaf A



Leaf B

b) The leaves below show the results of a starch test on green leaves picked from plants growing around the tree. The test was performed after the chlorophyll was removed from the leaves.

Which of the leaves were picked from crops growing under the tree? …………. (1)

Explain your choice of leaf. ………………………………………….……………… (1)

3. Jabulile is a member of her school's athletics team because she can run very fast.

a) During each race Jabulile's breathing changes.

i) What **two** changes occur in Jabulile's **breathing**? …………………………… (2)

ii) How do the changes that occur in Jabulile's **breathing** help her during each race?

……………………………………………………………………………………………… (4)

iii) Describe what happens in Jabulile's chest that makes her able to breathe in and out.

……………………………………………………………………………..……….. (4)

iv) What **four** differences would you find if you compared the air Jabulile breathed out and the air she breathed in?

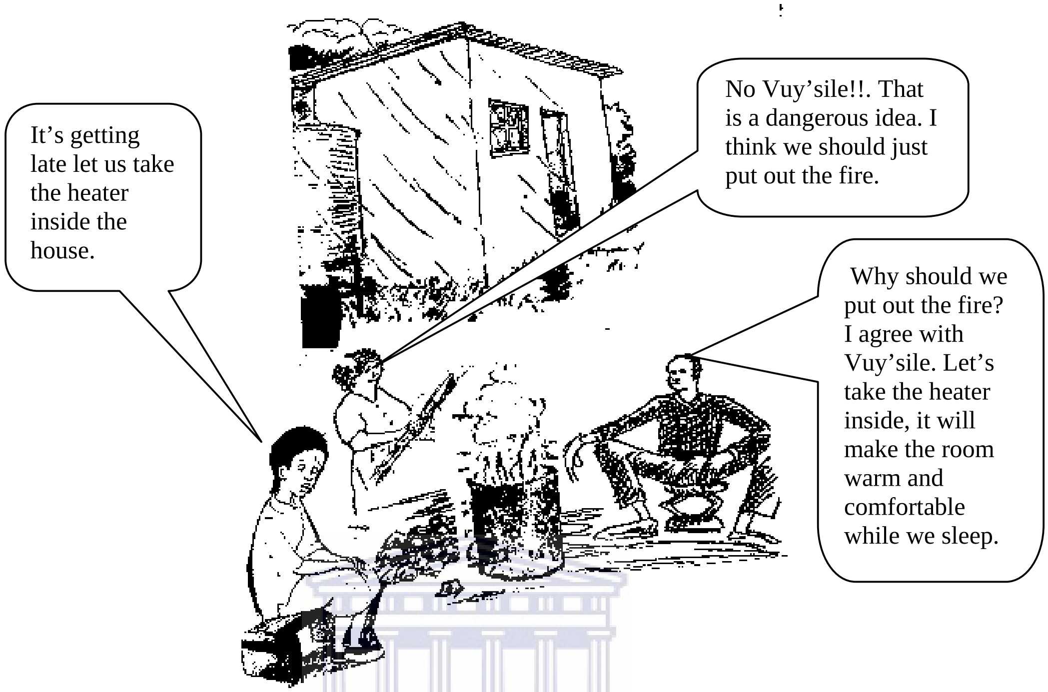
…………………………………………………………………………………………… (4)

b) Jabulile's heart beats faster during a race. How does the increase in heart beat help Jabulile during the race? …………………………………………………….………….……..... (3)

4. Many schools that use iron sheets for roofing their buildings paint the roofs. Why do you think the iron sheets-roofs are painted? ……………………………….… (1)

Explain your answer…………………………………………………..………………. (2)

5. The diagram below shows Themba, his two sisters Sonile and Vuy'sile, sitting around a **small** brazier *(immbawula*) outside their small one-room-house with a small window. They are talking to one another as they warm themselves.



a) What **two** substances are necessary for the brazier to provide warmth?

…………………………………………………………………………….…..……………... (2)

b) Give **two** reasons why Sonile feels taking the brazier inside the house is a dangerous idea?

Explain your answer. ………………………………………………………………… (4)

c) In putting out the fire Themba put the brazier inside a 25 L metal container and closed it tightly. Explain how their action helped put out the fire. ……………..………...… (2)

Total 35 Marks

>>>

Pupil Questionnaire

(Writing space was provided)

Please answer the following questions to express your feelings about the test.

1. Which questions of the test do you like?
2. What do you like about these questions?
3. Which questions do you **not** like?
4. Why do you **not** like these questions?
5. Describe other feelings or ideas you have about this test.
6. If you were told that all topic tests in science would ask questions that use situations you see around you as was done in this test, what would you say?
   Explain your answer.

APPENDIX IV

Letter to the Head Teachers and Ministry of Education

UNIVERSITY OF SWAZILAND
DEPARTMENT OF CURRICULUM AND TEACHING
P/BAG 4, Kwaluseni,
Tel: (+268) 5184011
Fax: (+268)5185276
E-mail: victoria@educ.uniswa.sz

25 February 2004

The Head Teacher
………………….
………………….
…………………..

Dear Head Teacher

Re-Study on the use of alternative assessment strategies for contextualised science teaching approach in pursuit of Ph. D. degree

I wish to request permission to carry out an investigation with one Form II science class on the use of alternative strategies of assessing students' learning of science concepts and skills when following a contextualised science (application in everyday life) teaching approach.

In order to carry out the study students will need to be taught using a contextualised approach. Contextualised curriculum materials for teaching science at Form II level have been developed and used as an alternative or substitute for the current Integrated Science Programme material. The use of a contextualised science teaching approach has been shown not to disadvantage students' learning of science.

The study will focus on the following topics: Electricity, Air and Life, Support and Movement, and Acids and Hydrogen. Student learning will then be assessed using contextualised assessment items as well as performance tasks to ascertain the viability of using such assessments for learning science in a contextualised way.

I will provide the curriculum material in the form of students' workbooks and teachers' guides for the four topics.

Please note that, from time to time, the participating teacher will need to attend workshops to discuss the teaching approach used, the use of the assessment tasks and their experiences and progress.

This study is in line with current science curriculum reform efforts going on in Swaziland. The curriculum being developed for science has a strong emphasis on contextualising science teaching. The study will thus allow teachers to practice the teaching approach and become familiar with the characteristics of contextualised science teaching. The teachers will also get experience in integrating assessment into learning, by implementing assessment tasks as part of the learning process.

Thank you.

Yours sincerely

Victoria Kelly (Ms)

APPENDIX V

Letter from the Ministry of Education

MINISTRY OF EDUCATION

Kingdom of Swaziland

Inspectorate Division

25 March 2004

Ms Victoria Kelly
UNISWA
Private Bag 4
Kwaluseni

Dear Madam

<u>**Re-Request for Permission to Carry Out a Research Study in Secondary Schools**</u>

I refer to your letter dated [illegible] February, 2004 on the above request.

I am pleased to inform you that your request was carefully considered in your favour. You are therefore granted the permission you have requested for.

We wish you well in your endeavour and hope that you will share the findings of your study in the interest of the education system.

Yours Sincerely

B. S. Ndlovu
For: Principal Secretary

MINISTRY OF EDUCATION
BOX 39
25 MAR 2004
MBABANE
SWAZILAND

APPENDIX VI

Teacher workshop invitation and programme

UNIVERSITY OF SWAZILAND
DEPARTMENT OF CURRICULUM AND TEACHING
P/BAG 4, Kwaluseni,
Tel: (+268) 5184011 Ext 2191
Fax: (+268)5185276
E-mail: victoria@educ.uniswa.sz
Cell 608 4309/ Home 505 7740

March 22, 2004

Dear Participating Teacher

Re-Workshop on context-based teaching to be held on Friday 26 and Saturday 27, March 2004.

As you have agreed to participate in the study on “Alternative assessment approaches within a context-based teaching approach”, I am kindly requesting you to attend the first induction workshop. This is a two-day workshop. The first day will be dedicated to the orientation and familiarisation with the teaching material namely the unit on Electricity. The programme for this day is attached.

Day two will be on the use of the alternative assessment approaches.

We shall have the workshop on the use of the Electricity material at the **In-service Department laboratories on Friday 26 March 2004 beginning at 8.30 am.**

You are kindly requested to carry out the TASK outlined on the attached sheet in preparation for day one of the workshop.

Day two (i.e. Saturday 27 March 2004) will focus of the assessment approach to be used during the study. It will take place **at School M2 Science Laboratory**. Arrangements are being made to have students taking the test while we observe and assess them. The workshop will begin at 08.30 to allow us to discuss the assessment approach before students arrive.

Please bring your teacher guides on Electricity to the workshop. Other workshop material will be provided. Tea and lunch will also be provided. Your travel will be reimbursed.

Thank you for your cooperation.

Yours sincerely

Victoria Kelly

The Programme for Friday 26 March 2004 stands as follows

08.30 Registration

08.45 Introduction

09.00 Discussion of Workshop Preparatory Task 1

09.30 Demonstration-teaching and discussion of a contextualised lesson EL6

10.15 Tea

10.30 Documentation of Lessons

11.00 Discussion of different contextualised lessons (EL1, EL2, EL3)

11.30 Demonstration-teaching and discussion of investigative lesson (EL7

12.15 Lunch

13.15 Peer-teaching and discussion of contextualised lesson EL8 and EL9

14.15 Group practice and discussion of investigative lesson EL10

15.30 Review and Summary of Workshop and follow-up activities

16.00 Closure

Workshop Preparatory Task 1

Before the workshop you will have read the lessons material for EL6: Mr Hlophe's Car Lights and the SWISP Activity 7.4: Conductors and Non-conductors.

Consider and discuss you opinions of the following:

- What is similar about the **materials?**
- What is different about the **materials?**
- What is similar about the **learning activities** suggested?
- What is different about the **learning activities** suggested?
- What is similar about the **teaching approach** suggested?
- What is different about the **teaching approach** suggested?
- What is similar about the **science idea**s pupils learn about?
- What is different about the **science ideas** pupils learn about?

APPENDIX VIIA

Student interview schedule/ Questionnaire

Dear student,

We have worked together a few times on your science work and I would like to find out what your ideas are about how we have been doing things.

Work in your groups and discuss your group ideas about the questions below and then write down your responses/answers on the sheet of paper given. Please allow one another a chance to say something.

1. a What do you think is good about using practical tasks to test you on what you know and can do in science?
   b Which part of the practical tasks do you like? (Remember each task had three parts: The planning, the testing, answering questions). Please explain your answer.

   c. What did you like about it?

2. How did the practical task benefit/help you in your learning of science?
3. What did you **not like** about the practical tasks?
4. a What kind of preparation (teaching or studying) does a student need to be able to write the practical tasks?
   b How ready were you when you wrote each of the three practical tasks (two on electricity and one on air and life)?

5. For the topic on Electricity you wrote two practical tasks, a topic test (and possibly some end of lesson tasks), and for Air and Living Things you wrote one practical task, a topic test and end of lesson tasks. All these pieces of work were marked and marks were recorded. How do you feel about taking several/many tasks as tests and work to be marked for each topic?
6. You worked on the practical tasks in groups. How do you feel about working on tasks as a group?
7. What would you recommend or suggest about the future use each of the following type of tasks to check students' understanding of science:
   a. practical tasks
   b. end of lesson exercises?
8. What else would you like to say about the way you have been learning and tested in Electricity and Air and Life?

Thank you

Ms V. Kelly

APPENDIX VIIB

Teacher interview schedules

Teachers' Reflection Guide

Dear Participating Teachers,

This is the final lap of the research work we have been working on together. Thank you for your continued support and dedication to the project. We need to meet and discuss **our experiences, observations and ideas about the use of performance tasks, context-based test items and end of lesson exercises as some of the other ways teachers can use to assess students' learning.**
You will remember, at the beginning of the project you were requested to use the guidelines below in box to keep a diary/journal of your experiences. We will discuss these experiences during the meeting.

Dear Participant Teacher,

Please keep a journal of your experiences while teaching using these materials and tests to help evaluate the use of the contextualised materials and the assessment strategies employed. This would help identify the strengths and weakness of implementing these curriculum approaches as well any requirements that may need to be met if such approaches are to be nationwide.

Possible Journal entries are:
Describing how you felt as you used the contextualised materials and the assessment questions;
How successful your lessons are, why they are successful or not successful;
What difficulties you had with the materials and/or the assessment questions
Your opinions of the teaching approach regarding learning of concepts; development of skills, usability and suitability of contextualised teaching approach for teaching science at secondary school, and context-based, embedded and performance assessment approaches; problems encountered, support requirements/needs
School readiness for classroom assessment that leads to end of term/year grades as suggested by the assessment approach used. What changes, if any, would be necessary to be able to efficiently use the assessment approaches studied;
Any other idea that you consider important (no matter how great or small the importance is).

In addition to discussing the above experiences we shall also use the following questions to guide our discussion:

1. What did you find good, beneficial, and advantageous, when using end of **lesson activities (which you collected and marked), performance tasks (practical test) and the context-based topic tests** for assessing learning within a context-based teaching approach. Please look at these in terms of both the teacher and the students?
2. What are the weaknesses of each of the approaches listed in Question 1 for assessing students learning when following the context-based teaching approach?
3. Has the use of these **assessment approaches** affected students' learning and/ or performance? Please explain.
4. Has the use of these **assessment approaches** affected your teaching of science? Please explain.
5. What requirements should schools and teachers meet to be able to use the three assessment approaches listed in Q1 above?
6. What kind of assistance (if any) does a teacher need when using these assessment approaches?
7. What are your views about the future use of the approaches listed above to assess students' learning?

8. What other feelings or views do you have about the use of the three approaches listed above to assess students' learning?

It might benefit the discussion if you could spend a little of your time to jot down responses to these questions. I will need to collect these responses to help when transcribing the discussion.
Thank you very much.
Victoria Kelly

Teacher Interview II Schedule Context based assessment

I realise that we did not discuss context-based tests and questions adequately during the group discussion. I wish to request for your time to talk about it again, but focus on both the teacher and students.

Questions to guide our thinking are:

1. What is good about the context-based questions/test for assessing students learning? (*Advantages, strengths*)
2. What is not good? (Disadvantages, challenges, weaknesses)
3. Did you identify any problems for students in answering/responding to the questions? (*Nature of problems)*
4. How did you deal with these problems? *(in your marking, giving grades)*
5. Did you experience any challenges while marking students' answers to the items? Explain.
6. What do you think about using contextualised questions in future tests?

Also explore

-what role do you see students' everyday experiences playing in student' answers to questions?
- how adequate are context based questions in assessing science learning? (Appropriateness of questions for science assessment)
- what is your position of expertise in using context based tests?

- what benefits, if any, do you think CBT have to students?

- what are your main concerns about the use of context based questions in science tests?

-what would you suggest for future use of context based questions in tests?

ANY OTHER POINTS YOU WISH TO RAISE?

APPENDIX VIII

Inter-coder reliability check

An Example of how inter-coder reliability was determined for the questionnaires. Inter-coder reliability checks for Questionnaire Question 8. The codes used by the coders are the same as those given in section 3.13.1 above.

Code notations were as follows: A= agreement between raters; dA= disagreement between raters; ADSE= affective disposition self-efficacy; ADME= affective disposition extrinsic motivation; ADSE= affective disposition intrinsic motivation; Met = metacognition; Patres= task resources; Recom = recommendation; TCIV= Task importance and value; TCC = task complexity; TCREQ1= pre-task requirements; TCREQ2= in-task requirements

School & group no.	Response	Coder 1	Coder 2	Coder3	My codes
8 M1 g2	We have gained more knowledge such that we are now able to answer some questions without studying	Tciv*	Tciv*	Tcc#	Tciv*-adse#
8 M2 3A g1	This two topics were helping us to know what is around us and how is made out of, and how it is good to us.	Tciv*	Tciv*	Met#	Tciv*
8 M2 3B g4	The way we do the practical test we able to understand better	Tciv#	Tcc*	Tciv#	Tcc*
8 M3 3A g1	i) It was interesting because Electricity and Air and Life are the things we observe in our environment.	Tciv*	Tcc#	Tciv*	Admi# Tciv*
8 M3 3A g1	ii) The lessons are too long, but we still benefit a lot.	Adme#	Tcc#	Recom#	Patres-t#; tciv#
8 M33A g4	Nothing	-	-	-	-
8 M3 3B g4	(i) We should continue to learn using this method.	Recm*	-	Recm*	Recom*-cont
8 M3 3B g6	(i) They should be applied in all schools so that all students will be something in future, do not rely on other people for knowledge. E.g. doctors. Every student should be educated for the benefit of her own country.	Tciv*	Tciv*	Recom*	Recom* Tciv*;
8 M4 3B g7	It is challenging because you have to be observant about what is happening around and what you did for us was good and it will help us in future.	Tciv#	Tciv#	Recom#	Tc-req2#
Number of agreeing and number of disagreeing codes and percentage agreement.		5A:3dA (63%)	4A:3dA (57%)	3A:5dA (38%)	7A:4dA (58%)

* implies agreement between my codes any one coder or all coders; # implies disagreement between any of my codes and those of the other raters; **-** UN means uncoded

Summary of inter-rater codes for student interview schedule/questionnaire.

Qn	Coder 1	% A	Coder 2	% A	Coder 3	% A
Q1	22A:9dA	71	16A:12dA	57	18A:12dA	60
Q2	7A:2dA	78	7A:3dA	70	3A:4dA	43
Q3	11A:5dA (2UC)	69 (61+)	9A:3dA (3UC)	75 (60+)	10A:3dA :2 UC	77 (67+)
Q4	5A: 0dA: (8UC)	100 (38+)	5A:1dA: (7 UC)	83 (39+	5A:9dA: (1 UC)	36 (33)
Q5	4A:3dA	57	5A:1dA: (1UC)	83 (71+)	2A:5dA	29
Q6	12A:2dA: (1UC)	86 (80+)	12A:0dA: (1 UC)	100 (92+)	11A:2dA	85
Q7	13A:2dA: (1UC)	87 (81+)	12A:4dA: (1UC)	100 (92+)	9A:7dA:1 UC	56 (53)
Q8	5A:3dA	63	4A:3dA (2 UC)	57 (44+)	3A:5dA:1 UC	38 (33+)
All Qns	**79A: 26dA: (12UC)**	**75 (68)**	**73A:21dA: (15 UC)**	**78 (70)**	**59A:47dA (4UC)**	**56 (54)**

A=agree; dA=disagree; UC=uncoded; UM=unmatched; no corresponding code in other coders;

+ implies percentage agreement based on all responses including coded/uncoded /unmatched codes. (uncoded responses were counted as one code for calculation purposes).

Calculation of percent agreement between codes was performed thus:

$$\frac{\textit{Number of codes that agree}}{\textit{Number of codes that agree and disagree (and unmatched codes)}} \times 100$$

APPENDIX VIII

Inter-coder reliability check

An Example of how inter-coder reliability was determined for the questionnaires. Inter-coder reliability checks for Questionnaire Question 8. The codes used by the coders are the same as those given in section 3.13.1 above.

Code notations were as follows: A= agreement between raters; dA= disagreement between raters; ADSE= affective disposition self-efficacy; ADME= affective disposition extrinsic motivation; ADSE= affective disposition intrinsic motivation; Met = metacognition; Patres= task resources; Recom = recommendation; TCIV= Task importance and value; TCC = task complexity; TCREQ1= pre-task requirements; TCREQ2= in-task requirements

School & group no.	Response	Coder 1	Coder 2	Coder3	My codes
8 M1 g2	We have gained more knowledge such that we are now able to answer some questions without studying	Tciv*	Tciv*	Tcc#	Tciv*-adse#
8 M2 3A g1	This two topics were helping us to know what is around us and how is made out of, and how it is good to us.	Tciv*	Tciv*	Met#	Tciv*
8 M2 3B g4	The way we do the practical test we able to understand better	Tciv#	Tcc*	Tciv#	Tcc*
8 M3 3A g1	i) It was interesting because Electricity and Air and Life are the things we observe in our environment.	Tciv*	Tcc#	Tciv*	Admi# Tciv*
8 M3 3A g1	ii) The lessons are too long, but we still benefit a lot.	Adme#	Tcc#	Recom#	Patres-t#; tciv#
8 M33A g4	Nothing	-	-	-	-
8 M3 3B g4	(i) We should continue to learn using this method.	Recm*	-	Recm*	Recom*-cont
8 M3 3B g6	(i) They should be applied in all schools so that all students will be something in future, do not rely on other people for knowledge. E.g. doctors. Every student should be educated for the benefit of her own country.	Tciv*	Tciv*	Recom*	Recom* Tciv*;
8 M4 3B g7	It is challenging because you have to be observant about what is happening around and what you did for us was good and it will help us in future.	Tciv#	Tciv#	Recom#	Tc-req2#
Number of agreeing and number of disagreeing codes and percentage agreement.		5A:3dA (63%)	4A:3dA (57%)	3A:5dA (38%)	7A:4dA (58%)

* implies agreement between my codes any one coder or all coders; # implies disagreement between any of my codes and those of the other raters; - UN means uncoded