

EduP-53



**Investigating and improving assessment practices in Physics in
secondary schools in Mozambique**

By Francisco Maria Januário
22381806

Submitted in fulfilment of the requirements for the degree of
PhD: Assessment and Quality Assurance

In the Department of Curriculum Studies
Faculty of Education
University of Pretoria
PRETORIA

July 2008

Supervisor: Professor Sarah Howie, University of Pretoria
Co-supervisors: Professor Emeritus Tjeerd Plomp, University of Twente
Professor Inocente Mutimucuo, Eduardo Mondlane University

TABLE OF CONTENTS

| | |
|--|------|
| TABLE OF CONTENTS..... | i |
| List of Tables | iii |
| List of Figures..... | iv |
| List of Acronyms | v |
| Declaration..... | vii |
| Summary | viii |
| Dedication..... | xiii |
| Acknowledgements..... | xiv |
| | |
| CHAPTER 1 | 1 |
| INTRODUCTION | 1 |
| 1.1 Introduction of the research problem..... | 1 |
| 1.2 Research problem and aims | 4 |
| 1.3 Research approach | 7 |
| 1.4 Overview of the following chapters..... | 13 |
| | |
| CHAPTER 2 | 13 |
| THE CONTEXT OF THE STUDY | 13 |
| 2.1 Country background information..... | 13 |
| 2.2 Educational context..... | 15 |
| 2.3 Conceptualisation and rationale of the problem | 23 |
| 2.4 Importance of the study | 27 |
| | |
| CHAPTER 3 | 29 |
| LITERATURE REVIEW AND CONCEPTUALISATION OF THE STUDY | 29 |
| 3.1 Introduction..... | 29 |
| 3.2 Definition of terms..... | 33 |
| 3.3 Theoretical orientation in assessment | 39 |
| 3.4 Assessment strategies in Physics | 44 |
| 3.5 Some intervention studies in science education | 54 |
| 3.6 Summary and conceptualisation of the study | 59 |
| | |
| CHAPTER 4 | 66 |
| RESEARCH DESIGN AND METHODS | 66 |
| 4.1 Introduction..... | 66 |
| 4.2 Research paradigm..... | 69 |
| 4.3 Overview of the research design..... | 70 |
| 4.4 Validity and reliability | 119 |
| 4.5 Ethical issues..... | 120 |
| 4.6 Conclusion | 121 |

| | |
|---|---------|
| CHAPTER 5 | 123 |
| ASSESSMENT PRACTICES OF MOZAMBICAN PHYSICS TEACHERS | 123 |
| 5.1 Introduction..... | 123 |
| 5.2 Findings of the Baseline Survey | 125 |
| 5.3 Conclusions..... | 142 |
| CHAPTER 6 | 148 |
| IMPROVING TEACHER ASSESSMENT PRACTICES IN PHYSICS IN MOZAMBIQUE..... | 148 |
| 6.1 Introduction..... | 149 |
| 6.2 Design and evaluation of the first prototype..... | 151 |
| 6.3 Design and evaluation of the second prototype | 157 |
| 6.4 Design of the third prototype | 191 |
| 6.5 Final appraisal resulting in the fourth prototype..... | 194 |
| 6.6 Conclusions and implications for further development..... | 202 |
| CHAPTER 7 | 208 |
| CONCLUSIONS AND RECOMMENDATIONS | 208 |
| 7.1 Summary of research questions and approach..... | 208 |
| 7.2 Reflections of the study | 220 |
| 7.3 Conclusions of the study..... | 226 |
| 7.4 Recommendations..... | 231 |
| REFERENCES | 237 |
| APPENDICES | 250 |
| Appendix A: Questionnaire for teachers and school directors | 250 |
| Appendix B: Classroom observation schedule | 255 |
| Appendix C: Interview schedule for teachers..... | 263 |
| Appendix D: Interview schedule for school directors | 265 |
| Appendix E: Interview schedule for pedagogical officers..... | 266 |
| Appendix F: Guide for expert appraisal | 268 |
| Appendix G: University students' questionnaire (before tryout) | 269 |
| Appendix H: University student's follow-up interview (before tryout)..... | 270 |
| Appendix I: Teacher's evaluation questionnaire (after tryout)..... | 271 |
| Appendix J: Follow-up interview with teachers (after tryout) | 274 |
| Appendix K: Students' questionnaire (after tryout) | 276 |
| Appendix L: Letter from the Ministry of Education and Culture to schools..... | 278 |
| Appendix M: Letter from Eduardo Mondlane University to schools..... | 279 |
| Appendix N: Letter of consent from the researcher to teachers | 280 |
| Appendix O: Ethical clearance | 283 |
| Appendix P: Physics Assessment Materials (on a CD - Rom)..... | 284 |
| Appendix Q: Certificate of language editing..... | 284 |

List of Tables

| | |
|---|-----|
| Table 2.1: Mozambique profile | 15 |
| Table 2.2: Formulae for calculating student marks | 18 |
| Table 3.1: Database search engines | 32 |
| Table 3.2: Components of a rubric | 46 |
| Table 4.1: Sample of Baseline Survey..... | 79 |
| Table 4.2: Data collection matrix | 85 |
| Table 4.3: Quality criteria related to phases in educational design research..... | 92 |
| Table 4.4: Curriculum components | 101 |
| Table 5.1: Consistency level of teachers' responses between questionnaire and interview (n=12)..... | 130 |
| Table 6.1: Design specifications of Physics assessment materials..... | 153 |
| Table 6.2: Summary of lessons for the second prototype..... | 159 |
| Table 6.3: Overview of the sub-questions for the evaluation design..... | 165 |
| Table 6.4: Characteristics of participants of the classroom tryout | 167 |
| Table 6.5: Teachers' general impression about the second prototype of the PAM materials..... | 173 |
| Table 6.6: The five levels of the Guskey (2000) model as applied to this study..... | 196 |

List of Figures

| | |
|--|-----|
| Figure 1.1: Steps towards achieving a research goal | 9 |
| Figure 1.2: Research model of the study | 11 |
| Figure 2.1: Internal borders of Mozambique and its provinces | 14 |
| Figure 2.2: Elements of the school environment | 23 |
| Figure 2.3: Components of the study..... | 25 |
| Figure 4.1: The cyclical process of educational design research..... | 90 |
| Figure 4.2: The original model by Mafumiko (2006)..... | 94 |
| Figure 4.3: General research design of the study..... | 95 |
| Figure 4.4: Research model of the Intervention Study | 98 |
| Figure 4.5: The vulnerable curriculum spider web..... | 101 |
| Figure 4.6: Assessment components..... | 102 |
| Figure 4.7: A sphere set in motion..... | 108 |
| Figure 4.8: Forces on a moving object | 109 |
| Figure 4.9: A soccer ball kicked into the air..... | 111 |
| Figure 4.10: Identification and comparison of forces acting on moving objects | 112 |
| Figure 5.1: Teachers' responses by type of assessment practice (n=12)..... | 126 |
| Figure 5.1: Teachers' responses by type of assessment practice (n=12)..... | 126 |
| Figure 5.2: Teachers' responses on the assessment of students' activities (n=12)..... | 132 |
| Figure 5.3: Teachers' responses on students' evaluation of their work (n=12)..... | 136 |
| Figure 5.4: Teachers' responses on the use of assessment results (n=12)..... | 137 |
| Figure 6.1: Execution of the lesson: a practice-oriented lesson plan for the teacher | 160 |
| Figure 6.2: Worksheet for students..... | 163 |

List of Acronyms

- ACP – *Actividades de Controle Parcial*
Activities of Partial Control
- ACS – *Actividades de Controle Sistemático*
Activities of Systematic Control
- ADDIE – Analysis-Design-Development-Intervention-Evaluation
- CNECE – *Concelho Nacional de Exames, Certificação e Equivalências*
National Council of Exams, Certification and Equivalences
- DINEG - *Direcção Nacional do Ensino Geral*
National Directorate of General Education
- EP1 – *Ensino/Escola Primário(a) do 1 Grau*
Lower Primary Education/School
- EP2 – *Ensino/Escola Primário(a) do 2 Grau*
Upper Primary Education/School
- ESG – *Ensino Secundário Geral*
General Secondary Education
- ESG1 – *Ensino Secundário Geral do 1 Ciclo*
General Secondary Education, Cycle 1
- ESG2 – *Ensino Secundário Geral do 2 Ciclo*
General Secondary Education, Cycle 2
- INDE – *Instituto Nacional de Desenvolvimento da Educação*
National Institute for Development of Education
- IPO – Input-Process-Output
- MEC – Ministry of Education and Culture
- MinEd – *Ministério da Educação*
Ministry of Education
- NGO – Non Government Organisation
- PAM – Physics Assessment Materials

POE – Predict-Observe-Explain

UNDP – United Nations Development Programme

SNE – *Sistema Nacional de Educação*
National System of Education

UEM – *Universidade Eduardo Mondlane*
Eduardo Mondlane University

UP – *Universidade Pedagógica*
Pedagogic University

Declaration

I declare that this thesis is my own, unaided work. It is being submitted for the Degree of Doctor of Philosophy in the University of Pretoria, South Africa. It has not been submitted before for any degree or examination in any other University.

Francisco Maria Januário

July 2008

Summary

Assessment, an integral part of teaching and learning, is a planned process of identifying, gathering and interpreting information about the performance of students. However, concerns have been raised about how assessment is being conducted in schools and so the aim of this study is to investigate and improve assessment practices used by Grade 12 secondary school Physics teachers in Mozambique, Africa. The study addresses the question of *what assessment practices do Grade 12 teachers in Physics in Mozambique apply and how can they be improved?* and it adopted a twofold research approach. A Baseline Survey aimed at gaining an overall impression of the assessment practices used by secondary school Physics teachers and an Intervention Study aimed at producing improvements on teacher assessment practices. The preliminary research followed a survey research method, while the intervention applied an educational design research approach. In the survey three questions were investigated: (i) *What assessments practices do Grade 12 Physics teachers apply?* (ii) *What is the quality of the assessment practices?* and (iii) *How relevant can the assessment practices be for student learning?* To address these questions a purposive sample of 12 Physics teachers, four school directors and five educational officers was selected. The survey was conducted in six secondary schools purposefully selected throughout the country and data were collected via interviews, questionnaires, classroom observations and written notes. The Intervention Study was designed to answer the question of *how teacher assessment practices can be improved.* This phase of the study involved a design, a classroom tryout, and a systematic evaluation of a series of exemplary Physics assessment materials (prototypes) in a context of demonstration experiments. The prototypes were developed for the concepts of force and inertia and their validity and practicality were verified using appraisal by experts, university students, teachers, and students. Classroom tryout was conducted with two teachers and their 62 students in two secondary schools.

Baseline Survey findings indicate that the most used assessment practices in schools are paper-and-pencil tests, verbal tests, and homework, while projects, portfolios, and peer-

assessments are the less frequently used ones. Oral communication during lessons, written work, presentations, notebooks, laboratory work, and ability to solve problems were used as quality criteria for the teachers' assessment. It was shown that the most frequently assessed student activity was written work, followed by the ability of students to solve problems, while laboratory work was the activity that was never assessed by many of the researched teachers. Another quality criterion used was the type of feedback given by teachers to students, which indicated that teachers were giving expressed (both congratulatory and critical), personal and timely feedback. It emerged that teachers often involve the students in the evaluation of their performance through reflection of assessment results and in addition, encourage students to engage in active learning.

Findings from the Intervention Study indicate that (i) teachers liked the presentation and structure of the materials following the Predict-Observe-Explain (POE) strategy and regarded their personal commitment as crucial for achieving the desired experimental results; (ii) students also liked most the POE strategy because it allowed them to develop their own explanations of the observed events and highlighted the role of teachers during the tryout as crucial for the success of the experiments.

The main conclusion of this study is that assessment practices undertaken by Physics teachers in Mozambican secondary schools are of poor quality and there is a need for improvement. This must be done by developing and applying exemplary assessment materials with the potential to improve performance assessment practices associated with demonstration experiments in Physics. The study recommends that the Ministry of Education and Culture and teacher training institutions should promote the training of teachers in developing exemplary assessment materials for their own use in schools. These materials should contain specific guidelines on how to conduct effective assessment practices.

Key words: Assessment practices, demonstration experiments, force, formative assessment, formative feedback, inertia, performance assessment, Physics assessment materials, practical work, prototypes.

Resumo

A avaliação, uma parte integrante de ensino e aprendizagem, é um processo planificado de identificação, recolha e interpretação de informação acerca do desempenho dos alunos. No entanto, tem se levantado preocupações sobre a forma como a avaliação é levada a cabo nas escolas. Neste sentido, o objectivo deste estudo é investigar e procurar melhorar as práticas de avaliação usadas pelos professores de Física da 12ª classe no ensino secundário em Moçambique, África. O estudo aborda a questão sobre *que práticas de avaliação os professores de Física da 12ª classe em Moçambique usam e como elas podem ser melhoradas* e adopta uma abordagem de investigação dualista. Um Estudo de Base (pesquisa exploratória) destinado a obter uma impressão geral das práticas de avaliação usadas pelos professores de Física no ensino secundário e um Estudo de Intervenção com o intuito de produzir melhoramentos nessas práticas. A investigação preliminar obedeceu ao método de Inquérito, enquanto a intervenção empregou a abordagem de Pesquisa de Concepção Educacional – outrora conhecida por Pesquisa de Desenvolvimento. No Estudo de Base foram investigadas três perguntas: (i) *Quais são as práticas de avaliação usadas pelos professores de Física da 12ª classe?* (ii) *Qual é a qualidade dessas práticas de avaliação?* (iii) *Quão relevantes essas práticas de avaliação são para a aprendizagem dos alunos?* Para responder a estas perguntas foi seleccionada uma amostra por conveniência composta por 12 professores de Física, quatro directores de escola e cinco técnicos de educação. O Inquérito foi levado a cabo em seis escolas secundárias convenientemente seleccionadas ao longo de todo o país e os dados foram recolhidos por meio de entrevistas, questionários, observações na sala de aulas e notas escritas. O Estudo de Intervenção foi concebido para responder à pergunta sobre *como é que as práticas de avaliação dos professores podem ser melhoradas*. Esta fase do estudo envolveu a concepção, o ensaio na sala de aulas e a avaliação sistemática de uma série de exemplares de materiais de avaliação de Física (protótipos) no contexto de experiências de demonstração laboratoriais. Os protótipos foram desenvolvidos para os conceitos de força e inércia e a sua validade e praticabilidade foram verificadas com o auxílio de especialistas da área, estudantes universitários, professores e alunos do

ensino secundário. O ensaio na sala de aulas foi realizado com dois professores e seus 62 alunos em duas escolas secundárias.

Os resultados do Estudo de Base indicam que as práticas de avaliação mais usadas nas escolas são os testes de papel-e-lápis, as perguntas orais e o trabalho de casa, enquanto os projectos, os portfólios e a avaliação dos colegas são as menos frequentemente usadas. A comunicação oral durante as aulas, o trabalho escrito, as apresentações, as notas nos cadernos, os trabalhos laboratoriais e a habilidade dos alunos de resolver problemas foram usados como critérios de qualidade para a avaliação levada a cabo pelos professores. Emergiu que a actividade dos alunos frequentemente avaliada pelos professores era o trabalho escrito, seguida da habilidade dos alunos de resolver problemas, enquanto o trabalho laboratorial foi a actividade que nunca era avaliada pela maioria dos professores alvo da pesquisa. Outro critério de qualidade usado foi o tipo de retroalimentação (*feedback*) dado pelos professores aos alunos, o qual indicou que os professores davam uma retroalimentação expressa (certo ou errado), pessoal e em tempo. Os resultados mostraram igualmente que os professores, muitas vezes, envolvem os alunos na avaliação do desempenho deles através de discussões de reflexão acerca dos resultados das avaliações e encorajam os alunos a se empenharem na aprendizagem activa.

Os resultados do Estudo de Intervenção indicam que (i) os professores gostaram da apresentação e da estrutura dos protótipos seguindo a estratégia Previsão-Observação-Explicação (POE) e consideraram o seu cometimento pessoal como crucial para o alcance dos resultados experimentais desejados; (ii) os alunos também gostaram muito da estratégia POE porque lhes permitiu formular as suas próprias explicações dos eventos observados e referiram o papel do professor durante o ensaio como crucial para o sucesso das experiências.

A principal conclusão deste estudo é de que as práticas de avaliação levadas a cabo pelos professores de Física nas escolas secundárias Moçambicanas são de uma qualidade pobre e precisam de ser melhoradas. Isto deve ser dado pela concepção e uso de exemplares de

materiais de avaliação que tenham potencial para melhorar as práticas de avaliação de desempenho em associação às experiências de demonstração laboratoriais em Física. Este estudo recomenda que o Ministério de Educação e Cultura e as instituições de formação de professores promovam a formação e capacitação de professores em matérias de concepção de exemplares de materiais de avaliação para o seu próprio uso nas escolas. Estes materiais devem conter instruções específicas sobre como conduzir práticas de avaliação efectivas.

Palavras-chave: Avaliação de desempenho, avaliação formativa, experiências de demonstração laboratoriais, força, inércia, materiais de avaliação de Física, práticas de avaliação, protótipos, retroalimentação formativa, trabalho prático.

Dedication

To my parents Maria Francisca Semende (1935 – 1986) and Januário Julai for sending me to school; and my elder sister Vitorina Januário for carrying me and my school bag during my first year of schooling.

Acknowledgements

This dissertation is the result of a research started since the time the MHO projects of the Netherlands Government supported the re-opening of the Faculty of Education at Eduardo Mondlane University in Mozambique in 2002. Since then almost six years have elapsed and time has passed so swiftly that it seems like yesterday. Following a sandwich PhD study programme made me spend all these years living between Maputo and Pretoria and becoming a stranger to many people including my wife and homeboys. Although I take full responsibility for the contents of this dissertation, so many people were involved in it that I hardly afford to call it 'my research'. I would like to thank the people listed below.

- I thank my supervision team led by Professor Sarah Howie, which included Professor Emeritus Tjeerd Plomp and Professor Inocente Mutimucuo for their continued guidance and stimulating discussions along the journey. Eish... Professor Plomp, those track changes and accompanying memos of countless pages to 'keep the story line' are unforgettable!
- Thank you to my wife Eurídia E. de Azevedo for her love and strength to raise a naughty first-born, and to Vuyo himself - my son - who secures the continuing line of my genetic tree. These are the people who suffered most from my absence and ...uff... Vuyo, remember how you would cry profusely when I sometimes forgot to buy you a yellow Hummer car toy on my way home from Pretoria.
- To my colleagues from the Faculty of Education at UEM, Eugénia Cossa, Débora Nandja, Aguiar Baquete, Adalberto Alberto, Cristina Tembe, Carlos Lauchande, Xavier Muianga, and Ernesto Mandlate thanks for helping me keeping my mental sanity particularly during the last six months of writing this dissertation.
- To Professor Mouzinho Mário, Dr Arlindo Siteo, and Viriato Chevane for helping me walk the PhD path; your support and encouragement cannot be quantified. Professor Mário and Viriato Chevane, your effort in proofreading the 'dirty' first complete draft without a reward has paid off.

- I thank my colleagues at the Centre for Evaluation and Assessment, Faculty of Education, at the University of Pretoria (Groenklof Campus), Vanessa Scherman, Elsie Venter, Zélda Snyman, Zanele Matlou, Surette van Staden, Ilse Rautenbach, Lisa Zimmerman, Ronald Pillay, Kim Draper, Lilian Tabane, and Tina Lopes for your endless support. Your 'Ubuntu' and generosity during the coffee breaks helped me to unwind from the paranoia of writing. 'Ngiyabonga bafowethu'.

Thank you also to:

- Professor William Fraser for believing in me and for giving me words of encouragement;
- Wout Ottevanger, Marinus Kool and Wim Kouwenhoven for their assistance and friendship, and for encouraging me to focus more on my human potential;
- NUFFIC for their financial support without which this study project would not have taken place;
- Jeannie Beukes from the Faculty of Education Administration for providing substantial support particularly during the last month of submitting the dissertation and during the examination process;
- Bernice Mcneil and Cilla Nel for assisting me with the language editing; you are the persons who really suffered with my broken English and I thank you all deeply;
- officials from the Ministry of Education and Culture in Mozambique for accepting to participate in this study; and
- Finally, but not least, the Boards of the schools, teachers, and students who participated in this study for allowing me to interfere with their school's teaching timetables.

CHAPTER 1

INTRODUCTION

CHAPTER 1

INTRODUCTION

This chapter introduces the study on investigating and improving assessment practices in Physics in secondary schools in Mozambique. In Section 1.1 the research problem is introduced with a focus on the structure of the Mozambique educational system and the problematic situation connected with secondary education and qualifications of secondary school teachers. Section 1.2 presents the research problem and aims of the study leading to the main research question of the study. A more detailed elaboration of this question follows in the subsequent chapters. The research approach of the study is discussed in Section 1.3 where the pragmatic knowledge claim is stated as the scientific basis for the study. The schematic representation of the steps to be followed in order to answer the main research question is also presented in this section as a research model of the study. The model is elaborated into four main components namely, research goal, research object, research perspective and research activities. Finally, Section 1.4 presents an overview of the chapters.

1.1 Introduction of the research problem

After many years of colonisation, and in line with many other countries, Mozambique has implemented its own National Education System. The Mozambique National System of Education (SNE) comprises ten subsystems, namely: general education, adult education, technical and vocational education, teacher education, higher education, special education, distance education, school health and sanitation, sports, and school production and food. There are four levels of education: primary, secondary, intermediate and higher education. The subsystem of general education is composed of primary education (lower and upper) and secondary education (cycles 1 and 2). General secondary education, cycle 1 (ESG1),

comprises Grades 8, 9 and 10, and the general secondary education, cycle 2 (ESG2), comprises Grades 11 and 12. In cycle 1 the students have ten subjects namely Biology, Chemistry, Drawing, English, Geography, History, Mathematics, Physics, Physical Education, and Portuguese. In cycle 2 French and Philosophy are added, while students can choose one of the three existing streams, depending on what they would like to pursue at university level: sciences, humanities, or social sciences (MinEd, 2004). The purpose of this study is to investigate and improve assessment practices used by secondary school Physics teachers in Mozambique and will focus on the Science stream of ESG2. In 2006 the enrolment figures for general secondary education (ESG) were 293,179 students, which can be broken down into 257,729 students enrolled for ESG1 and 35,450 students enrolled for ESG2. The public school network under the Ministry of Education and Culture (MEC) was composed of 265 schools of ESG. Of these, 216 were of ESG1 and 49 schools of ESG2 (MEC, 2007). The ESG2 schools have the designation of *Complete Secondary Schools* because they offer all grades of secondary education, i.e., from Grades 8 to 12. The efficiency of the education system in general is extremely low and is characterised by high repetition and dropout rates. The system is also affected by significant disparities in gender and equity (MEC, 2007; Palme, 1992). For instance, with the exception of Maputo and surroundings, the enrolment rate for girls varies. In the transition from primary to secondary education and within the subsystem of secondary education (from ESG1 to ESG2) the system suffers a significant loss of girls. As per December 2006 the participation of girls in secondary education was estimated as 40.1% compared with 46.8% of primary. Within secondary education the figures of girls participation dropped from 42.2% in ESG1 to 38.9% in ESG2 (MEC, 2007).

In relation to teacher training the impact on the teaching profession has also been a matter of concern. A substantial percentage of graduates do not enter the teaching profession as they find jobs in the expanding market economy or in government ministries. From the total number of Pedagogical University (UP) secondary qualified teachers and those from other training institutions (mainly foreign) who are working in the public education sector, only about 30% are teaching in secondary schools (MinEd, 2004). Approximately 15% of those who enter UP are already serving as teachers. It is widely accepted that the four-year

course of *Licenciatura* is too long and academic for a qualification course. UP's strategic plan recognises that it will be unable to meet MEC projections for trained secondary teachers and that the overall shortfall will grow by about 60 teachers per year until 2008, when it will be 340 teachers. There are UP plans to increase the number of graduates in Mathematics, Portuguese, and Sciences, but this depends on sufficient qualified candidates with Grade 12 applying for the course. Although there are actually some promotion measures from MEC enabling some in-service teachers to enroll at UP without Grade 12, the number of the graduates for these fields is still insufficient.

As the number of schools expands, qualified teachers are becoming increasingly difficult to recruit. The number of unqualified teachers has increased more quickly and the number of qualified teachers has declined, especially outside Maputo. For example, in three schools in Tete Province with nearly 4500 students in total there are only three qualified teachers, while two schools in Maputo Province with nearly 2000 students have 35 qualified teachers. One in six teachers has no, or insufficient, teaching qualification. Men outnumber women in the secondary teaching profession by 5:1 and the increase in the numbers of unqualified teachers has maintained this imbalance. There are also differences in the availability of qualified teachers in different subjects. Estimates based on a representative sample of 1600 teachers suggest that less than 30% of teachers in English and Mathematics are qualified, with particular difficulties in recruiting teachers of Drawing. Female teachers are under-represented in the teaching of Biology, Chemistry, Mathematics, Physics and Drawing (37% of the total teachers is female).

In terms of retention of teachers, the situation can be characterised by a significant number of teachers who are leaving schools for better-paid employment (MEC, 2007). According to the Ministry of Education and Culture, a review of possible incentives is needed in the short term and should include: (i) improved living and working conditions, (ii) better contractual regulations, (iii) better career opportunities, (iv) access to credit, (v) accommodation close to the schools, and (vi) bursaries for their children. However, as teachers improve their educational background, they tend to find better paid work, thus teacher development has a potentially damaging side effect on the education system,

which needs to be countered in any review of the teacher development process. Some teachers, although qualified, remain on contract terms for up to two years waiting for their new position to be formally established. Many unqualified teachers are now being recruited on provisional contracts (seasonal teachers) after graduation from secondary school (end of Grade 12) and they need a system that will give them an opportunity to achieve the status of fully qualified teachers.

After this review of problems that the Mozambican education system, and especially secondary education, is confronted with, it is important to state that Mozambican teachers strive to meet the same learning outcomes for all students, at least for the secondary education graduates. One assumption underlying this research is that good assessment practices used by teachers in their classrooms can well serve as one of the means of realising these learning outcomes. It is believed that all teachers – whether they are qualified, under-qualified, unqualified, pre- and in-service teachers - can be made familiar with the characteristics and usefulness of formative assessment. The on-going review of the secondary education curriculum by the MEC has already identified the revision of the assessment practices as one of the strategic priority for the secondary education (MEC & INDE, 2007). The present study proposes an investigation aimed at improving teacher assessment practices at this level with the focus on Physics education in Grade 12. The research is of an exploratory nature involving a selected number of secondary school teachers.

1.2 Research problem and aims

All planned or unplanned assessments carried out by the teacher at school level, and the final examination undertaken by the MEC at a national level are governed by regulations set out in a document called "Regulation of Assessment for the General Secondary Education" published in the Bulletin of the Republic No. 35/96 of August 28, 1996 (*Governo de Moçambique, 1996*). This document presents an orientation on objectives, forms, frequency, methods as well as the approval criteria. Accordingly, the objectives of assessment are: (i) to verify the degree of the mastery of the curriculum objectives; (ii) to

contribute to the improvement of the teaching quality and to the evaluation of the teacher's work; and (iii) to verify the effectiveness of the methods and teaching/learning means.

A review of studies that took a close look at classroom assessment practices in general (albeit on primary level) indicates that there are some deviations from the application of the “Regulation of Assessment” document (INDE, 2005; Lauchande, 2001; MinEd, 1998; Palme, 1992; Popov, 1994). Specifically, the review revealed the following points:

- In general, the assessment is fundamentally based on the memorisation of concepts, formulas and mechanisation of procedures partly due to the teachers' weak scientific and pedagogic competence, and to their lack of skills in developing appropriate assessment instruments. This situation is more accentuated in experimental subjects, such as Physics, where teachers do not effectively assess the student abilities to manipulate, observe, generalise and establish relationships. This is due, on the one hand, to the teachers' weak preparation for assessing these abilities, and on the other hand, to the lack of teaching material and equipment such as microscopes and some other laboratory equipment.
- Students show great difficulties in providing correct answers to essay questions and their writing skills are below average.
- It is argued that the Regulation of Assessment is punishing students because it gives more emphasis on summative than formative assessment activities.
- The quality of teaching is very low. Teachers claim not to be involved in educational decision-making, and they have no opportunity to participate in regular upgrading courses.
- The teachers show lack of skills in designing and administering valid formative tests, i.e., tests that assess what is supposed to be assessed. There are also problems in how test items are formulated, especially those requiring analysis and comprehension levels of cognition.
- Tests and other assessment practices used by the teachers at the school level are not always in line with what is assessed by the national examinations.
- MEC does not always have access to reliable information about what really happens in classrooms.

All these problems relate to questions of how the intended curriculum – what teachers are expected to teach and students to learn - is implemented by teachers in their classrooms and is attained by students; how the learning process is monitored; and how students' achievement is assessed and certified in schools. In short, the problems relate to the need for improvement of formative and summative assessment in schools.

The argument for this study was that the improvement of teacher assessment strategies could help to monitor qualitative improvements of students' results and of the performance of the educational system as a whole. The central aim of this study was to investigate how the assessment practices of teachers can be improved, with a focus on Grade 12 Physics teachers. To address this aim, a twofold approach was applied. The core of the study was to investigate, through an intervention approach, the characteristics of effective assessment practices for Physics teachers. The intervention consisted of developing and trying out consecutive prototypes of assessment practices that Physics teachers may use in their classrooms. But, an important starting point for this was to have a good understanding of what was actually going on in the classroom. At the beginning of the research, a Baseline Survey was conducted in order to know what assessment practices Physics teachers were primarily undertaking.

So the main research question of this study was formulated as set out below.

What assessment practices do Grade 12 teachers in Physics in Mozambique apply and how can they be improved?

Available literature suggested that any improvement of assessment practices can only be an important means of improving teachers' work in the classroom if it is accompanied by changes in the instructional process (Airasian, 2001; Chatterji, 2003; Popham, 2002; van den Akker, 1999). For example, improved teachers' assessment strategies in Physics can use formative approaches as a way of improving student learning in the classroom. Assessment is important for the instructional process because (i) it determines whether

students are moving satisfactorily toward instructional outcomes that teachers are seeking to promote; and (ii) teachers can discern where to direct their instructional energies to ameliorate students' weaknesses and what already mastered skills or knowledge can be omitted from the lesson. Therefore, the importance of the study lies on the intervention phase where improved assessment strategies are developed, following both formative and summative assessment approaches.

Having addressed the main research question, the following section outlines the research approach of the study.

1.3 Research approach

The present study applied a twofold approach, namely a survey for the preliminary study, and an educational design research approach employed for the Intervention Study. A preliminary analysis of the Mozambican education system and a review of literature were undertaken leading to the conceptualisation of both the Baseline Survey and the Intervention Study. The findings of the Baseline Survey were used to further elaborate and refine the conceptual framework for the Intervention Study. The research designs for baseline and for intervention studies are elaborated on in Chapter 4. This subsection presents only a short characterisation of the two phases.

The Baseline Survey was aimed at getting an overall impression of the assessment practices used by Physics teachers in schools. A limited and purposeful sample of Grade 12 Physics teachers selected from different schools and provinces were considered to be sufficient to represent the different contexts (urban-rural, north-centre-south, etc.) of the country. This means that the survey sample was not based on a random selection from the population of all Grade 12 Physics teachers, but on a purposefully and carefully selected sample that provides for the indication of all aspects of Mozambican Grade 12 Physics education.

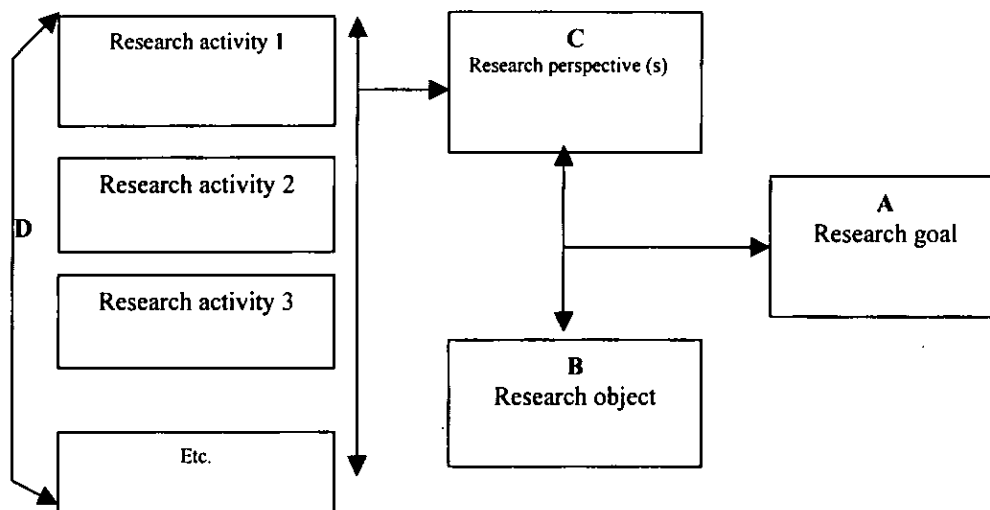
The Intervention Study aimed at producing improvements on teacher assessment practices. Its scientific position is rooted in the pragmatic knowledge claim (see Creswell,

2003). According to this claim, knowledge arises out of actions, situations, and consequences, where the main concern is with applications and solutions to problems. In principle, both qualitative and quantitative methods can be applied to collect and analyse data with the main aim of understanding the complexities of the current situation and to produce findings that contribute to the solution of the problem at stake. More importantly, the study approach was geared towards “what works” in schools and classrooms and how it will work on the basis of intended consequences. Within this framework, the Intervention Study applied an educational design research approach (see Bereiter, 2002; Plomp, 2006; Reeves, 2000; Richey et al., 2004; van den Akker, 1999). Educational design research in the context of this study is a research approach in which the search for characteristics of an effective intervention is conducted, while working on that intervention. Four phases can be distinguished in such an approach (Plomp, 2006): (i) preliminary research, (ii) prototyping phase, (iii) assessment phase, and (iv) systematic reflection and documentation.

For the purpose of this study, only two phases were considered namely: preliminary research and prototyping phase while systematic reflection and documentation took place throughout the study. The preliminary research phase of the Intervention Study built on the findings emerging from the literature review, document analysis, and the Baseline Survey leading to the conceptual framework and the operational research questions of the Intervention Study. The prototyping phase comprised iterative design with formative evaluation of several prototypes in a cyclical way applying a model with analysis-design-development-intervention-evaluation (ADDIE) elements. The emphasis of this phase was on refining and optimising the intervention by verifying whether the intervention met the prescribed design specifications. The systematic reflection and documentation phase portrayed the entire study (both baseline and intervention phases) in order to support a retrospective analysis and the specification of design principles. Specifically, all undertaken activities and the emerging findings of the two phases were used to draw inferences and to formulate design principles on how assessment practices in Mozambican Grade 12 Physics classrooms can be improved. The study did not consider the assessment phase that comprises a summative evaluation of whether the intervention works in

classrooms and with teachers who were not part of the prototyping phase. The reason for the exclusion of this phase is the limitation of time.

A research model adapted from Verschuren and Doorewaard (2003, taken from Plomp, 2004) influenced the structure of this study. The model is a schematic representation of the research goal and the general steps needed to achieve this goal (Figure 1.1). This model emphasises that a research framework can be developed step by step and it distinguishes four distinctive steps, which are worked through in reverse order whilst doing the research. The first step (A) is the summary of the research goal where, depending on the research question, new theories, principles and/or hypotheses are formulated (to be tested or developed in the research), or a problem context is diagnosed and an intervention is suggested. The second (B) and third (C) steps are the identification of the research object and research perspective respectively. The fourth step (D) corresponds to a number of research activities to be undertaken in order to investigate the problem and to generate a solution to the problem.



(Source: Plomp, 2004 adapted from Verschuren & Doorewaard, 2003)

Figure 1.1: Steps towards achieving a research goal

Although the original model of Verschuren and Doorewaard clearly depicts steps to be followed to achieve any research goals and the corresponding research activities, the

model indicates neither causality nor possible relationships between the elements in the model.

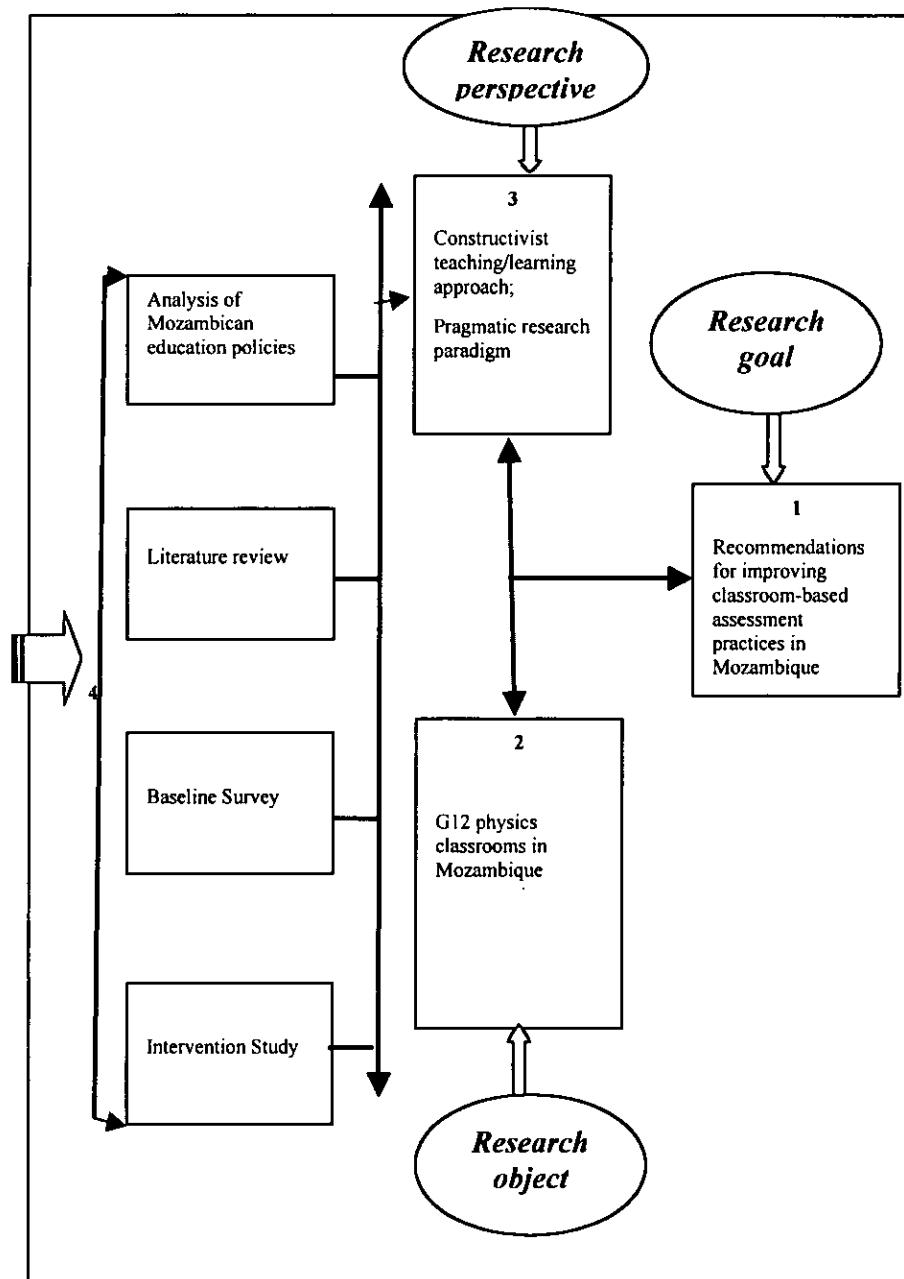


Figure 1.2: Research model of the study

The model for this study (Figure 1.2) presents the study process in terms of research goal, research object, research perspective, and context analysis and intervention. In the model, the research goal (1) is defined as the formulation of 'recommendations for improving classroom-based assessment practices in Mozambique'. It is assumed that undertaking an Intervention Study, aimed at helping teachers to design and carry out some selected performance assessment practices in Physics would improve their classroom practices in

general and would result in a number of design principles for designing such assessment practices. The research object (2) refers to the assessments practices that 'Grade 12 Physics teachers apply in the classroom'. The practices can be either formal or informal, with the focus on the formative type of assessments. As is the case with any other problem, to achieve relevant, suitable and applicable solutions to educational problems a triangulation is applied to ensure that the various problems' perspectives are seen from more than one angle at the same time. In Section 1.1 it was mentioned that the Mozambican education system, in general, is confronted with problems related to the lack of familiarity of both teachers and students with the characteristics and usefulness of formative assessment. To focus this study, Grade 12 and the subject Physics were chosen as the means of studying how to improve classroom assessment practices. This step is further elaborated in Chapter 2, where the context of Mozambican education is discussed. The research perspective (3) depicts the lens through which the researcher can look at the research object under investigation. In this study, this perspective refers to the 'constructivist approach of learning and teaching' as the context within which the classroom-based assessments are investigated by helping Grade 12 students to perform authentic tasks with the aim of solving real-life problems. This is in line with the Mozambique government policy for the revised curriculum that advocates a constructivist approach with strong emphasis on authentic assessment. The research perspective also includes the application of a pragmatic research approach, according to which the emphasis is on what works at the time of the intervention and how to research on the basis of intended outcomes. This step is further elaborated in Chapter 3 (Section 3.4) as findings from literature review on contextualising assessment practices in Physics. The context analysis and the Intervention Study (4) consist of a number of steps: the 'analysis of Mozambican educational policies' regarding assessment, the 'literature review', i.e., the summary of other research and arguments of other scholars on the issue, the 'Baseline Survey' of the actual status of classroom-based assessment in Grade 12 Physics in Mozambique, and the 'Intervention Study' aimed at improving teachers' assessments in schools. This step is further elaborated in Chapters 2 and 3 (document analysis and literature review) and in Chapters 5 and 6 (findings from baseline and intervention phases).

The research model in Figure 1.2 serves as an important conceptual guide for the steps necessary to follow in order to analyze assessment practices applied by secondary school Physics teachers in Mozambique, and the way to improve these. As the information is gathered and improvements are suggested, the model provides a guide to help in exploring the relationships between the learning and the assessment of Physics.

1.4 Overview of the following chapters

This dissertation is organised into seven chapters. This chapter provides an introduction to the problem and how to address this problem. The overview of the Mozambique context, and the problem viewed in its educational context particularly in relation to teacher qualifications are presented in Chapter 2. The review of the literature resulting in a conceptual framework for the research is set out in Chapter 3. This chapter includes reviews of the publications and research done internationally, which are used to support the choices of the research methods and to address what other scholars have written about the topic. Chapter 4 discusses the research design, and the research procedures chosen for the baseline and intervention phases of the study. The findings from the Baseline Survey are reported in Chapter 5. Chapter 6 reports on the outcomes of the Intervention Study. The development and try out of the various prototypes as well as their expected practicality and expected effectiveness in the classroom setting are all discussed in this chapter. Chapter 7 presents the conclusions in the light of the research question, discusses the findings, and presents a number of recommendations.

CHAPTER 2

THE CONTEXT OF THE STUDY

CHAPTER 2

THE CONTEXT OF THE STUDY

This chapter outlines the situation prevailing in secondary education in Mozambique within the overall education system. It starts by presenting the geographical, political, and socio-economic status of the country (Section 2.1). Section 2.2 discusses the current education context, focusing on assessment system and practices, curriculum reform, and the importance of linking curriculum, instruction, and assessment. The conceptualisation and rationale of the problem are described in Section 2.3. Finally, Section 2.4 discusses the importance of the study for the Mozambican context.

2.1 Country background information

Mozambique, with an area of 799,380 Km², is located on the eastern coast of Southern Africa south of the equator. It is bordered by Tanzania in the north, Malawi and Zambia in the north-west, Zimbabwe in the west, South Africa and Swaziland in the south-east, and also by South Africa in the south. The eastern part consists of nearly a 2,500 Km coastline facing the Indian Ocean. Data from the National Institute of Statistics (INE, 2007) indicate that the total population of the country is 20.500 million inhabitants. About 56.7% of the population is illiterate.

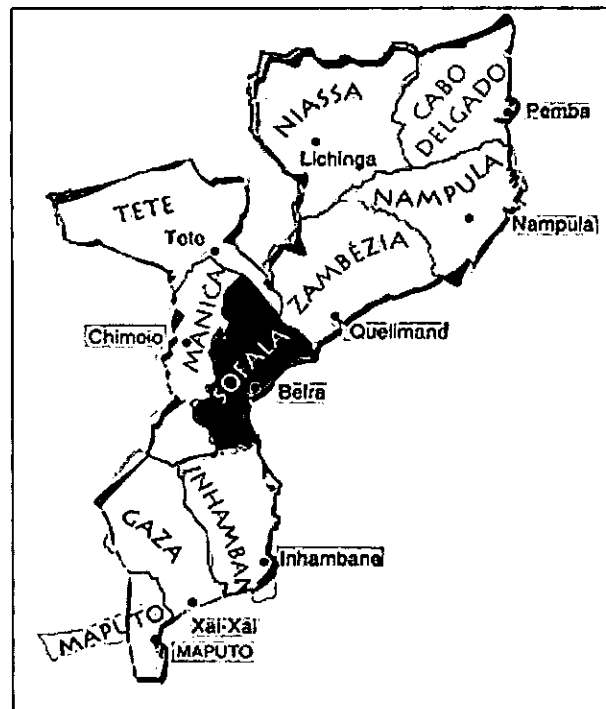


Figure 2.1: Internal borders of Mozambique and its provinces

The country's internal borders are defined by eleven provinces, namely Cabo Delgado, Gaza, Inhambane, Manica, Maputo City, Maputo Province, Nampula, Niassa, Sofala, Tete, and Zambezia. The most populous provinces are Nampula (20% of the national population) and Zambezia (19%). The official language of the country is Portuguese. Sixteen local languages are spoken throughout the country and of these Emakhuwa (26.3%), Xichangana (11.4%), Elomwe (7.9%), and Cisena (7.0%) are the most spoken (INE, 2007). In terms of health conditions, the country is characterised by a diverse distribution of health units. Data from the Ministry of Health, quoted in the Statistical Yearbook (2003), indicate that in 2003 about 53.3% of the country is covered by health posts, 43.1% by health centres, and 3.6% by hospitals. Table 2.1 summarises the country's political, social, and economic situation.

Table 2.1: Mozambique profile

| | |
|---------------------|--|
| <i>The country</i> | Area of 799,390km ² ; date of independence: 25 June 1975. The annual average temperature is 23.8°C. The minimum temperature absolute is 6.8°C and the maximum absolute temperature is 45.3°C. |
| <i>Government</i> | Armando Emilio Guebuza has been the president since February 2005. Mozambique is a Republic and adopted a new constitution in 2005. The parliament consists of 250 members. FRELIMO is the political party in power and the biggest. |
| <i>Capital city</i> | Maputo, with 1,162 000 inhabitants in 2003. |
| <i>The people</i> | Mozambique has a population of 20.5 million inhabitants. In 2003 the illiteracy rate was 53.6 and life expectancy at birth was 46.3. The population is densest in the North and along the coast with Indian Ocean. |
| <i>Currency</i> | The country's own currency is METICAL (MZM). 1 US dollar is equivalent to 24,800 MZM (per August 2005). |
| <i>Languages</i> | The official language is Portuguese. Sixteen local languages are spoken throughout the country with Emakhuwa (26.3%), Xichangana (11.4%), Elomwe (7.9%), and Cisena (7.0%) being the most widely spoken. |
| <i>Education</i> | The enrolment rates by 2004 for general secondary education were 107,301 students: 95,201 students in ESG1 and 12,100 students in ESG2. There are 143 public schools from which 27 are of ESG2. |
| <i>Economy</i> | The human development index in 2002 was 0.354. The adjusted real GDP per capita was 0.360 (UNDP, 2004). Agriculture is the basis of the economy. In 2003 the main export products were cashew nuts, prawns, lobsters, cotton, and wood. Import products are transport and electrical equipment, machinery, vegetable and petroleum products and cereals. |

The following section discusses the current situation in education with the main focus being on assessment practices, curriculum reform, and the importance of linking curriculum, instruction and assessment.

2.2 Educational context

The Mozambican Education Policy defines the provision of education to the population as its main goal while trying to ensure acceptable standards of teaching quality. Secondary education is regarded as an integral component of social and economic development. It

provides essential preparation for mid-level employment and post-secondary education and training, including training for teachers, creating the practical skills that will facilitate rapid integration into society, particularly into the employment market. In line with the Government's objectives, the secondary education sub-sector pursues limited expansion in enrolments, aimed at raising minimum quality standards with a focus on increasing equity in access, particularly for girls and other disadvantaged groups. On the other hand, the sub-sector strives to improve the quality of instruction by equipping schools, investing in teacher training and qualification, providing pedagogical support and supervision, reforming the curriculum, and ensuring the provision of teaching and learning materials. The current situation shows that there is great pressure for an expansion of the secondary education. Increasing numbers of students finish their primary education and are willing to continue their education in the system. The present situation of the secondary education is characterised by low efficiency, high repetition rates with an average of 35% in the ESG1 and 25% in the ESG2, high failure rates and low performance of teachers (MEC&INDE, 2007). There are also huge disparities in the provinces. Despite the fact that community, private and religious schools offer alternative opportunities to some children, their contribution is still insignificant in regard to the improvement of access to the ESG and to the reduction of geographic and gender disparities. The majority of children from poor social groups, particularly from rural areas, do not have access to secondary education and there are significant geographic (North, Centre, South, and rural vs urban) and gender disparities (MEC, 2007).

2.2.1 Assessment system and practices

The Regulation of Assessment for the General Secondary Education document of the MEC highlights two main functions of assessment, namely a formative and a summative function. The formative function, integrating continuous and diagnostic functions, is aimed at undertaking a systematic analysis of the students' results and the reasons for these results. The summative function is orientated towards assessing the level of performance of students and the attribution of a final classification. At the end of the year students obtain a final mark for each subject. This is a pass mark when the student has

equal to, or greater than, 10 out of 20 points and a fail mark when it is less than 10. Most importantly, if a student fails the two core subjects Portuguese or Mathematics s/he cannot transfer to the following grade. Transition from Grade 11 to Grade 12 will take place if a student has (i) a pass mark for Portuguese and Mathematics, and (ii) not more than two fail marks in total. Regarding assessment procedures for teachers in general, the MEC document indicates that, at classroom level, the assessment system consists of three main forms: (i) Activities of Systematic Control (ACS), (ii) Activities of Partial Control (ACP) and (iii) Exams. The ACS and ACP are teacher-made assessments and are undertaken in every grades (from 8 to 12), while the exams are prepared by MEC and only administered at the end of grades 10 and 12 for certification and placement purposes. ACS is meant to be a formative assessment and is to be applied by a teacher to assess parts of a unit of the syllabus. Assessment methods can be:

- oral;
- written, with prior announcement, with a maximum duration of 45 minutes;
- written, without previous announcement, on the theme of the previous lesson or on the homework, with a maximum duration of fifteen minutes;
- homework;
- practical laboratory activities or of another type;
- verification of the student's exercise books;
- activities on selected themes in several units of the syllabus; or
- other activities designed to prove the student's performance in part of a unit.

ACP has a more summative character, as it is meant to assess student performance in the units of the teaching programs. Methods of assessment can be:

- written, with maximum duration of 90 minutes;
- practical activities; or
- research activities.

The examinations are prepared by the Ministry to assess students' knowledge, abilities, and attitudes developed in the education process of each cycle. They can be written or oral depending on the nature of the subject. The written examinations at the end of Grade 12

are the same nationally for all students and schools (public and private) of the same level. The national examinations of all subjects (including Physics) are centrally developed and distributed by the MEC through the Provincial Directorates of Education and Culture. In the MEC, multidisciplinary groups of officials from the National Directorate of General Education (DINEG) develop four exam proposals for each subject, taking into consideration teacher proposals. The teachers who submit proposals are carefully selected from schools. The exam proposals are then sent to the National Council of Exams, Certification and Equivalence (CNECE), a unit responsible for examinations and level certifications within the MEC. This department is then responsible for the analysis and approval of the submitted proposals and for the final selection of the exam papers. The examinations are administered in each school by the school's Commission of Teachers, and are graded by the same Commission. Provincial and district inspectors are sent to schools to monitor the whole process, in particular to control whether the scoring guides previously prepared with the exam papers are correctly followed. Concerning the number of written tests during the school year the situation is as follows:

In the Cycle 2 (Grades 10 to 12) the number of ACP per semester is three for all subjects. As for ACS, the teacher may administer as many tests as s/he wants, but three assessments should be registered. The student grades are based on the 0-20 points scale. The semester and the yearly mean are calculated according to the formulae (rounded) set out below.

Table 2.2: Formulae for calculating student marks

| | |
|---|---|
| 1. Semester Mean: | $\frac{\text{Mean of ACS} + 2 \times \text{Mean of ACP}}{3}$ |
| 2. Yearly Mean: | $\frac{\text{Mean of the 1st semester} + \text{Mean of the 2nd semester}}{2}$ |
| 3. Final Mark (G12): | $\frac{2 \times \text{Yearly Mean} + \text{Exam Mark}}{3}$ |
| In the case of the Languages, the exam mark is obtained in the following way: | |
| Exam Mark: | $\frac{\text{Mark of the Written Exam} + \text{Mark of the Oral Exam}}{2}$ |

ACP = Activities of Partial Control, ACS = Activities of Systematic Control, G12 = Grade 12

Annually, schools should send, before a date previously set, the statistical data of the examinations to the CNECE in the MEC, after having completed the grading of the examinations.

2.2.2 Curriculum review and reform

The MEC's general policy for secondary education is well outlined in the Secondary Education and Secondary Teacher Education Strategic Plan document (MinEd, 2001). In terms of the secondary school curriculum, the document indicates, on the one hand, that the present curriculum is characterised by an excessive number of subjects. The ESG1 curriculum has eleven subjects, and ESG2 has three streams of seven subjects each. Life skills do not feature sufficiently in the curriculum of either ESG1 or ESG2. One of the consequences of this situation has been high repetition and dropout rates, especially in the early years of ESG1 (Palme, 1992). Students in ESG2 tend to choose social studies, perceiving it to be easier. In summary, the current ESG2 curriculum is extremely academic and demands a high level of theoretical knowledge, without promoting practical skills that would facilitate the integration of graduates into the labour market. This diminishes the possibility of developing a critical mass of mathematic and Science students, and is subsequently reflected in the number of new entries to university, particularly of girls. The present curriculum, particularly in ESG2 serves the secondary education by selecting students for post-secondary education. Competencies in languages are very influential in the educational system. While a competency in the Portuguese language is a determining factor in tests and examination results, the demands of globalization and rapid economy growth mean that the employers attach demand in English competency even greater than few years earlier. There is little gender visibility in the curriculum – across subjects, in extra-curricular activities, and in the preparation of instructional materials. The curriculum and the assessment of student learning levels are crucial to educational provision. But curriculum issues are inextricably related to the major policy issue of the purpose of secondary education. Curriculum changes can affect the number of teachers needed in each subject, and considerations related to the size and

nature of secondary education (formal schools and open and distance education). The Ministry of Education and Culture recognises that, to achieve effective results, a co-ordinated curriculum and assessment strategy needs to be designed.

Regarding the external efficiency of the system the government recognises that there is a need for much more study on the issue of the relevance of the curriculum to the world of work. On the one hand, reports (MEC & INDE, 2007; MinEd, 1998; MinEd, 2001) state that there will be a critical shortage of workforce to sustain and improve economic progress in Mozambique; on the other, there are indications of comparatively high levels of unemployment among educated workforce. There is a need to produce secondary level graduates to ensure an adequate supply of teachers and other public servants (in the health and extension agriculture services). Current educational thinking is that a good general, flexible secondary education is needed, as the future job market is uncertain and subject to rapid technological change.

The assessment system, particularly the examinations, is seen to be potentially vulnerable and characterised by inefficiencies and a low level of control, which results in high rates of repetition and dropouts, particularly in the early classes of secondary education. Therefore, the MEC is conscious that the system of assessments and examinations is in need of major overhaul and this can only be achieved as part of the secondary education curriculum reform (MEC & INDE, 2007; MinEd, 2001). Summative assessment at the secondary level is mentioned in the MinEd document as being one of the key components in monitoring the quality of secondary education. This type of assessment refers to ACP and normally takes place at the end of a learning period, usually consisting of three main tests that are written at the end of programme units, and aims to determine how much of the content of the subject the students know.

The most pressing issue is the steady increase of unqualified teachers in schools, which indicates two urgent, related concerns: On one hand the importance of considering the needs of teachers (both qualified and unqualified) should be taken into account seriously, and on the other there is the need for a qualification programme for those teachers.

Although there are some in-service training initiatives (e.g., by Eduardo Mondlane University, Distance Education Unit under the MEC), UP should reconsider its courses to improve the impact. There are indications that UP is developing a strategic plan which includes a move to establish the three-year course as a basic teacher qualification for teachers of all subjects. It is expected that successful graduates would then be able to apply for postgraduate study through a one to two year Master's Degree. These are, however, long-term plans with long-term expected results. More urgent action should be taken, especially with those teachers working in the classrooms today.

Currently the MEC is working on a curriculum revision for secondary education to be implemented soon (as at August 2007). According to official documents available in the MEC (INDE, 2005; MEC & INDE, 2007; MEC, 2007; MinEd, 1998; MinEd, 2001) the strategic priority for secondary education is the revision of the curriculum aimed at incorporating less of an academic (theoretical) and more of a practical orientation; the revision of the assessment practices and the student learning outcomes; and the expansion of the in-service training opportunities for secondary school teachers. An example of a measure aimed at improvement included in the draft documents is that teachers must be made familiar with the characteristics and usefulness of formative assessment (INDE, 2005; MEC & INDE, 2007). Teachers will have to learn how to develop formative assessments that will be used in the classroom to inform and enhance the learning process.

2.2.3 Importance of linking curriculum, instruction, and assessment

To address the problems mentioned above, it is important that in the revised curriculum, curriculum (what should be taught), instruction (what is being taught and how), and assessment (assessing student learning) are linked. This is reflected in Popham's position (2002), when he states that assessment in the classroom is central to student learning. Firstly, it determines whether students are moving satisfactorily toward instructional outcomes that the teacher and the educational system through the national curriculum are seeking to promote. Secondly, assessment is important for the instructional process because teachers can discern (i) where to put direct instructional energies to ameliorate

student weaknesses and (ii) what already mastered skills or knowledge can be instructionally avoided. These functions of assessment illustrate the importance of the linkage between curriculum, instruction and assessment and the crucial role played by teachers in making this connection work effectively. Although teachers have shown difficulties in adapting to a student-centred approach of instruction due to the 'tradition' in Mozambican education and the characteristics of the system during their training, which advocated a more teacher-centred approach, the present MEC policy recommends a more constructivist and student-centred approach.

It is the author's perception that improvement of teacher assessment practices, both formative and summative, could help monitor improvements in the quality of students' results and the performance of the educational system, and achieve the intended learning outcomes.

It has been argued that one of the major challenges facing curriculum improvement is creating a balance and consistency between the various components of a curriculum (refer to Chapter 4, subsection 4.3.2). These components may range from aims and objectives of learning, content, learning activities, teacher role, materials and resources, grouping, location, time, to assessment. However, the choice of focusing in this study on improvement of the assessment component is based on the conviction that this component deserves separate attention at all levels. But the other components of the curriculum and the process of instruction have not been neglected because a careful alignment between assessment and these aspects is critical for any successful curriculum change. Furthermore, the improvement of assessment practices can easily raise learning standards (Weeden, Winter & Broadfoot, 2002).

Section 2.2 above has provided the educational background of the study. The following section addresses the conceptualization and rationale of the problem.

2.3 Conceptualisation and rationale of the problem

It is against the background of poor student results and the lack of effectiveness of present assessment practices in schools, as described in previous section, that this study aims at investigating formative and summative assessment practices in Physics teaching in Grade 12, and at developing an intervention aimed at improving the assessment practices for Physics in ESG2. The problem is perceived as a problem at school level. Before elaborating on the problem, first the environment of the school is considered as far as it may have an influence on the problem area (see Figure 2.2).

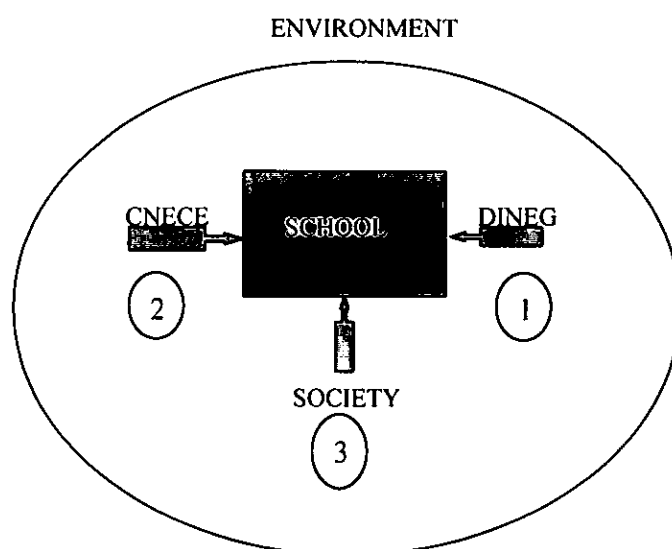


Figure 2.2: Elements of the school environment

Three institutions are crucial when taking into consideration the educational environment of the school.

- 1 DINEG: National Directorate of General Education – this is a unit in the Ministry of Education and Culture with responsibility for the coordination of the curriculum review process for the secondary schools. It also has a role in supervising schools but it has no direct influence on teaching and learning processes.
- 2 CNECE: National Council of Exams, Certification and Equivalence – this is a unit within the Ministry responsible for setting the goals, objectives and standards of assessment by

the examinations in schools. The unit is also responsible for certifying students' credentials. Therefore, this unit determines to a large extent the focus of classroom assessment practices.

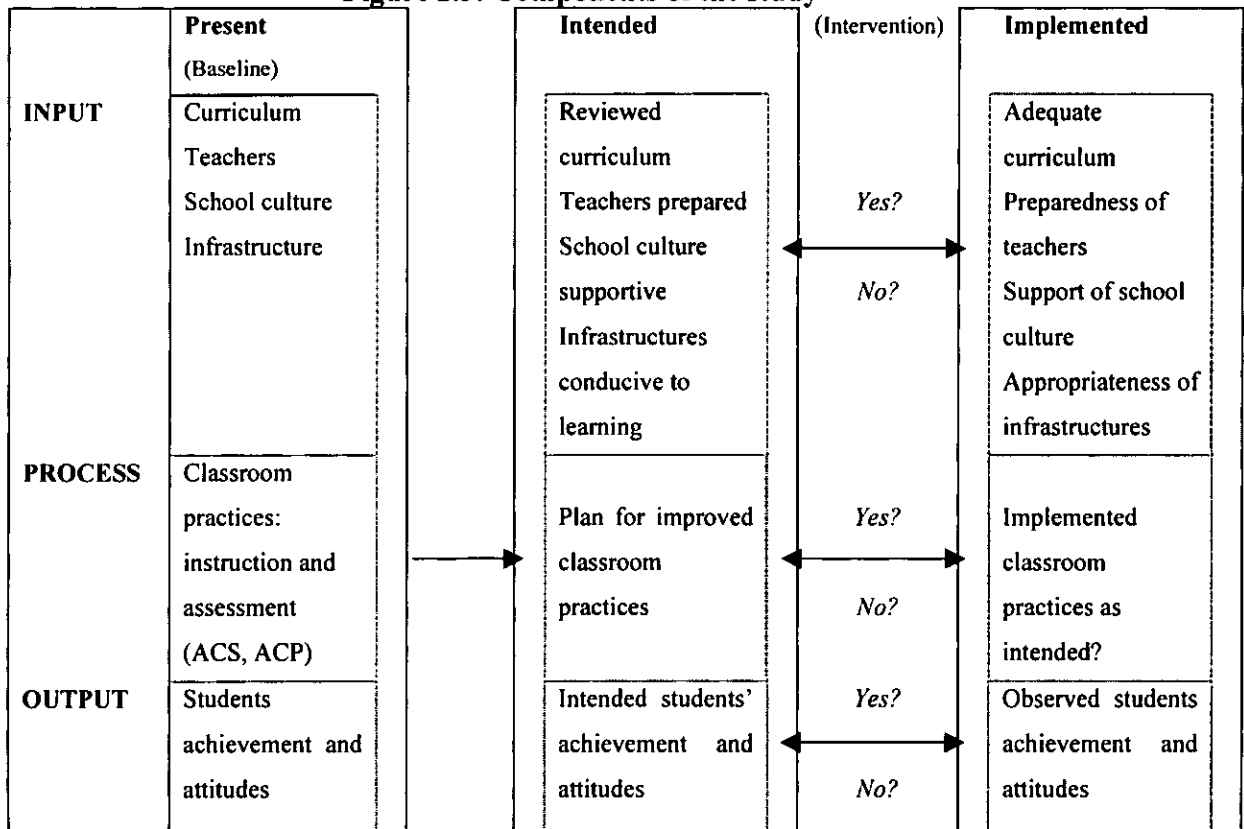
3

Society: (non-) academic community, employers, and donors - this entity encompasses parents, higher education, employers, NGO's and non-profit organisations who play a role in school problem awareness and who expect that schools meet the educational expectations of the society.

All these entities together establish the context for the schools and they have 'external' influences on the school. They have a particular interest in school results and may influence, in one way or another, the course of the study. Therefore one constantly needs to be alert about their possible influence on the problem under investigation.

It is essential to have a good understanding of the present assessment practices in schools and classrooms in order to design an effective 'intervention in assessment'. This implies that the present study should include a Baseline Survey to develop a good understanding of the research problem prior to the Intervention Study of this research. The two parts are reflected in Figure 2.3.

Figure 2.3: Components of the study



The school and classroom practices are looked at from an input-process-output (IPO) model. The model categories were derived from Howie (2002) and Shavelson et al., (1987) and are broadly considered. In later stages, they will be refined on the basis of the literature review and other research activities.

Although one has to acknowledge that the school environment does have influence on the school, for instance by setting policies and regulations for the school, this research focuses on processes within the school and will accept policies and regulations as setting the boundaries for what is possible in the schools. In other words, the research focuses on the level of the school, the *meso* level. In the IPO model for the problem being studied, the stakeholders are therefore:

- a) the students as the subjects who need feedback to monitor their own learning;

- b) the teachers who need the achievement and performance data for providing feedback to students and for themselves on the improvement of their instructional and assessment practices; and
- c) the leadership in the school who has to arrange for the necessary conditions to make the intervention happen, such as school policy changes (if necessary), facilitating staff development and providing infrastructures.

Given the aims of the study, it is necessary to consider the IPO model for three different situations reflecting the different aspects of the problem which are discussed below.

1. The 'Present' is aimed at gathering baseline information on the conditions in place allowing that the current teachers' classroom practices take place, which finally lead to students' results. At the input level, there are: the present curriculum (under review), the teachers, the school culture, and the infrastructure as the environment, setting the boundaries of the instructional processes. At the process level, there are the classroom practices, comprising instructional processes and assessments (ACS and ACP) as the activities to be primarily investigated by the study. At the output level, there are the students' achievement and attitudes towards their learning of Physics.

2. The 'Intended' consists of a plan for improved classroom practices based on state-of-the-art literature and a thorough understanding of the present situation provided that, if the desired conditions (reviewed curriculum, teachers prepared, school culture supportive, infrastructures conducive to learning) are in place, the intended students' achievement and attitudes will be attained.

3. The 'Implemented' is the study of the actual implementation of the intervention developed under 'Intended'. Data were collected to find out whether classroom practices, including the desired input conditions and students' achievement and attitudes, were implemented as intended.

2.4 Importance of the study

Within this conceptualisation, the importance of this study for Mozambican context can be summarised as set out below.

Firstly, the improvement of assessment practices could help to monitor qualitative improvements in student results and of the performance of the educational system. It is expected that the assessment practices being investigated by this study will be an important means of improving teacher assessment practices, and must necessarily be accompanied by changes in the instructional process. For example, effective assessment prototypes for Physics can use formative and summative approaches as means to contribute to improving student learning in the classroom. The students and the teachers are therefore considered to be the prime beneficiaries of the study. Teachers need the data for providing feedback to students and for themselves and the students are the subjects who most need the data for monitoring their own knowledge. Implementing new strategies is always a difficult task to accomplish. Therefore, the study results also provide a framework as to how to support teachers in making the new assessment approaches more relevant for their classroom practices.

Secondly, good mastery of Physics and good final results in Physics are important for students themselves because of the implications related to the policy dealing with grade-to-grade promotion. Furthermore, good results are part of the high school diploma, which is a gateway to the university level. For the teachers, school results not only help them to take decisions on what instructional objectives to pursue but, as Popham (2002:11) puts it, “the results influence public perceptions of school effectiveness and respond to the pressures from above”.

Thirdly, the improvement of teacher assessment practices by developing prototypes can serve as a supporting tool for the Ministry in monitoring the quality of education. The MEC has to rely on assessment data as indicators of the performance of the system. Furthermore, the improvement in student results is also important for the MEC in particular and for the society in general. For the Ministry, these results are used as

indicators of how well the system is performing. For society (donors, employers, and (non-) academic community), good results make the financial contribution to education noteworthy and, more generally, success in school signifies order and control. It evokes a traditional set of educational, social, economic, and moral values.

Finally, the study is a contribution to the research in Mozambique, particularly in terms of secondary education where studies of this nature are still scarce compared to those dealing with primary education. The combination of studies at both levels will enable decision-makers to plan adequately and to monitor the performance and quality of the education system as a whole.

CHAPTER 3

LITERATURE REVIEW AND CONCEPTUALISATION OF THE STUDY

CHAPTER 3

LITERATURE REVIEW AND CONCEPTUALISATION OF THE STUDY

This chapter presents the literature relevant to a discussion of assessment strategies in Physics and the conceptual framework of the study. It starts by presenting the research questions and the sources used to find out what is already known about the research questions (3.1). The method used to review the literature, as well as the keywords used to search the information, are also presented in this section. The definition of terms used in this research, and a number of different learning theories used as a platform to guide the discussion are presented and discussed in Sections 3.2 and 3.3 respectively. Section 3.4 begins with presenting arguments in favour of constructivism as central to Science learning (3.4.1). It also presents and discusses arguments about the roles of both teachers (3.4.2) and students (3.4.3) in assessment, and contexts within which different assessment strategies can be considered (3.4.4). The contributions from various authors who worked within the African context are presented and discussed in Section 3.5 as methodological illustrations for the educational design research approach to research on interventions. Drawing from the literature review findings and contextualising the problem posed by the research questions, Section 3.6 presents the issues used to address the main research question adequately.

3.1 Introduction

The research questions of this study are: What assessment strategies do secondary school Physics teachers in Mozambique apply, and how can they be improved? In order to find an answer to these questions, two research phases were considered. Firstly, the study investigated through a survey approach the assessment practices used by secondary school

teachers in schools. The research question addressed by the Baseline Survey, as referred to in Chapter 1, is:

What assessment strategies do Grade 12 teachers in Physics in Mozambique apply and what can be said about their quality?

The second component of the study was characterised by an intervention approach aimed at improving some of the assessment practices used by Physics teachers in schools. The Intervention Study addressed the following research question (refer also to Chapter 1):

How can teacher assessment strategies be improved?

Both the Baseline Survey and the Intervention Study were informed by the literature review.

To find out what is already known from the literature about the two study phases, various sources were reviewed both in printed form and electronically. General open source software articles and information were used in order to gain a global idea of the current assessment strategies used by Grade 12 teachers in Physics, as well as the quality of these strategies. Keywords like ‘assessment strategies’, ‘assessment toolkits’, ‘science toolkits’, and ‘mechanical Physics and assessment’ (for Grade 12 or its equivalent) were used to obtain information on what is known about the research question addressed by the survey approach and what the relevant arguments of several scholars are on assessment for learning Physics in general. Thus, the information obtained from the review of these keywords was used to inform the Baseline Survey, particularly in terms of the characteristics of assessment strategies and teaching practices.

The research question addressed by the intervention component of the study was also informed by the keywords referred to above, but the following additional search terms were also used: ‘performance assessments’, ‘formative assessment’, ‘summative assessment’, ‘prototypes’ and ‘demonstration experiments in secondary education’. The term ‘demonstration experiments’, is included in this list because, in Science education

literature, it has been used with different interpretations. Some authors have been using it with the same meaning as 'laboratory experiments'. This term in this dissertation, refers to activities in the Physics classroom in which students observe, carry out small experiments, interpret phenomena or events occurring to objects, and report the findings, and are all guided by a teacher and teaching materials. In general, relevant considerations following from these keywords are that the future of assessment nowadays depends on the ability to assess student skills in performing tasks; on the power of formative assessment in monitoring student learning; on the role of summative assessment for accountability purposes; and on the importance of demonstration experiments in science subjects (Airasian, 2001; Black & Atkin, 1996; Gardner, 2006; Popham, 2002; Race et al., 2005). Besides using scientific literature from the academic libraries of the University of Pretoria and the Eduardo Mondlane University, other sources were explored by using the following search engines on the internet:

- Academic Information System (University library);
- CAB abstracts;
- ERIC;
- Google scholar;
- ISI web of science;
- Science direct; and
- Tucks (electronic journals for which the University of Pretoria has a subscription).

Table 3.1 indicates which keywords were used in which database search engines.

Table 3.1: Database search engines

| | Academic Information System | CAB abstracts | ERIC | Google scholar | ISI web of science | Science direct | Tucks |
|--|-----------------------------|---------------|------|----------------|--------------------|----------------|-------|
| Assessment strategies | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Assessment toolkits | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Formative assessment | ✓ | ✓ | ✓ | ✓ | | | ✓ |
| Demonstration experiments in secondary education | | ✓ | | ✓ | ✓ | ✓ | |
| Mechanical Physics and assessment | | | | ✓ | ✓ | | |
| Prototypes and Physics | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Performance assessments | ✓ | | | ✓ | | | ✓ |
| Science toolkits | | | ✓ | ✓ | ✓ | ✓ | |
| Summative assessment | ✓ | ✓ | | ✓ | | | ✓ |

A significant number of books and articles were found from all these sources. A subsequent selection was made on the basis of years of publication, the education level, and the context. Although the review considered non-African contexts (especially European and American), the focus was on those references reporting research undertaken within and about an African context, or with some similar characteristics, written during the period 1996-2006, and on Grade 12 in Mozambique.

During the review of the databases and search engines, it emerged that some relevant authors were quoted frequently. As the central areas of this research are classroom assessment (both formative and summative), performance assessment, and Physics learning, the literature review focused on authors in these areas, such as: Airasian (2000, 2001), Black (1998), Black and Atkin (1996), Black et al., (2003), Dekkers (1997), Harlen (2006), Kathy and Burke (2003), McMillan (2001), Moskal (2003), Muller (2006), Mutimucuo (1998), Popham (2002), Race et al., (2005), Stiggins (1987), Treagust et al., (1996), and Weeden et al., (2002).

Current writings (such as theses) in the field of Science teaching and learning and reporting investigations within the African context and elsewhere often follow an educational design research approach. As this approach is relevant when addressing the intervention research question, a number of dissertations (viz. Mafumiko, 2006; Motswiri, 2004; Ottevanger, 2001; Teclé, 2006) have been included in the literature review as well.

Given the research questions for this study, the literature review focused specifically on issues related to the 'assessment strategies' used for Science subjects, the type of 'learning evidence' in 'alternative', 'authentic', 'formative', and 'performance' kinds of assessment, and the role of both teachers and students in assessment. In summary, all of the authors reviewed emphasise the need to improve the current teachers' assessment strategies in schools.

Having presented the research questions being addressed by both the Baseline Survey and the Intervention Study, as well as the sources for the literature review, the next sections deal with: definitions of terms mostly used in the dissertation (3.2); the arguments of several authors on the theoretical orientation of assessment in education (3.3); the roles of teachers and students in assessment; and the need to improve current assessment strategies in Science, with particular emphasis in Physics teaching and learning (3.4).

3.2 Definition of terms

In this dissertation, several key terms are frequently used. For the purpose of the present study and for the reader's convenience, definitions of key terms used in this study are presented in this section. However, before doing this, it is relevant to clarify the distinction between assessment *for* learning and assessment *of* learning. According to Black et al., (2003) assessment *for* learning is any assessment where the first priority is to serve the purpose of promoting student learning. This kind of assessment is usually informal, embedded in all aspects of teaching and learning, and conducted differently by different teachers as part of their own individual teaching styles. Assessment *of* learning is for grading and certification, occurs in formal settings or rituals, involves non-frequent

tests, is isolated from normal teaching and learning, is carried out on special occasions, and is conducted by methods over which individual teachers have little or no control. This study is about assessment *for* learning.

In this section, the following concepts and terms will be defined and discussed namely: *assessment, authentic assessment, classroom assessment, formative assessment, paper-and-pencil tests, peer-assessment, performance assessment, portfolios, projects* and *prototypes*.

Assessment

The term 'assessment' is defined in the Glossary of the 1999 Standards for Educational and Psychological Testing as "any systematic method of obtaining information (from tests and other sources) to draw inferences about characteristics of people, objects or programs" (Chatterji, 2003). Airasian (2001) defines assessment as the process of collecting, synthesising, and interpreting information to aid in decision-making. For this author, assessment involves more than administering, scoring and grading paper-and-pencil tests, and includes the full range of information teachers gather in their classrooms. This information helps teachers to understand their students, monitor instruction and establish a viable classroom culture. Given the two definitions, the researcher's perception of assessment is that, when the term is used to refer to assessment processes, the activities include writing items or designing an assessment tool, making observations or gathering data using an assessment tool, scoring responses from an assessment tool, developing a scale with specified properties, and administering an instrument using prescribed guidelines. Briefly, the term assessment in the present study is used to refer to a variety of ways teachers gather, synthesise, and interpret information.

In this study, all kinds of assessments used by teachers in schools, from the traditional approaches of paper-and pencil tests to a more constructivist and dynamic process of gathering information following some prescribed guidelines are termed assessment practices or assessment strategies.

Authentic assessment

This is the form of assessment in which students are asked to perform real-world tasks that demonstrate meaningful application of essential knowledge and skills. McMillan (2001) for instance, defines authentic assessment as an assessment that is constructed to be more consistent with what people do in situations that occur naturally outside classroom. An authentic assessment usually includes a task for students to perform and a rubric by which their performance on the task will be evaluated. In order to determine whether authentic assessment is successful, the school must ask students to perform meaningful tasks that replicate real world challenges to see if students are capable of doing so (Wiggins, 1993). The Mozambican curriculum goals stated in the Physics Syllabus for Grades 11 and 12 emphasise that “the starting point for students’ knowledge acquisition is practical work and the nature and its phenomena is the stepping stone for proving any formulated hypothesis” (MinEd, 1997:1). This shows that authentic assessment drives the curriculum. Teachers have to determine the tasks that students will perform to demonstrate their mastery, and then a curriculum is developed that will enable students to perform those tasks well. Some educators (Meyer, 1992; Stiggins, 1987) distinguish authentic assessment from performance assessment by defining performance assessment as performance-based, but with no reference to the authentic nature of the task.

Classroom assessment

Popham (2002) defines ‘classroom assessment’ as the assessment that comprises a number of assessment decisions taken by the teacher during the teaching and learning process. These decisions occur mainly within the classroom environment, or are informed by the classroom climate. In some educational systems the term ‘classroom assessment’ is referred to as ‘non-standardised assessment’ meaning that the assessments are constructed by teachers, specifically for classroom use, and focused on the particular type of instruction provided in that classroom (Airasian, 2001). The information from classroom assessment is used to provide feedback about the performance of students in a single class, not of students in other classes. In the present study, the term classroom assessment is used to refer to all kind of assessments undertaken by the teachers in the classroom during

the instruction process. Therefore, the terms ‘classroom assessment’ and ‘teacher assessment’ are used interchangeably.

Formative assessment

According to Black et al., (2003:122) “formative assessment is a process, one in which information about learning is *evoked* and then *used* to improve the teaching and learning activities in which teachers and students are engaged”. This definition takes the idea of formative assessment beyond the ‘micro-summative’ assessments of classroom tests and homework. It broadens the sources of evidence and solidifies the notion of what should be done with the evidence. The sources from which evidence can be drawn do not exclude information gathered from formal assessments, but more important than the source of evidence is the idea that the information obtained affects subsequent teaching and learning activities.

Paper-and-pencil tests

Paper-and-pencil tests are assessments in which students write their responses to questions or problems (Airasian, 2001). Examples of paper-and-pencil tests are essays, multiple-choice tests, written assignments, written reports, a drawn picture, or a filling in of a worksheet. In general, paper-and-pencil tests are of two types: selection type – where the student responds to each question by selecting an answer from the choices provided, and supply type – which requires a student to produce or construct a response to a question or task.

Peer-assessment

Generally, peer-assessment is an assessment of the work of others by people of equal status and power. The term can also be used to describe approaches to accountability carried out on behalf of government agencies. In the context of student learning, peer-assessment may be divided into the giving and receiving of feedback and making formal estimates of worth of other students’ work (Brown et al., 1997). It usually involves an element of mutuality and, beneath the processes of giving and receiving feedback, there are implicit criteria of what counts as ‘good’ for different purposes and contexts. The

students' notions of 'good' and 'poor' can be generated in large group classes and tutorials and so provide a basis for reflective learning and more formal peer-assessment. Thus, peer-assessment is rather a tool for learning than a tool for summative assessment.

Performance assessment

The length of responses in items where the student has to supply a response can vary substantially. When, for instance, a student is required to produce complex constructions such as Science experiments reports, book reviews and/or class projects, these assessments are termed performance assessments. As the term suggests, performance assessment requires a demonstration of student skills or knowledge (Moskal, 2003). Performance assessment can take many different forms, which include written and oral demonstrations and activities that can be completed by either a group or an individual. A factor that distinguishes performance assessments from other extended response activities is that they require students to demonstrate the application of knowledge to a particular context. Through observation or analysis of a student's response, the teacher can determine what the student knows, what the student does not know, and what misconceptions the student has with respect to the purpose of the assessment.

Portfolios (in education)

Portfolios, in general, constitute the chief method by which certain professionals such as models, artists, photographers, architects, and journalists display their skills and accomplishments. In the domain of education, a portfolio is a systematic collection of a selected student work (Popham, 2002). It engages students in assessing their own progress or accomplishments over time and establishes ongoing learning goals. A portfolio is not an unrelated collection of student's work but contains consciously selected examples of work that are intended to show student's growth toward important learning goals.

Projects

Projects are purposeful activities of a student or a group of students with a limited duration. They can be finished products which include all written tasks that the students do and which reflect a certain development process of collecting, interpreting and reporting data (Brown et al., 1997). Thus, they are usually carried out in the final year of the course but also may be used in the first year of the course to encourage students to become active, independent students. Projects may be laboratory-based, library-based, work-based, studio-based or community-based. The main purpose of the projects is to develop enquiry-based student skills. While projects enables a student to explore deeply a field or topic, develop an initiative, provide personal ownership of learning and enhance time- and project-management skills, they have a disadvantage of being time consuming to set up, monitor and provide feedback.

Prototypes

A prototype is a model upon which other similar materials are based. It represents all products that are designed before the final product is constructed and fully implemented in practice (Nieveen, 1999). In its initial stage, a prototype can be developed, discussed and modified as required to build consensus. It is, therefore, designed with particular care. In the process of developing a prototype, developers come to an agreement on what to show and how to show it. In the context of the present study, prototypes are physical exemplary materials on teaching and assessment strategies for teachers to use in the classroom that demonstrate the acquisition of student skill and are based on a set of established performance criteria.

This section has presented and discussed the terms or concepts frequently used in the dissertation. The next Section (3.3) provides a number of perspectives of different authors on assessment, which serve as a theoretical orientation to the study.

3.3 Theoretical orientation in assessment

In this section, three aspects are regarded as being relevant for providing a theoretical orientation in classroom-based assessment for this study. The aspects are: (i) the objective of assessment and the process of giving feedback to students; (ii) the need for teachers to conduct effective assessment *for* learning; and (iii) the teachers' preparedness on conducting assessment that can generate evidence of authentic learning.

In terms of the first aspect and according to several authors (Black, 1998; Black et al., 2003, Kathy & Burke, 2003; Lin & Gronlund, 2000), assessment may be conducted to serve different purposes, such as assessment to satisfy demands for public accountability; assessment to report an individual's achievements; and assessment to support learning. The focus of assessment in this study falls within the latter purpose (supporting learning) because the study aims to improve those assessment practices that teachers apply. The rationale of focusing on this purpose is that the main aim of schools is to promote student learning and the teacher needs constant information about what the students know. Ideally, assessment should provide short-term feedback so that obstacles can be identified and tackled at an early stage in the learning process. This is particularly important where the learning plan is such that progress with one week's work depends on a grasp of the ideas discussed in the previous week. This type of assessment aims at improving learning, and is called formative assessment or assessment *for* learning.

It is clear that this assessment is the responsibility of the classroom teacher, but others, inside and outside the school might support this work by providing appropriate training and methods for conducting such an assessment. In Science subjects like Physics, however, evidence that a formative assessment is really improving learning must be accompanied by a type of assessment where students are asked to perform real-world tasks and demonstrate the meaningful application of knowledge and skills. This leads to what McMillan (2001) calls authentic assessment and its success depends very much on support that the teacher must receive from various educational stakeholders inside and outside the school. Therefore, in providing such support to teachers, Nuttall (cited in Kathy & Burke,

2003), argues that it is relevant for teachers to know how to generate evidence of authentic learning. Authentic learning, as the Physics Syllabus for Grades 11 and 12 recommends, is crucial for learning science and specifically experimental subjects like Physics. Nuttall (1987) also describes a number of criteria for tasks that validly assess learning, namely: (i) tasks that are concrete and within the experience of the individual; (ii) tasks that are presented clearly; and (iii) tasks that are perceived as relevant to the current concerns of the student. The value of these tasks, in researcher's opinion, allows students to demonstrate good performance because they promote interaction between students and the teacher. In addition, they allow the teacher to get into the students' thinking and reasoning and to evaluate their potential.

Bell and Cowie (2001) distinguish between two types of formative assessment, namely planned formative assessment and interactive formative assessment. These authors suggest that planned formative assessment is used to elicit permanent evidence of students' thinking, and such assessment occasions are semi-formal and may occur at the beginning and end of a topic. A specific assessment activity is set for the purpose of providing evidence that is used to improve learning. All the information is elicited through the task set and the teacher and the student act on this information with reference to the topic itself, with reference to the students' previous performance, and with reference to how the students and the teacher are proposing to take learning forward. Interactive formative assessment is described by Bell and Cowie (2001) as taking place during student-teacher interaction. This refers to the incidental or ongoing formative assessment that arises out of learning activity and cannot be anticipated.

As is the case with planned assessment, in interactive formative assessment the purpose is to improve learning by mediating the student learning. The process involves the teachers noticing, recognising and responding to students' thinking and it is more teacher- and student-driven than curriculum-driven. Unlike the kind of permanent information that accrues from planned assessment, this kind of assessment generates information that is ephemeral. The latter kind of formative assessment is crucial for this study because it is

important for enhancing student learning, and therefore, the teacher must be supported in knowing how to react in relation to what is deemed at the time to be worth noticing in the student. Unlike in the planned formative assessment where there is a longer time gap in responding, in the interactive formative assessment, the teacher's response is immediate, and the kind of planning that can still be made is on how to facilitate dialogue and tasks between him/her and the students. In an interactive assessment, students are given opportunity to argue about the assessment tasks and to challenge teachers' responses to their questions.

As for the importance of immediate and ongoing feedback, Race et al., (2005) elaborate on how quality feedback can best be given to students. Amongst the several aspects of quality feedback referred to by these authors, they mention the following aspects of quality feedback: (i) time - the sooner the feedback is given the better; (ii) personality - it needs to fit each students' achievement; (iii) expressed - whether congratulatory or critical; and (iv) empowerment - both congratulatory and critical feedback must not dampen learning, but rather strengthen and consolidate it.

In conclusion, for the assessment objectives and feedback, three major aspects provide orientation to the review on assessment for this study. Firstly, the assessment is carried out to support learning, therefore, the provision of feedback should be on a short-time basis so that obstacles in the learning process can be tackled in good time (Black, 1998; Race et al., 2005). Secondly, teachers need to have at their disposal certain students' tasks that can validly assess particular learning and generate evidence of authentic learning (Kathy & Burke, 2003; Lin & Gronlund, 2000; Popham, 2002). Thirdly, immediate and ongoing feedback is crucial to facilitate student-teacher interaction (Bell & Cowie, 2001; Race et al., 2005).

With reference to carrying out an effective assessment *for* learning, it is worth mentioning that it requires having students actively engaged in finding solutions to problems they face and developing the ability to construct knowledge. In this process, the role of the teachers as facilitators is crucial in monitoring the assessment practice. James (2006), on the relationship between assessment practice and the ways in which the processes and

outcomes of learning are understood, argues that three theories of learning and their implications for assessment practice can be distinguished. These are discussed below.

- **Behaviourism:** this is where the environment for learning is the determining factor, the learning is the conditioned response to external stimuli, and rewards and punishments are the powerful ways of forming or eradicating habits. The implications for assessment practice are that the progress is measured by timed tests, performance is interpreted as either correct or incorrect, and poor performance is remedied by more practice in the incorrect items.
- **Constructivism:** this is where the learning environment is determined by prior knowledge - what goes on in people's minds - emphasis is on 'understanding', and problem solving is the context for knowledge construction through deductive and inductive reasoning. The implications for assessment are that self-monitoring and self-regulation are relevant dimensions of learning, and the role of the teacher is to help 'novices' to acquire 'expert' understanding of conceptual structures and processing strategies to solve problems. When students are involved in the construction of their own learning through formative assessment, they develop the ability to monitor and regulate their learning agenda.
- **Socio-culturalism:** this is where learning occurs in an interaction between the individual and the social environment. Thinking is conducted through actions that alter the situation and the situation changes the thinking. The implication is that, prior to learning, there is a need to develop social relationships through language, because it represents the central element to our capacity of thinking.

It has been argued that the latter theory is not yet well worked out in terms of its implications for teaching and assessment (James, 2006). Teaching and learning tasks need to be more collaborative and students need to be involved in the generation of problems and of solutions, because the current perspective of assessment within this perspective is still inadequately conceptualised. For the context of this study, the constructivist theory of learning is recommendable. The reason is that the Mozambican teaching system emphasises the importance of considering children's prior knowledge before helping them understand other conceptual structures. The implication of this choice for assessment is

that this construction of children's own learning can be easily facilitated through formative assessment advocated by this study.

Regarding teachers' readiness to conduct assessment, i.e., what teachers need to know in order to assess learning and to generate evidence of authentic learning, James and Pedder (2006) explain that, as a pre-condition to enhance assessment for learning, changing pedagogical practice should be taken into account, particularly with the roles of both teachers and students. Teachers should be supported in developing skills to plan assessment, interpret learning evidence, and provide feedback to students and support them in assessing their work and that of their peers. This means that teachers need to practise new roles and try and evaluate new ways of thinking. Students should also be helped to take on new roles as students. They should understand the learning goals and identify the criteria used for assessing their progress, develop skills of peer and self-assessment, and make progress through constructive formative feedback from peers and their teacher. This implies developing a language and eagerness for talking about teaching and learning.

Within this framework, the present literature review seeks, amongst other issues, to understand how teachers, who have been trained following a behaviourist theory of learning (Buendía Gomez, 1999; Siteo, 2006), can facilitate student learning (and assessment) in a constructivist approach, as advocated by the Mozambican syllabus and recommended by recent literature, without neglecting socio-cultural relationships. In fact, for improved Physics learning, an effective formative assessment carried out in conjunction with authentic assessment, as argued earlier on, can better be achieved if implemented taking into account different learning theories. Students learn Physics better when their prior knowledge is taken into account and when their ability to perform real-tasks is encouraged (Dekkers, 1997; Mutimucio, 1998); and this is also the reason why authentic performance assessment is crucial. The combination of formative and authentic assessments is very important because, while the performance of real-world tasks is supported by authentic assessment, the learning of necessary basic concepts and principles is merely dealt with by formative assessment.

3.4 Assessment strategies in Physics

The purpose of this section is to discuss arguments of several authors regarding different assessment strategies used to assess Physics learning in varied contexts. It begins with examining the principles of constructivism theory seen as central theory relating to Science education in general, and to Physics learning in particular. The importance of performance and authentic assessments in enhancing Physics learning is also discussed in this part. The second and third parts discuss the role of the teacher and the role of the students in assessment, respectively. The fourth part presents several arguments of the authors reviewed about different assessment strategies and the context in which they are used. Lessons drawn from this section are summarised in the fifth part.

3.4.1 Constructivist views of learning

Constructivism clarifies the views about the nature of human knowing, particularly the nature of scientific knowledge, as well as a view about learning processes and validation procedures of the acquired knowledge, and it is seen as a powerful theoretical resource that maximises student learning (Mutimucuo, 1998; Treagust et al., 1996). Both psychological and epistemological principles of constructivism emphasise that knowledge cannot be separated from knowing subject. The epistemological principle states that the function of cognition is adaptive and enables the student to construct viable explanations of experiences of the world. The psychological principle states that students do not passively receive knowledge but they actively build up their knowledge by a cognizing subject (Treagust et al., 1996). When the assessment of the knowledge is taken into account this process of building up knowledge becomes somehow problematic due to the different levels of activities (both manipulative and mental activities) that the students may be involved and to the different realities (realities accessible via sensory organs, via theoretical understanding and via instrumentation) where they are living in. Therefore, in assessment there is a need for consensus about levels of activities and different realities and for this process of knowledge construction, what the student already knows is of central importance. Consequently, knowledge about the world is seen as a human

construction, and this view of a student as an active agent, constructing his/her own reality, determines life processes and changes of all living beings (Mutimucio, 1998; Piaget, 1972).

Literature on Science education and the perception of the issues pertaining to assessment strategies in Physics were considered to expand the view on student difficulties when it comes to assess the learning. Because constructivism is seen as the theoretical framework on which most research into student thinking and learning is based, the discussion on assessment strategies on Physics, either in Africa or internationally, is based on this theory (Mutimucio, 1998; Treagust et al., 1996). International literature (Airasian, 2001; Moskal, 2003; Stiggins, 1987) emphasises the role of performance assessment as crucial in assessing Physics learning. But for students to be able to learn Physics better they should not only be asked to do performance tasks. They should primarily be able to understand basic concepts related to the subject. This type of knowledge can be and is mainly assessed through paper-and-pencil tests, but sometimes with other kinds of formative assessment. In fact, available research findings indicate that assessing students' basic skills seems not to be the problem with Mozambican secondary school teachers (INDE, 2005; Lauchande, 2001). According to this literature, paper-and-pencil tests and oral questioning have been used by these teachers with successful results at this respect. The problem is that students are not assessed on their ability to perform real-world tasks, i.e., their skills or proficiency in doing something. This is the rationale of choosing performance assessment for this study as one of the assessment strategies that can help teachers to improve student learning of Physics. This choice will be argued in Chapter 5 as a conclusion from the Baseline Survey. Performance assessments call upon the students to demonstrate specific skills and competencies, i.e., to apply the concepts and the knowledge they have acquired. The form of assessment in which students are asked to perform real-world tasks that demonstrate meaningful application of essential skills and knowledge is authentic performance assessment and it is closely related to the constructivist theory of learning. According to Muller (2006) one of the most critical features of an authentic assessment is that it usually includes a task for students to perform with a rubric against which their performance on the task will be evaluated.

On the other hand, Muller (2006) discusses the process of evaluating student performance through rubrics. The author argues that performance assessments are typically criterion-referenced measures where a student performance on a task is determined by matching the student performance against a set of criteria to determine the degree to which the student performance meets the criteria for the task. To measure student performance against a pre-determined set of criteria, a rubric, or scoring scale, is typically created which contains the essential criteria for the task and appropriate levels of performance for each criterion. A rubric is comprised of two components: criteria and levels of performance. For instance, in a task where a student is asked to assemble an electric circuit (manipulative activity), one of the criteria could be the student ability to obtain all the necessary equipment for the task and the other criterion could be the student ability to assemble the circuit properly. For both criteria, it can be defined whether the level of performance is poor or excellent. The rubric describes the task itself, i.e., the assembling of the electric circuit. Each rubric has at least two criteria of good performance and at least two levels of performance at which that performance was achieved. The criteria are the characteristics of good performance on a task. For each criterion, the teacher or the evaluator applying the rubric can determine to what degree the student has met the criterion, i.e., the level of performance. Table 3.2 shows the components of a rubric as well as the criteria and the levels of performance that can be used to assess it.

Table 3.2: Components of a rubric

| Rubric: Assembling an electric circuit | | |
|--|--|--|
| Criteria | Levels of performance | |
| Criterion 1: Student ability to obtain all the necessary equipment. | Poor (e.g., from 0 to 1): Basic equipment necessary for the electric circuit not found | Excellent (e.g., from 2 to 4): All equipment for the assembling of the electric circuit in place. |
| Criterion 2: Student ability to assemble the circuit properly. | Poor (e.g., from 0 to 1): Electric circuit not accurately assembled and some essential parts missing, leading to a malfunctioning of the circuit. | Excellent (e.g., from 2 to 4): Electric circuit accurately assembled and working properly. |

In between the levels of performance 'poor' and 'excellent' some other levels can be considered namely 'satisfactory' and 'good' with redistribution of points to fit in the scale of 0 to 4. Each criterion has its own scoring scale (e.g., from 0 to 4 points) and the scale is applied along the different levels of performance with a previously established maximum score (e.g., 20 points). Hence, a rubric appears to be a scoring scale used to assess student performance in terms of a task-specific set of criteria.

In conclusion, both the psychological and the epistemological principles of constructivism emphasise that students need to be actively involved in building up their knowledge. In this process of knowledge construction, the discussion on assessment strategies taking place in Africa and internationally, emphasises the role of performance assessment as being crucial for assessing Science subjects. As was referred to earlier on in this subsection, performance assessment alone cannot help students learn Physics. Other types of assessment strategies - for instance, paper-and-pencil tests, verbal tests, and portfolios - are also important and are needed to assist students to learn the necessary knowledge and skills in order to do a performance assessment. This study, however, focuses on the use of performance assessment because the Baseline Survey findings (see Chapter 5) have indicated that teachers are already achieving some success by using those other types of assessment leading to the successful completion of a performance assessment. Assessment with these characteristics will be argued in Chapter 5 as a conclusion from the Baseline Survey. Performance assessment deals with the role of students on demonstrating their skills and competencies, as well as the role of the teachers in monitoring the learning process and in evaluating the levels of such performances.

The following subsections addresses firstly, the role of the teacher in assessment and secondly, that of the students. Thirdly, arguments of various authors about different assessment strategies applied to different contexts are presented and discussed as platform for designing the improved assessment strategy advocated by this study.

3.4.2 *The role of the teacher in assessment*

Before elaborating any further, it is important to explain the need to improve the current assessment practice in schools. Effectively, assessment remains the weakest aspect of teaching and learning in most subjects, including Physics. Nearly all school assessment policies have weaknesses, which are reflected in corresponding weaknesses in the assessment practice of teachers. For example, Weeden et al., (2002) argue that, in many classrooms, the issue is not that teachers are not assessing enough, but that they are not using the information they collect to help students learn.

The problem being pointed out by the present study is linked to the collection of learning evidence and it is twofold: (i) that Grade 12 Physics teachers have been using limited types of assessment (mostly paper-and-pencil tests); and (ii) that also with these limited types, the collected evidence has only been used for accountability, and not for promoting learning. A strong argument of this study is that any support for the role that a teacher can play in the classroom should be directed towards trying to find a solution to this problem.

Race et al., (2005) on the problem of the limited types of assessment used by teachers, for instance, discuss the importance of putting assessment into context and they stress the need for teachers to consider several aspects of assessment. These aspects include knowing why to assess, what to assess, and what the quality is of feedback they should provide to their students. Regarding the issue of *why to assess*, these authors argue that teachers might be supported to understand, amongst other aspects, that assessment is carried out to guide student improvement and help them to learn from their mistakes (formative assessment). In relation to *what to assess*, teachers' role might be to assess the 'process' of how the students are achieving the learning outcomes and in a holistic approach rather than the 'product' (outcome) itself. As for the *feedback quality*, the authors emphasise the need for a timely, personal, expressed and empowering feedback for student learning. The authors' argument is that the use of a variety of assessment types will support and inform student learning. Contextualised learning is in line with a

constructivist approach, but one cannot expect to assess different learning skills with limited assessment types and taking into account different student backgrounds.

In relation to the collection of evidence, Harlen (2006) elaborates on how teachers must be encouraged to collect student learning evidence as a normal part of class work either in an informal or formal formative assessment. In particular, the author highlights the importance of criterion-referenced assessment through which the student's achievement is described in terms of what s/he can effectively do, as opposed to norm-referenced assessment that is based on the ranking of students in order of their achievement. Student- and criterion-referenced assessment must be the basis for judging the evidence, the feedback must be judged and used by both students and teachers, and the assessment in general should be directed *for* learning. The fact is that, according to the available Mozambican literature (INDE, 2005; Lauchande, 2001), current teacher assessment practice in Science subjects has a more formal summative character whereby the collection of evidence is done as a separate task or test. The basis of judgment is criterion-referenced and the assessment is generally an assessment *of* learning.

3.4.3 The role of students in assessment

The arguments of several authors in Science education (Dekkers, 1997; Harlen, 2006; Race et al., 2005) suggest that an effective assessment for learning strategies depends, among other aspects, on active student involvement, their ability to assess their colleagues and themselves, and on the profound influence assessment has on the motivation and esteem of the students.

Dekkers (1997), for instance, argues that research on student knowledge of the world requires a basis of scientific knowledge that the teacher and students share, as well as effective communication between them. The author is of the opinion that, for the establishment of such a base of knowledge, it is important to start with establishing which knowledge is shared. If there is no such basis, no mutual understanding can develop.

On the issue of student involvement, Race et al. (2005) argue that the provision of feedback either to individual or to groups of students helps to improve their active participation in the learning process. The authors also emphasise that, if a teacher increases the students' participation, he/she must allow them to interrogate and challenge his/her comments. In subjects like Physics, this is crucial because interrogation and the expression of the student thoughts help to mediate their reasoning. Peer-assessment is another factor regarded as relevant by these authors for successful assessment. They argue that students learn more intensely when they have a sense of ownership of the agenda, and by assessing their peers, they learn from each other's successes and weaknesses. Providing or negotiating assessment criteria, gradually introducing peer-assessment, and making peer-assessment marks meaningful (i.e., making the marks count) are, according to the same authors, some of the aspects regarded as being useful for a successful peer-assessment. In fact, meaningfulness of assessment marks is also crucial in increasing students' motivation. On this respect, Harlen (2006) claims that most of the roles that students can play in assessment in particular, and in learning in general, have much to do with their motivation. It represents the construct that impels students to spend the time and effort needed for solving problems and for learning. The author also argues that students do not only gain motivation as an input from education, but it is also an outcome if they are able to adapt to the world of changing conditions that occur beyond formal schooling. When such changes occur rapidly, the motivation of students to learn new skills will be stronger and their enjoyment of encountering new challenges will be greater. Consequently, assessment is seen as one of the key factors that affect student motivation. Still according to Harlen, some authors such as Stiggins (2001), claim that teachers can enhance or inhibit student desire to learn more quickly through their use of assessment than through any other instructional means they can use.

Students' self-esteem is another important factor in learning and assessment (Race et al., 2005). It is defined as the way people value themselves both as people and as students, and shows the confidence that the person feels in being able to learn. This means that any role to be played by students in assessment strongly depends on the level of their own self-

esteem. Those students who are confident about their ability to learn will approach any assessment task with an expectation of success and a determination to overcome problems.

3.4.4 Assessment strategies and the context

Airasian (2001) has pointed out that teachers normally use three main strategies to gather their assessment information namely, observation, oral questioning, and paper-and-pencil tests. According to this author, the paper-and-pencil methods is the most important method teachers use to collect assessment information and they are of two general types of methods, namely, selection and supply. In the selection type, students respond to each question by selecting an answer from the choices provided. In supply or constructed response type, the student produces a response to a question or task. In the selection type, the advantage is that it provides the maximum degree of control for the question writer, in a supply-type item, the question writer only has control over the question itself, since the responsibility for constructing the answer resides with the supplier.

The observation method is one of the major strategy teachers use to collect assessment data about students, instruction, and learning. It involves watching or listening to students carry out some activity, or judging a product a student has produced. For example, when students submit a Science project or set up laboratory experiment, the teachers also observe and judge what the students have produced. Both planned and unplanned observations have the advantage of allowing teachers to observe a particular of student behaviour which is thus considered as important information gathering techniques in classrooms.

The oral questioning is another method mostly used by teachers not only to collect assessment information but also for guiding instruction. It can be used to review a previously taught topic, brainstorm a new topic, find out how well the lesson is being understood by the students, and to gain the attention of the disturbed students. The advantage of this strategy is that it allows the teacher to gather information related to

assessment without the intrusiveness of administering paper-and-paper assessments. Formal oral examinations, for instance, are used in subject areas such as foreign language, singing, and speech. Oral questioning techniques are also seen as vital to complement all other information gathering strategies.

Other authors discuss many other assessment strategies. With the portfolios strategy, for instance, it is worth considering arguments by Kemp and Toperoff (1998). These authors define portfolios as collections of student work representing a selection of performance. Portfolios may be a folder containing a student's best pieces, and the student's evaluation of the strengths and weaknesses of the pieces. It may also contain one or more works-in-progress that illustrate the creation of a product, such as an essay, evolving through various stages of conception, drafting, and revision. According to Kemp and Toperoff (1998), recent changes in education policy in the United States, which emphasise greater teacher involvement in designing curriculum and assessing students, have been an impetus to increased portfolio use in schools. They are valued as an assessment tool because, as representations of classroom-based performance, they can be fully integrated into the curriculum. And unlike separate tests, they supplement, rather than take time away from instruction. Moreover, many teachers, educators and researchers believe that portfolio assessments are more effective than 'old-style' tests for measuring academic skills and informing instructional decisions. Popham (2002) distinguishes some advantages and disadvantages of portfolios. Portfolio assessments are difficult to evaluate because they are tailored to individual student's needs, interests, and abilities and they take time to carry out properly. On the other hand, however, they are a way of documenting and evaluating growth and allow student self-evaluation and personal ownership.

Popham (2002) also supports performance assessment. He points out that many teachers consider short-answer and essay tests a form of performance assessment, which means that they equate this kind of assessment strategy with any form of constructed-response assessment. Other authors (Airasian, 2000; Moskal, 2003) contend that genuine performance assessments must have at least three characteristics. These are: multiple evaluative criteria, in which the student performance is judged using more than one

evaluative criterion; prespecified quality standards, where each of the evaluative criteria on which a student performance is to be judged, are explicated before judging the quality of the performance; and judgmental appraisal, where genuine performance assessments depend on human judgements to determine how acceptable a student performance really is. For whatever reasons, many advocates of performance assessment prefer that the student tasks should represent real-world rather than school-world situations (Airasian, 2001; Moskal, 2003; Stiggins, 1987; Wiggins, 1993). Authentic assessment and alternative assessment are phrases used by some authors (McMillan, 2001; Meyer, 1992; Stiggins, 1987) to describe performance assessment. Authentic assessment is so considered because the assessment tasks are more closely linked to real-life and not to school life, while in alternative assessments the tasks are alternative to those of traditional paper-and-pencil tests.

3.4.5 When reading from this section

From the several arguments presented and discussed in the above section two major lessons emerge as relevant for addressing the main research question of the present study:

Firstly, that a number of assessment strategies can be used to assess Physics learning and the constructivist theory of learning appears to support some of them. Depending on the assessment objectives being pursued (supply answers or constructed elaborated responses, learning by doing, guiding instruction, evaluating student work) and on the context in which the assessment takes place (normal classroom situation, laboratory setting, out-of-school environment) one can make adequate decisions on which learning theory better suit which assessment strategy. For this particular study, constructivism is appropriate to guide students to learn in a laboratory setting with the aim of achieving the goal of learning by doing. In fact, the Baseline Survey of this study was carried out taking into account both the assessment objectives sought by the teachers and the context in which these teachers were assessing their students. The constructivism theory, seen as central for Physics learning, was influential in this study during the design and development of the exemplary assessment materials (see Chapter 4).

A second lesson from the literature reviewed is that the roles of both teachers and students are crucial for the success of any assessment strategy. It is important for the teachers to know why to assess, what to assess, and what the quality is of feedback they should provide to their students. A criterion-referenced student assessment must be the basis for judging the learning evidence, the feedback to students must be judged and used by both students and teachers, and the assessment in general should be directed *for* learning.

For the students' perspective, the literature suggests that an effective assessment depends mainly on their active involvement, their ability to assess their colleagues and themselves, and on their motivation and self esteem. Therefore, any improvement of assessment practices can only be effective if it includes those assessment practices or strategies that emphasise formative approaches as a way of improving student learning. This study then, addressed these aspects during the Intervention Study through an instructional strategy of students' knowledge construction named Predict-Observe-Explain (see Chapter 4, Section 4.3, under design guidelines for the intervention).

Having presented the arguments of various authors on the topic under investigation, the following section reviews the literature on intervention studies conducted within an African context in the field of Science education.

3.5 Some intervention studies in science education

The content of this section is presented through two perspectives, namely the findings of the reviewed intervention studies, and the methodological and substantive implications of such an approach for the present study.

During the review of the literature, some writings related to research on interventions were sourced. Most of this literature was in the form of PhD theses written within an African context and emphasising an educational design research approach as the most suitable for intervention studies. The writings are rooted in the field of assessment in Science

education, and are particularly related to probing students' understanding of Science using this approach. From all the authors reviewed in this field, the findings from research by Mafumiko (2006), Motswiri (2004), Ottevanger (2001) and Tecle (2006) are relevant to the research reported in this thesis.

For example, the aim of Mafumiko's study - *Micro-scale experimentation as a catalyst for improving the chemistry curriculum in Tanzania* - was to investigate the possible use of a low-cost approach to practical work that could contribute to improving the teaching and learning of chemistry in Tanzanian secondary schools. The study focused on designing and evaluating an intervention of micro-scale experiments to support curriculum materials at this level. The main research question addressed by this study was 'what are the characteristics of micro-scale chemistry materials that contribute to the initial implementation of practical work in chemistry education in Tanzanian secondary schools?' The key findings were that (i) teachers and students were able to implement most of the lesson activities according to advice provided in the curriculum materials (classroom implementation); (ii) teachers regarded having access to the support materials in advance as very helpful for preparation of the lessons in general (opinions about the study approach); (iii) teachers considered micro-scale experiments as a useful way of conducting practical work because it enabled them to involve a large number of students with minimum resources (opinions about conducting practical work); (iv) students involved in the micro-scale experiments found them helpful in enhancing their learning of chemistry, made the subject enjoyable and, hence, increased students' participation in the lessons (opinions about the approach); (v) students found their involvement in chemistry micro-scale experiments as increasing their confidence in doing more experiments as well as their awareness on safety and environment (opinions about learning of chemistry). In general, these findings indicated that for the teachers, the materials provided adequate support information during the preparation of students' practical work with the development approach which needed less time, and less sophisticated equipment.

Tecle (2006) in *The potential of a professional development scenario for supporting biology teachers in Eritrea* addressed the question of 'what are the characteristics of a

professional development scenario that effectively supports biology teachers in Eritrea implementing a more student-centred approach?’ Similar to Mafumiko’s study, this study adopted an educational design research approach to guide the analysis, design, evaluation, and revision processes of the professional development scenario (intervention). This study found that teachers regarded prototyping of professional development scenario as being important and useful in providing them support on subject matter knowledge, lesson organisation, using concepts maps, and handling group activities. However, in some cases teachers were observed encountering problems with group work activities and throughout the tryouts, the issue of time continued to be problematic. Teachers needed more time, particularly in drawing conclusions from the activities. In general, teachers appreciated the summative evaluation workshop because it provided them with exemplary materials, a forum for active discussion, the opportunity to observe exemplary practice, and a learning environment for practicing and augmenting the skills for teaching practically-oriented biology lessons.

Motswiri (2004) conducted an investigation on *Supporting chemistry teachers in implementing formative assessment of investigative practical work in Botswana* and addressed the research question of how can exemplary curriculum materials support senior secondary chemistry teachers in Botswana with the implementation of formative assessment of students’ investigative practical work. This study also followed an educational research approach where a prototyping process was used for an orientation study aimed at (i) articulating initial design specifications for the envisaged exemplary materials, (ii) developing and trying out several versions of prototypes, and (iii) field testing the final version. Findings from this study indicated that teachers were critical about the congruence of the exemplary materials (the intended practice appeared to be incongruent with the teachers’ current practice), but they were positive about their clarity (the materials were regarded understandable) and their cost (the suggested implementation was possible within the limitations of available resources in the science laboratories). However, the teachers were not often observed to demonstrate formative assessment orientations in terms of asking questions related to helping students to reflect critically on results they expected, activities they carried out, and results they obtained. They seemed to

be in need for more support in terms of formative assessments with particular emphasis on time management. Their use of practical work was more frequent in the body of the lesson (where students were helped with probes to participate in planning and experimenting) than in the lesson introduction and lesson conclusion. Students, in contrast, enjoyed the lessons, especially during lesson introduction and learned subject-specific knowledge in terms of investigation procedures.

Ottevanger (2001) on *Teacher support materials as a catalyst for science curriculum implementation in Namibia* addressed the question of what are the characteristics of materials that adequately support teachers in the initial implementation of Science curriculum innovation in the classroom. With the same research approach used in the other studies, Ottevanger's study found that (i) from the scientific process point of view the teacher support materials have led to well-organised lessons in the majority of cases and were useful as a resource in offering extra information on the topic of the lesson. This was seen by the author as a positive step forward in the Namibian context. (ii) The connection between the specific experiment and the relevant theory needs to be further strengthened. (iii) Students' involvement in the lessons increased during the lessons supported by the developed materials. (iv) Students indicated that they liked using materials from local context, doing group work and cooperating with other students. They also referred to the fact that their teacher appeared to be better prepared than in their usual classes. (v) Although teachers seemed to address the time issue in their own ways, this appears to be a continuous problem in completing lessons. In conclusion, this author claims that in the context they were used, teacher support materials containing procedural specifications have shown themselves able to act as a catalyst in the initial implementation of the new curriculum in the classrooms.

The relevance of all these interventions for the present study can be described from both methodological and substantive perspectives. Although the focus of all these studies was on improving student-centred learning, and not particularly on assessment strategies, their methodology can still be successfully applied to the present study. There are similarities between the present study and the literature discussed that support this argument. Firstly,

all studies focused on design and formative evaluation of exemplary materials where the search for characteristics of an effective intervention was conducted while teachers and students were working on that intervention. Secondly, there was a decision to focus on a certain topic or theme to concentrate on. Thirdly, the tasks carried out either by students or by teachers were designed in a standardised manner. Fourthly, the methodology included anticipation of potential implementation problems through application of a systematic process of formative evaluation of the products. Fifthly, the support materials for teachers were designed in such a way that they provided help at four support levels, namely, subject knowledge, lesson preparation, teaching methodology, and assessment and feedback. All these aspects are concurrent and characterise the methodology of research on interventions.

Substantively, from all the contributions and arguments presented in the reviewed studies, the emphasis is on the importance of developing teacher support materials and on the design and tryout of authentic material in a classroom environment. With regard to the importance of the materials, most of the teachers considered such materials as very useful for their lessons, they could be used as broad guides for future lesson preparations, and they represented an opportunity for them to engage in a learning process while working in their own environment. In relation to the design and tryout of the materials, it is worthy mentioning a number of practical aspects that arose during the processes which needed to be carefully monitored. These aspects include: the role of the teacher in guiding student activities; the role of students as group workers; the time involved in discussions; and the overall monitoring of both teacher and student behaviour as compared to the normal classes. All of these aspects relate to what Black and William (2006) and the Assessment Reform Group (ARG) (1999) from the UK refer to as effective assessment *for* learning, as opposed to assessment *of* learning. ARG (1999) argues that the heart of learning evidence lies in the power of formative assessment and that any feedback for students is only effective if used to guide improvement. In addition, effective assessment for learning depends on effective feedback to students, on their active involvement, and on the adjustment of teaching to the results of student assessment. Harlen (2006) discusses assessment *for* learning as a cycle of events with the students in the centre of it. The cycle

starts with the goals or objectives of the assessment task through which the teachers intend to collect evidence. Then, in possession of enough evidence, it follows an interpretation process aimed at judging the students' achievement so that decisions about next steps can be properly taken. Finally, teachers decide about how to take next steps related to student activities in the learning process, which in turn are directed towards the assessment goals. In fact, one of the issues that the present study wishes to address is related to limited assessment strategies used by Mozambican teachers to collect evidence of learning in schools. The point is that, seemingly, teachers do not only collect insufficient evidence for learning, but also the little evidence being collected, is of poor quality. It seems that teachers apply assessment strategies that allow them to obtain learning evidence only of basic knowledge and skills and even this evidence is not used for improving student learning. The interpretation of the evidence to judge students' achievement is only criterion-referenced, and the ultimate assessment goal is to report on students' achievement. So, one may conclude that the assessment process of Mozambican teachers does not represent Harlen's complete cycle of assessment events and this leads to an ineffective assessment for learning.

Having reviewed some of the intervention studies carried out in the African context and presented the lessons that can be derived from these studies, the next section provides a platform on how these lessons can be used to conceptualise the study and to guide the formulation of preliminary operational research questions of this study.

3.6 Summary and conceptualisation of the study

The topic of this study is to investigate assessment practices used by Grade 12 teachers in Physics in Mozambique and, if needed, to develop an intervention aimed at improving the quality of classroom assessment. Where the literature was reviewed from various angles, the findings were summarised from the perspective of the research questions. This section summarises what was learnt from the reviewed literature as a whole and provides direction about the conceptualisation of the study.

To begin with, one of the most relevant theories in students' knowledge construction is constructivism (Mutimucuo, 1998; Treagust et al., 1996). It was argued that it represents a powerful theoretical resource that may maximise student learning. In fact both the current Mozambican Grade 11 and 12 syllabuses for Physics, and the secondary school curriculum under review, acknowledge and recommend its utilisation. The problem, however, is that while it is recommended for the students in schools, training institutions are still educating teachers within the paradigm of behaviourism. The challenge of this study is to improve assessment practices of teachers and ultimately to help them implement the recommended curriculum.

In the second place, depending on the assessment objectives to be achieved and the context in which the assessment takes place, student achievement in Physics learning should be assessed using different assessment strategies and in varied learning contexts. Therefore, in the process of investigating assessment practices being used by secondary school teachers, there is a need to be constantly alert to what the teacher actually needs to achieve taking into account the conditions in which she or he is working.

In the third place, there is the crucial role played by both teachers and students in assessment. Teachers must know why to assess, what to assess, and understand the importance of quality feedback which they provide to their students. Successful assessment practices take place with active involvement of the students, their ability to assess their peers and themselves, and on their motivation and self esteem. This is likely to occur when using those varied assessment practices that emphasise formative approaches.

In the fourth place, one of the most successful assessment practices in Science education is performance assessment because of its crucial role in assessing student's day-to-day activities. This type of assessment calls upon the students to demonstrate specific skills and competencies and requires them to perform real-world tasks that demonstrate meaningful application of essential skills and knowledge. So, the improvement of teacher assessment practices sought by this study, as will be argued in Chapter 5 (Section 5.3), implies taking into consideration the importance of performance assessment without,

however, neglecting the role played by all other different assessment strategies described in subsection 3.4.4.

In the fifth place, the reviewed literature has put an emphasis on assessment *for* learning, where the results are used to inform the teaching and learning process, as opposed to assessment *of* learning, which is mainly for grading and certification. In undertaking assessment *for* learning teachers must consider completing the entire cycle of assessment events if it is to enhance learning. This cycle of assessment should:

1. Determine the goals of learning and therefore of assessment.
2. Collect enough evidence of learning.
3. Judge whether students' achievement is sufficient.
4. Decide on the next steps in the process of learning and teaching.

Harlen (2006) in this respect, points out that if assessment is to be effective for learning, an entire cycle of goals-evidence-judgment of achievement-next steps in learning-goals has to be completed. The teacher must collect evidence related to goals; interpret the evidence in order to judge the student's achievement; use achievement data to influence decisions about the next steps in learning geared towards the goals. However, and according to some literature reviewed (see, for instance, INDE, 2005; Lauchande, 2001), it seems that Mozambican teachers do not follow the complete process of conducting an effective assessment to inform learning. Teachers do not seem to collect enough evidence of learning - due in part to the use of limited assessment strategies - and they do not use the information they collect to help students learn either. Therefore, the teachers do not complete the cycle of events that might characterise an effective assessment for learning.

Arguments from literature indicate that teachers must be helped to put assessment into context by considering aspects such as why to assess, what to assess, what quality of feedback they should provide to their students, and the curriculum perspective. In this respect, criterion-referenced assessment must be the basis for judging the student performance, the feedback must be judged and used by both students and teachers, and the assessment in general should be directed *for* learning. From this lesson, it is suggested that

this study addresses the issue of ‘collecting evidence for learning’ and its interpretation for judging the student achievement. Furthermore, formative assessment should be considered as the heart of learning evidence and, as supported by ARG (1999), feedback for students will only be effective if it is used to guide improvement.

A concluding remark is that, although all these authors stress the importance of using the collected evidence to take decisions about the next steps in learning, the literature review has shown that somehow there is neither sufficient research of the extent to which assessment strategies are being used for Physics as a subject, nor any reported professional support for teachers to assist them in the development of performance assessment material for use in an ordinary classroom environment. This means that the main question posed by this study remains unanswered in the review of the literature.

This study addresses these shortcomings by contextualising the problem with a focus on secondary school Physics for Grade 12. Apart from the constructivist approach, three other lessons learnt from the literature review can be highlighted as far as the characteristics of materials are concerned. Firstly, there is the need to help teachers by developing and letting them use exemplary support materials on performance assessments that can help students construct their knowledge. Exemplary materials of this nature should help teachers in several aspects of subject knowledge, lesson preparation, teaching methodology, and assessment and feedback. These characteristics should guide them for future lesson preparations, provide them with the opportunity to learn while working, and help them to learn how to develop the materials for topics other than the ones selected for this study. The design and tryout of the materials aspects such as the role of the teacher, the role of students, the class management, and the overall monitoring of student behaviour during lessons, are also aspects to feature in the materials. Secondly, the development of such materials should be done in an ordinary classroom environment to allow users to participate in the process while working in their normal routine. Thirdly the learning evidence should be used to feed the teaching and learning process and, hence, formative assessment is crucial.

In this study the following main research question was examined by the intervention:

How can the teacher assessment practices be improved?

In order to address this question adequately, and drawing from the literature, it is of paramount importance to design an intervention that builds upon teachers' present knowledge, skills and experiences with formative assessment in the classroom. As referred to in previous chapters, this implies a prior investigation using a survey approach and aimed at knowing what assessment practices Grade 12 teachers in Physics in Mozambique apply. Some operational research questions are formulated for the Baseline Survey, and are listed below.

- *What assessments practices do Grade 12 teachers apply?*
- *What can be said about the quality of the assessment practices?*
- *How relevant are the assessment practices for student learning?*

Although these research questions generate valuable baseline knowledge about the actual classroom assessment, this knowledge is descriptive by nature and does not provide indications as how to improve the teacher assessment practices as implied by the main research question. The improvement of teacher assessment practices is achieved through working together with teachers in producing and using assessment materials. Briefly, and as was referred to in Chapter 1 (Section 1.1), the main question is twofold, i.e., it implies knowing firstly, what assessment practices Grade 12 teachers apply (Baseline Survey) and secondly how to improve them (Intervention Study).

All baseline and intervention research questions are described in detail in Chapter 4 (Research Design and Methods). At this point, it is relevant to capture what the reviewed literature says about how to improve teacher assessment practices, and what are the most common assessment practices used to assess Science learning.

From the several assessment practices recommended by the literature and from which an intervention for improvement can be conducted, a choice is made concerning performance type of assessment as the focus of this Intervention Study. As was referred to earlier, the rationale for this choice is that, amongst all mentioned assessment strategies, performance assessments appear to be of vital importance in assessing the student understanding of key physical concepts. Performance assessment can also be seen as an adequate means of improving teacher practices of assessing physical and related skills, because all schools expect students to demonstrate a number of skills, from simple communications skills like reading, writing, and speaking, to more complex psychomotor skills like building a car-toy or setting up laboratory equipment. In spite of the importance of all these student skills, when it comes to Science subjects like Physics, the students must not only be able to grasp the concept or the process, but also to explain and use it to solve real-life problems. For example, after students have learnt to identify the direction of power in one electric circuit (e.g., via multiple-choice tests), they must be able to go through the process of identifying, by themselves, some other unknown directions of electric circuits given to them. This kind of hands-on demonstrations of concept mastery is essential in Physics.

In this regard, Airasian (2000) argues that there has indeed been growing emphasis on using performance assessment to determine student understanding of the concepts they are taught and to measure their ability to apply procedural knowledge. Gronlund (1998) also emphasises the role of performance assessment in providing a systematic way of evaluating reasoning and skill outcomes. These outcomes are important, for instance, for Physics because the subject is concerned with solving problems and developing laboratory skills. Moreover, the current educational trend to shift from norm-referenced assessment (ranking of students in order of achievement) to criterion-referenced assessment (description of what students can do) has created a need for a more direct assessment of how well students can perform. Therefore, it is important for this study to allow students to demonstrate, through performance assessment, their ability to do real-world tasks while observing all the procedures involved. It is also relevant to emphasise that, while all other assessment strategies can successfully be conducted in a classroom environment or as homework, effective performance assessment is most likely to succeed when: (i) it is

undertaken in a laboratory context where students can perform real demonstration experiments; (ii) a set of procedural steps is followed – ranging from specifying clear performance outcomes to selecting a proper method of observing, recording and scoring; and (iii) a systematic method of combining them with traditional tests is used. The specific use of any of the above-mentioned assessment strategies depends on the specific learning outcomes to be achieved. This means that to select an adequate assessment strategy, a number of intended learning outcomes must be prespecified. Effectively, for this study the performance outcomes have been identified as the need to demonstrate and develop explanations about force and inertia. Each of these outcomes corresponds to certain areas of performance being assessed. Performance outcomes commonly use verbs such as ‘identify’, ‘construct’, ‘demonstrate’ or appropriate synonyms.

In relation to what aspects of teacher assessment can be taken into account in order to know what assessment practices Grade 12 teachers actually do apply in a contextualised environment, the literature emphasises (as already has been concluded), amongst other formative assessment practices, the importance of performance assessment. However, it is also a lesson from the literature that a pre-requisite for students to be able to learn Physics better, is the prior understanding of the basic concepts related to the subject and, according to the literature, this can normally be assessed using mainly paper-and-pencil tests. Among the variety of other recommended formative assessment practices are observation methods, oral questioning, peer-assessment and portfolios.

These and other assessment practices are examined by the Baseline Survey reported in Chapter 5, while Chapter 4 describes procedural steps, approach, learning outcomes, and performance areas of the Intervention Study. Specifically, Chapter 4 presents the research design of the study (as a rationale for having two phases); the operational research questions of each component (following from this preliminary formulation); the research paradigm; and all methodological aspects of the two phases.

CHAPTER 4

RESEARCH DESIGN AND METHODS

CHAPTER 4

RESEARCH DESIGN AND METHODS

This chapter introduces the research design and methods of the study of investigating and improving assessment practices of Grade 12 Physics teachers in Mozambique. Section 4.1 presents the rationale for having two phases, the research approach for each phase, the general formulation of the research questions addressed by the phases and some reflections on research methodology. Section 4.2 discusses the research paradigm, which was chosen from the research questions addressed in the previous subsection. Section 4.3 elaborates on the research design of the study. The section starts by presenting the research design of the Baseline Survey (the first phase of the study). The research questions addressed by this phase of the study, population and sampling, data collection strategies, and data processing and analysis methods are also presented in this part. The section also discusses the research design of the Intervention Study (second phase), the educational design research as the approach followed in this phase, and the guidelines for designing the intervention. Section 4.4 presents arguments about the validity and reliability of the study while ethical issues are discussed in Section 4.5. Finally, Section 4.6 presents the conclusion and provides an orientation for the following chapter.

4.1 Introduction

The purpose of this study was to investigate the assessment practices of Grade 12 Physics teachers in Mozambique and how these practices can be improved. The rationale for this study, as discussed in Chapter 1, was that the quality of Physics learning demonstrated by students leaving secondary school is poor and there are reasons for believing that inadequate assessment practices are one of the main contributory reasons for this. As was referred to in Chapter 2, the problem was perceived as a problem at school level. Therefore, it was essential to have a good understanding of the present assessment

practices carried out by secondary school teachers in schools and classrooms in order to design an effective 'intervention in assessment'. The context in schools can be characterised by various influences from different educational and social entities (see Chapter 2, Section 2.3). In order to gather relevant information pertaining to assessment practices in such a diversified target population, a study by means of a variety of data collection strategies had to be undertaken so that findings can reflect the characteristics of the wider population. This implied that this research should have a preliminary Baseline Survey to develop a good understanding and insight prior to the Intervention Study aimed at designing an intervention that included developing Physics assessment prototypes for teachers to use in their classrooms to optimise the teaching and learning of Grade 12 Physics in the classroom.

The Baseline Survey focused on the identification of assessment practices currently used by Grade 12 teachers in Mozambican schools and their knowledge and skills on assessing students. Teacher knowledge and skills are addressed by investigating the quality and relevance of the classroom assessments. The main research question for the Baseline Survey was formulated as follows: *What assessment practices do Grade 12 teachers in Physics in Mozambique apply and what is their quality?* This research question is in line with the aim of the study which is divided into three specific research topics namely (i) the types of assessment practices (diagnostic, formative, summative) currently in use by Grade 12 Physics teachers in schools, (ii) the quality of these practices and (iii) their relevance for classroom practice. Therefore, the research question is also operationalised into three operational research questions, which are:

1. *What assessments practices do Grade 12 teachers apply?*
2. *What can be said about the quality of the assessment practices?*
3. *How relevant are the assessment practices for student learning?*

The design of the survey is based on the context in which Physics teachers are working in schools as well as on the insights of what the literature highlights as good practice in classroom assessment. More generally, the Baseline Survey lays down the groundwork for

the Intervention Study. Based on the assumption that the assessment practices will help teachers to develop abilities to monitor improvements in student learning and in the performance of the educational system, questions about the types of assessment practices, their quality, and their relevance for classroom practice were included in the Baseline Survey. This assumption was supported by views from the literature reviewed in Chapter 3 (subsection 3.4.4), other authors (Black et al., 2003; Popham, 2002; Weeden et al., 2002), and also from the Science education experts in general.

In light of the findings from the Baseline Survey, the main research question of the Intervention Study – which became the main research question of the study - is: *How can teacher assessment practices be improved?* The process of reviewing the literature on the importance of improving teacher assessment practices (Chapter 3, Section 3.5) has emphasised the need to support teachers both in terms of conducting effective assessment *for learning*, where the assessment results are used to enhance the teaching and learning, as well as the development of authentic assessment material in a classroom environment. According to van den Akker (1999), an evolutionary prototyping of curricular or assessment products and their subsequent representations in practice are viewed as more productive than linear development approaches. Formative evaluations of subsequent assessment versions are essential to such productiveness, and an educational design research approach is seen to enhance knowledge growth. Therefore this intervention, as discussed in Chapter 3 (Section 3.6), focuses on (i) the design and formative evaluation of exemplary performance assessment materials for demonstration experiments aimed at assisting teachers to improve their assessment practices and on (ii) a laboratory written report by students. A great emphasis has been put on lesson materials rather than on assessment materials throughout the intervention. This is because any assessment strategy can only be successful if it is applied using quality lesson materials. Good lesson materials provide adequate support information for the preparation of student assessments, they are broad guides for future lesson preparations (including assessment tasks), and teachers and students implement most of their lesson activities according to advice provided in the lesson materials.

The demonstration experiments and the students' report are designed to focus on only two Physics concepts namely force and inertia, whereas the intervention addresses the functions of assessment namely diagnostic, formative and summative assessment. The reasons for selecting these two Physics concepts are given in subsection 4.4.3. As stated earlier in Chapter 3 (Section 3.1) 'demonstration experiments' refers to students' activities of observing, carrying out small experiments, interpreting phenomena, and reporting findings, and are all guided by a teacher. Demonstration experiments may be performed either individually or in small groups of students and they take a few minutes to perform. When these experiments take longer to perform (from 30 min to hours), the literature refers to them as laboratory demonstration experiments. In this study, a decision was made to use the term 'demonstration experiments' because of the characteristics described above.

The intervention applies the methodological approach of educational design research suggested by van den Akker and Plomp (1993). The potential of educational design research is that the search for characteristics of an effective intervention is conducted while working on that intervention. The research approach is discussed in Section 4.3 while the research paradigm, considered suitable for addressing the research questions of the study, is discussed in the next section, Section 4.2.

4.2 Research paradigm

Several authors have argued that to choose the type of knowledge claim, the researchers have to adapt certain assumptions about what and how they will learn during their inquiry (Creswell, 2003; Lincoln & Guba, 2000; Mertens, 1998). According to these authors, this claim can be named ontology, epistemology, philosophical assumption, or paradigm. Philosophically, the researcher makes claims about the nature of the reality, i.e., what is knowable (ontology), what is the relationship between the researcher and the researched (epistemology), the language of the research (rhetoric), and what the process of studying the reality (methodology) will be.

According to Creswell (2003), there are four schools of thought about knowledge claims namely post-positivism, constructivism, advocacy, and pragmatism. Post-positivist researchers claim that causes probably determine effects. These researchers challenge the traditional notion of the absolute truth of knowledge and they claim that when studying the behaviour of human beings one cannot be positive about the claims of knowledge. Constructivists often address the process of interaction among individuals and focus on the specific contexts in which they live and work in order to understand their historical and cultural settings. Advocacy researchers believe that inquiry needs to be intertwined with politics, and the research should contain an action agenda that may change the lives of the researched, the institutions where they work, and the researcher's life. Pragmatists claim that knowledge arises out of actions, situations and consequences rather than antecedent conditions, and their main concern is with applications and solutions to problems.

Within this framework, the scientific position of this study (as referred to in Chapter 1, Section 1.3) is rooted in the pragmatic knowledge claim. The research process did not begin with any one system of reality to identify the type of research method to be applied; rather, it started by identifying the problems to be solved, i.e., from the research questions formulated and went on to identify the suitable research methods that were relevant in obtaining valid and reliable answers to these questions. The research was geared towards the best understanding of the research problem. The truth was not solely based on dualism between the researcher's mind and reality, but it was on what worked at the time. Both qualitative and quantitative methods were applied to collect and analyse data with the main aim of understanding the complexities of the current situation and to produce findings that contribute to a solution to the problem.

4.3 Overview of the research design

In line with the pragmatist claim described above, this study intended firstly, to find out what classroom assessment practices teachers are using in schools before choosing any particular type of assessment to monitor improvements and secondly, to generate a methodological approach and guidelines for the design and development of an adequate

study approach aimed at improving such practices. The most suitable strategy to identify current assessment practices is the survey approach, where a study was conducted using questionnaires, observations, and interviews for collecting data. As several authors quoted in Chapter 3 have argued (Airasian, 2001; Dekkers, 1997; Moskal, 2003; Stiggins, 1987), in order to expand teacher assessment practices in Science subjects, demonstration experiments are important to improve performance assessments, particularly in Physics (White & Gunstone, 1992). The intervention was aimed at improving teacher assessment practices in Physics, and was conducted in such a way that teachers and students could conduct demonstration experiments performing real-world tasks while working in their normal classroom schedule under existing conditions and materials. Thus, the improvement of assessment practices proposed by this study was investigated under ordinary classroom circumstances and not in a setting specifically created for this research. So, the study approach was geared towards what works in schools and how this can be improved on the basis of intended consequences.

Methodologically the study applied, for the survey approach, mixed methods in recognition of the fact that both quantitative and qualitative methods may have limitations and one can neutralise the limitations and biases of the other (Creswell, 2003). Triangulation was considered as a means to seek convergence of findings. Still regarding the survey, the principle of using different data sources and multiple data collection instruments was used to guarantee triangulation. For the intervention, the strategy consisted of formative evaluations of exemplary assessment materials (specifically designed for this study) where the quality was verified by investigating the validity, expected practicality and expected effectiveness of the materials produced (Nieveen, 1997; van den Akker, 1999; van den Akker and Plomp, 1993).

Subsection 4.3.1 presents the research design of the Baseline Survey. It discusses the methodological aspects of operational research questions, population and sampling, instrumentation and data analysis.

4.3.1 Research design for the Baseline Survey

In order to address the objective of the research, it was necessary to start by undertaking a preliminary identification of assessment practices currently used by Grade 12 Physics teachers in Mozambican schools. Specifically, this implies a search of the kind of assessment practices Grade 12 teachers of Physics currently apply at classroom level and what can be said about the quality of these practices. The country has a population of 120 secondary school teachers teaching Physics in Grade 12, distributed in 30 schools of General Secondary Education – cycle 2 (ESG2) (per June 2004). As the purpose of the Baseline Survey was to inform the Intervention Study about assessment practices used in this target population of teachers and schools, and not to gain a national representative picture, it was believed that a small survey of some purposively selected Mozambican secondary schools from different provinces would be sufficient. In other words, it has been assumed that the perspective of various purposively selected school contexts would be representative for the characteristics of teachers, students, and schools.

As indicated in Chapter 3 (Section 3.6), a preliminary research aimed at identifying the assessment practices used by secondary school Physics teachers in Mozambique, was undertaken. The aim is expressed in three operational research questions. Because the research questions are formulated in line with the three corresponding research purposes (see Section 4.1), it is useful and relevant to gain a good understanding of a number of characteristics of the assessment practices applied by teachers, viz. the types of assessment practices, their quality, and their relevance for learning. These three elements constitute the perspective from which the characteristics of assessment practices are viewed during the survey in schools.

The identification of types of teacher assessment practices to be looked at in the classroom was firstly informed by the literature review and later refined by the pilot phase of the data collection instruments. From the variety of formative assessment practices referred to by the literature as crucial for what teachers need to know in order to undertake a contextualised assessment, observation methods, oral questioning, peer-assessment, and portfolios are the most critical (see Chapter 3, subsection 3.4.3). While oral questioning,

peer-assessment and portfolios were directly observed from the classroom by the researcher, teachers' own observation methods were not easily observable. In order to accomplish this, the strategy was to analyse how teachers observed students' process of designing and developing finished products resulting from a certain planned activity. From this analysis, it was possible to record the teacher's own comments and suggestions for improvement. Finished products include all written tasks that the students do and which reflect a certain development process of collecting, interpreting and reporting data. These products are known in schools as projects. In addition to these assessment practices, paper-and-pencil tests were also investigated due to their potential in assessing student abilities to understand basic concepts.

As a result of lessons learnt from literature and improvements arising from the pilot phase the following types of teacher assessment practices used in the classrooms, namely portfolios, peer-assessment, verbal tests, paper-and-pencil tests and projects were investigated in this study. These assessment practices are deemed relevant and good types of assessment to be used in classroom assessments (Popham, 2002; Weeden et al., 2002). Since these practices had been already referred to by the literature as assessment strategies that teachers normally use in schools, the identification of the types for this study was done by verifying which ones were more frequently used by Mozambican secondary school teachers to assess their students.

The term 'quality' of assessment practices refers to all aspects of validity and reliability of these practices. As referred to earlier in Section 4.1, the quality of assessment practices includes the teacher knowledge and skills for assessing student work. Thus, the quality aspect was investigated by analysing how teachers were assessing oral communication during lessons, written work, presentations, notebooks, and laboratory work of their students. According to several authors, these student tasks, if regularly undertaken, are indicators of the quality of assessment, particularly for science subjects (Black et al., 2003; Race et al., 2005; Weeden et al., 2002). In the context of this study, the quality of assessment practices used by teachers was investigated by two means: (i) verifying the

frequency of use, by teachers, of these student tasks and (ii) checking their validity and reliability.

As in the two previous aspects, the element of *relevance* was investigated on the basis of what the literature suggests is good practice. Popham (2002) for instance, argues that if the teacher shares the goals to be achieved with the students and involves them in the evaluation of their own work, this allows students to know what is expected from them. Thus, the issue of assessment relevance was addressed by investigating (i) how teachers engage students in the evaluation of their performance, and (ii) how often they use the assessment results to guide the student learning.

This section is composed of four subsections. Subsection 4.3.1.1 elaborates on the three operational research questions referred to earlier in this chapter and how the research design will address these. Subsection 4.3.1.2 presents the population and the sample of the Baseline Survey. Subsection 4.3.1.3 describes the instrument development. The process of instrument development and piloting, triangulation, identification of data sources, as well as the procedures followed during the survey, are all presented and discussed in this subsection. Subsection 4.3.1.4 presents the methods used for analysing data.

4.3.1.1 Research questions for the Baseline Survey

The main research question and the three operational research questions are presented and described in this subsection. The main research question addressed by the Baseline Survey is:

What assessment practices do Grade 12 teachers in Physics in Mozambique apply and what is their quality and relevance?

This question was addressed by a survey of 12 teachers from six schools in Gaza, Zambézia and Cabo Delgado Provinces, and of five educational officers from the Ministry of Education and Culture in Maputo. Taking into account the aspects of which

characteristics of assessment practices were to be investigated (types, quality and relevance), the formulation of the operational research questions for the Baseline Survey followed the same classification. Three operational research questions were formulated. The first operational research question sought to identify the types of assessment practices used by teachers, namely:

a) What assessment practices do Grade 12 Physics teachers apply?

Assessment practices investigated in the classroom included the five types mentioned earlier, namely portfolios, peer-assessment, verbal tests, paper-and-pencil tests, and projects. As was mentioned earlier, the criterion used to identify the types of assessment practices used by teachers in schools was to verify and count how many times (i.e., how often) the teachers used each assessment practice during several classroom sessions. Therefore, in order to address this question the teachers were asked to give information, amongst other questions, on the following sub-question:

- *a. How often do you use each of the following assessment practices? Portfolios, peer-assessment, verbal tests, paper-and-pencil tests, projects*

By means of questionnaires, interviews and through classroom observations, it was possible to verify - by checking the frequency (daily, weekly, monthly, never) of using the different assessment practices - which assessment practices the teachers have been applied during the classroom assessment. Teachers were also allowed to describe other possible assessment practices, which they use.

The second operational research question addresses the quality of the assessment, namely:

b) What is the quality of the assessment practices?

Aspects of assessment quality include not only the frequency but also the characteristics of the assessment tasks in terms of students' knowledge or skills (reasoning, memory or

process) being assessed by certain assessment type or activity. These elements were addressed by the following sub-questions:

- *b1. How often do you assess the following student activities? Oral communication during lessons, written work, presentations, exercise books, laboratory work, solving problems*
- *b2. What can be said about the validity and reliability of the assessment practices?*

The validity and reliability of the student activities was verified by the kind of feedback given to students by teachers in those different activities. Aspects of expression (whether the feedback was congratulatory or critical), time (timely feedback or given afterwards), and personality (individualised or in-group feedback) were used for the most assessed activities (Race et al., 2005).

Finally, the third operational research question deals with the relevance of these assessment practices for learning, namely:

c) How relevant are the assessment practices for student learning?

The relevance of the assessment practices refers to those elements that express the level of students' involvement in their own assessment, as well as the follow-up actions to be undertaken by the teacher after handing the assessment results out. Two sub-questions were formulated to address this question, namely:

- *c1. How do you engage students in the evaluation of their performance?*
- *c2. How often do you use the assessment results, for what purposes, and how?*

Teachers were given several alternative options on students' involvement in the evaluation of their performance namely (i) I do not involve them at all; (ii) by handing the results out; (iii) by involving them in self-assessment; (iv) by sharing with them the goals to be achieved; (v) by explaining to them the implications of the evaluation; (vi) by reflecting with them on the assessment data. Particular emphasis was given to peer-assessment due to the impact of this type of assessment in self-assessment (Race et al., 2005).

The three operational research questions were investigated using various target populations, which are described in the following section.

4.3.1.2 Population and sampling

There were three target populations relevant for addressing the operational research questions of the Baseline Survey. These are listed below.

- Teachers – are the active subjects in the assessment processes being investigated.
- School directors - have the responsibility for implementing the government regulations on assessment and on monitoring the quality of teaching in their school. They also play a role in creating a supportive school culture.
- Education officers – are responsible for providing the infrastructure to schools and inspect whether schools do a good job in terms of quality education.

As the Baseline Survey was aimed at gaining an impression of the assessment practices of Physics teachers in schools, it was believed that a small survey of Physics teachers and school directors from six Mozambican secondary schools from various provinces, representing the different contexts (urban-rural, different regions in the country), would be sufficient for this purpose (see also Chapter 1, Section 1.3 and Chapter 4, subsection 4.3.1). As referred to in Chapter 2 (Section 2.1), the country is composed of eleven provinces, clustered into the North, Centre, and South. Three provinces were drawn from these regions – one from each – and two schools were selected from each of these provinces. Maputo City was only considered for pedagogical officers and assessment specialists. No schools were selected from Maputo City because schools in this area appeared to be exhausted by extensive research activities taking place at the time. In order to enable comparison between teachers' responses, schools were selected according to their capacity of having at least two teachers teaching Physics in Grade 12. However, there were still three schools in which only one teacher taught Grade 12 Physics. This was the case in Pemba and Montepuez Secondary Schools in Cabo Delgado and Mocuba

Secondary School in Zambézia. Where this occurred, an additional teacher was taken from Grade 11. In the end, two Physics teachers from each of the two selected schools in the three provinces (in total 12 teachers) were sampled. The school directors were also sampled for participation in the study. In total, the intended sample was composed of twelve teachers and six school directors. Only four school directors, however, participated in the study as such, because two of the school directors were also Physics teachers and, due to practical limitations, they could only participate in one capacity. Given the focus of the study the role of the teacher was considered more important and, therefore, they had to provide information as teachers and not as directors. It was very important to obtain as much information as possible about the assessment practices carried out by teachers and the number of teachers excluding these ones would have been insufficient for this purpose.

Besides, five educational officers (two pedagogical officers and three assessment specialists) from the Ministry of Education and Culture (MEC) were asked to participate in the study. The pedagogical officers and assessment specialists were purposefully selected from the MEC in Maputo City due to their responsibilities for monitoring the assessment system within the Ministry. Of these pedagogical officers, one is the Director of the National Institute for Educational Development (INDE) - an institution responsible for curriculum review for both primary and secondary education - and a former Head of the Department of Assessment and Certification in the Ministry, and the other is the National Education Inspector. Concerning the assessment specialists, all of them were science subject specialists working in different departments within the Ministry. Table 4.1 summarises the details of the realised sample for the Baseline Study.

Table 4.1: Sample of Baseline Survey

| Region | Province | Institution | Nr. of teachers | Nr. of school directors | Nr. of educational officers |
|--------------|--------------|-----------------------------------|-----------------|-------------------------|-----------------------------|
| South | Gaza | Joaquim Chissano Secondary School | 2 | 1 | - |
| | | Chókwè Secondary School | 2 | 1 | - |
| | Maputo City | Ministry of Education and Culture | - | - | 5 |
| Centre | Zambézia | 25 de Setembro Secondary School | 2 | 1 | - |
| | | Mocuba Secondary School | 2 | - | - |
| North | Cabo Delgado | Pemba Secondary School | 2 | 1 | - |
| | | Montepuez Secondary School | 2 | - | - |
| Total | 3 | 7 | 12 | 4 | 5 |

The samples of schools (and therefore the samples of teachers, school directors and educational (both pedagogical and assessment) officers) were *purposive samples*, as they were drawn with the purpose of obtaining insight into three important perspectives on the classroom practice namely instruction in relation to assessment, management's perspective in relation to teachers' preparedness for conducting appropriate assessments, and inspectorate regarding quality control of the teacher assessment practices, with the view to using the information to design an intervention study and not to generalise to full populations. With these samples, all activities were undertaken to address the main research question for the Baseline Survey formulated at the beginning of subsection 4.3.1.1.

4.3.1.3 Data collection strategies

This subsection comprises four parts. The first part presents the number and characteristics of data collection instruments used in this study, and a summary of the content of each instrument. The second part discusses the development process of the various instruments, the piloting process, and the validation of the instruments by experts. The third part starts by providing information on the type of data collected to answer operational research questions, the way it was collected, the triangulation process of instruments and data sources, and ends with a summary of all information in a data

collection matrix. The fourth part presents procedural information on the number and sequence of activities that were carried out in preparing and conducting the Baseline Survey for this research.

Instruments and data collection strategies

Five data collection instruments were developed for the Baseline Survey namely, a questionnaire for teachers, a classroom observation schedule, and three interview schedules for teachers, school directors, and pedagogical officers.

The questionnaire for teachers (Appendix A) consisted of four main sections. The first section contained information about the questionnaire itself (e.g., what is it about, why should it be filled in) and requested personal background information about the respondents (name, gender, age, school, etc.). The second section requested information about the types and quality of assessment practices used by teachers in the classroom. This information was sought through five closed questions of multiple choice and Likert scale type of items. The third section was about the relevance of the assessment practices and comprised of four questions of multiple choice items. The fourth and final section contained evaluation questions with one question containing multiple-choice items and another one being open-ended. It is worthy to mention that, although the majority of the questions were closed, they provided teachers with the opportunity to express their views and opinions by exploring the 'other specify__' type of items. Furthermore, in the evaluation questions teachers were asked to comment about any other issue that was not addressed in the questionnaire.

The classroom observation schedule (Appendix B) contained four sections. The first section contained background information about the teacher and the school, the second was about the physical appearance of the classroom and the teaching and learning environment. The third section presented the description of the students (e.g., number, gender, age), and the fourth section comprised a number of close questions (Likert scale items) related to the types of assessment practices undertaken by the teachers and their relevance for learning. More specifically, the questions addressed the extent to which

assessment practices were applied by teachers, their quality as demonstrated by teachers and students, their appropriateness for instruction, and their validity and reliability for student learning.

The interview schedules for teachers, school directors, and pedagogical officers (Appendices C, D, and E) comprised of 13, nine, and ten questions respectively. All schedules had an introduction stating the aim of the interview and the reasons why the interviewees and their respective schools were chosen to participate in the study. The introduction also indicated that the identity of the interviewees would remain anonymous and the information confidential. The interview questions addressed similar issues as in the questionnaires such as types, quality and relevance of assessment practices with the intent to cross check the information. There were three additional aspects addressed in the interviews with the pedagogical officers from the MEC. The first aspect was related to the objectives of the teachers' assessment as seen by the Ministry. The second asked how the translation of these objectives into practice is compared to the information provided by the national examinations undertaken by the Ministry. The third was meant to seek the opinions of the interviewees about the impact of the supervisory visits to schools.

Overall, the questionnaires and interviews were designed to gather information about the types, quality and relevance of assessment practices used by secondary school teachers in schools. The classroom observation schedule was also used by the researcher to triangulate the information on assessment practices given by teachers and school directors. But more importantly, it permitted the observation of the physical conditions of the classrooms, the teaching and learning environment, and the characteristics of the students. The different instruments used with different data sources allowed for cross checking of the information and increased its validity and reliability.

Development process and piloting

The first version of data collection instruments for Baseline Survey was developed by the researcher and the instruments were appraised by experts to ensure their validity. The instruments were all piloted before the data collection process. Questionnaires and

interview schedules for teachers and school directors, as well as, classroom observation schedules, were piloted with five Grade 12 Physics teachers randomly selected from schools around Maputo. School directors' interviews were piloted with two directors of some secondary schools also located in the Maputo area. The interviews for pedagogical officers were piloted with two pedagogical officers from INDE. The main objective of the pilot phase was to increase the validity of the instruments in terms of language, depth of assessment content approach, and time required for completing the instruments. For instance, one expert on designing assessment instruments and one Science educator specialist were asked to provide their comments and suggestions on how to improve the quality of the instruments (content validity). Both experts scrutinised all the instruments in order to determine their validity in terms of face validity. The reliability of the instruments was checked by verifying the consistency of the responses and to ensure that respondents answered related items in a similar way (internal reliability). More specifically, the piloting was intended to find out whether all members of the sample, especially teachers, would be able to understand the instruments and to complete the questionnaires in time.

After revision of the first version, the instruments were finalised. All the participants in the pilot phase were asked to comment on the content and practicality of the instruments. Overall, the pilot phase was instrumental in improving the validity and practicality of the data collection instruments by generating valuable suggestions for improving of the final version.

Data collection and triangulation

As was discussed earlier, in order to gather the information needed to answer the main research question of the Baseline Survey and to assure the validity and reliability of the information gathered the principle of triangulation of data sources and of instruments was applied.

The information needed to obtain answers for the *specific research question a.* included (x1) the types and frequency of usage of a certain type of assessment practice, (x2) the opinions about why teachers assess and evaluate the student performance, including about

the teacher preparedness level and (x3) the physical conditions under which the assessment practices are carried out. This information could best be obtained from questionnaires with the purposively selected teachers. More generally, teachers were asked to reflect about the different assessment practices they use in their classroom, their purposes, how they conduct them in practice, how they evaluate the final performance of the students, and under what physical conditions they work. To validate the information from the teacher questionnaires a number of interviews were conducted with the teachers. In the interviews they also had to respond to the questions about types and quality of assessments with the aim of supporting or refuting their arguments expressed during questionnaires. Classes were observed and interviews conducted with school directors (who are also teachers) and with pedagogical officers for the same purpose.

For the answers to the specific research question b. the information needed included (y1) data about the frequency of assessing, by teachers, of certain students' activities, and the teachers' opinions firstly about whether the assessment practices used allow students to demonstrate their performance (y2) and secondly about the characteristics of their scoring procedures of these practices (y3) – whether they are clear, consistent and unbiased. The information about the quality of assessment practices was mainly obtained through questionnaires (to teachers and school directors) and interviews (to school directors and pedagogical officers). Pedagogical officers were also interviewed to gather their views about the level of preparedness of teachers in designing and administering classroom assessments and as sources of information complementary to the information provided by teachers and school directors. The validation of the information from both data collection instruments was done by researcher's observations of the classroom practices and by written notes provided by the assessment specialists.

Regarding the information necessary to address the *specific research question c.* data included (z1) the use by teachers of assessment results to monitor student learning, and (z2) the teacher preparedness to design classroom assessments including the coverage of relevant topics, and student involvement in the evaluation of their own work. This information was obtained from teachers and school directors through questionnaires. To

validate this information two means were used: (i) the first was the interviews with school directors, both to cross check the information provided by other teachers and the observations during lessons (external reliability) as well as to enable teachers to express their own views about assessment practices currently in use in their schools; (ii) the second was written notes from the assessment specialists who were asked to state in writing not only how they perceive the classroom assessment as taking place in the schools, but also the MEC's philosophy of what they consider as desirable assessment practices for schools as well as the level of teacher accomplishment of them particularly as assessment specialists have been playing a major role in the supervision of the teaching quality and in the development of national assessments.

As referred to earlier, apart from interviews, questionnaires and written notes from assessment specialists, the researcher also conducted classroom observations. Eight classroom observations were conducted with eight teachers; the remaining four teachers did not have their classes observed because they were not available at the time of observations. Classroom observations were deliberately not announced in advance to avoid simulation of lessons. This was also the reason why only one class was observed per teacher where the objective was mainly to obtain information about the unplanned assessment practices, the teaching strategies, and the physical conditions of the typical Physics classroom. Verifying teachers' formal assessment practices was not necessarily the objective of the observations, because formal assessments are planned and announced in advance and they did not take place on the dates of the observations.

Overall, while teachers mainly provided information about type, quality, and frequency of assessment practices, school directors, pedagogical officers and assessment specialists were asked to give their opinions about the quality of assessment practices, the use of assessment results for monitor student learning, and the level of teachers' preparedness in designing acceptable assessment practices.

Thus, in order to guarantee valid and reliable information and for triangulation purposes, a variety of data collection strategies, instruments, and data sources were used to answer the

formulated research questions. This is summarised in the following data collection matrix (Table 4.2).

Table 4.2: Data collection matrix

| Instrument | From teachers via: | | | From school directors via: | | From pedagogical officers via: | From assessment specialists via: |
|--|--------------------|----|----|----------------------------|----|--------------------------------|----------------------------------|
| | Qn | Iv | Ob | Qn | Iv | Iv | Wn |
| Information variables | | | | | | | |
| a. Assessment practices used | | | | | | | |
| -types and frequency (x1) | ✓ | | ✓ | | | ✓ | |
| -opinions about why do they assess, how much of the teaching time they spend on assessment and how do they evaluate the student performance (x2) | | ✓ | | | ✓ | ✓ | ✓ |
| -level of teacher preparedness | | ✓ | ✓ | | ✓ | ✓ | |
| -physical conditions (x3) | | | ✓ | | ✓ | ✓ | |
| b. Quality of assessment practices | | | | | | | |
| -frequency of certain student activities (y1) | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ |
| -allowing students to demonstrate performance (y2) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| -clear, consistent and unbiased scoring procedures, etc. (y2) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | |
| c. Relevance of assessment practices | | | | | | | |
| - frequency of using student assessment results (z1) | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ |
| -level of student involvement in evaluating their own work (z2) | ✓ | ✓ | | ✓ | ✓ | | ✓ |
| -coverage of relevant topics and appropriateness for instruction (z2) | ✓ | ✓ | | ✓ | ✓ | | ✓ |

Qn = questionnaire, Iv = interview, Ob = classroom observation, Wn = written notes

In practice, questionnaires and classroom observation schedules were the main data collection instruments. Interviews were conducted after the completion of the

questionnaires and the observation of classes in order to guarantee the reliability of the information. To avoid copying of information, the interviews were conducted on the same day with all teachers, one after another.

Research procedures

The main research activities of this study started with the literature review that was carried out between June 2003 and April 2004. The document analysis at the MEC, aimed at obtaining statistical information about the number of secondary schools teaching Grade 12 Physics in the country as well as their number of Physics teachers, was undertaken from May to June 2004. Then the instrument development phase was initiated where the first version of the self-completed questionnaires, classroom observation schedules, and interview schedules were developed by the researcher. These were then piloted. Then the main fieldwork was conducted during which a number of activities were undertaken sequentially. As was referred to earlier, firstly questionnaires were administered to the twelve selected teachers from the six schools. Then interviews were conducted with teachers and thereafter, classroom observations were undertaken, focusing on teachers' instructional practices. These took place in August 2004 and March 2005. The interviews with the school directors took place from August to October 2005. In November 2005, the interviews with pedagogical officers from the MEC took place. Finally, the written notes which had been requested by assessment specialists and from the MEC (who were previously asked to provide their thoughts in writing about Physics classroom assessment) were collected in February 2006.

The subsection 4.3.1.4 presents the methods used to analyse the data.

4.3.1.4 Data processing and analysis methods

Self-completed questionnaires from the teachers were analysed quantitatively using categorisation of questions and calculation of frequencies (Bardin, 1977) following a set of procedures for describing, synthesising, analysing, and interpreting data (Gay & Airasian, 2003). Frequencies were produced from the analysis and were presented in

graphs and tables. The software package, Statistical Package for the Social Sciences (SPSS - version 8.0), was used to directly capture all quantitative data from different data collection instruments. Data were analysed and presented using frequency tables.

Interviews and classroom observations were analysed qualitatively through summarisation of questions and categorisation of the responses (Miles & Huberman, 1994). Thematic content analysis was employed to analyse the data (Bardin, 1977; Race et al., 2005). Contact summary forms with excerpts for illustration were then filled in to review the interview and classroom observation notes and an overall summary of the main findings was produced. Two main concurrent flows of activity were followed, namely (i) a process of deciding on the meaning of each item of data or set of data by noting similar patterns or explanations and (ii) a process of assembling information that allows to draw conclusions and to take further action. Findings from the Baseline Survey are presented and discussed in Chapter 5.

Subsection 4.3.2 is the research design of the Intervention Study. The section introduces the educational design research approach, elaborates on the design of Physics assessment materials prototypes, and ends with a discussion of design guidelines for the intervention.

4.3.2 Research design for the Intervention Study

The study on investigating and improving assessment practices in Physics in secondary schools in Mozambique focuses, on designing and developing Physics assessment materials aimed at helping teachers to improve their assessment practices in schools. Following from the main research question of this study, as formulated in Chapter 1 (Section 1.2), the aim of this study is twofold, namely to identify assessment practices used by secondary school Physics teachers in Mozambique and to undertake an intervention aimed at improving these practices. The Baseline Survey described in previous sections of this chapter addressed the first part of the aim. The Intervention Study dealt with in this section addresses the second part of the study aim and its research question is formulated as:

How can the teacher assessment practices be improved?

In order to find an answer to this question, and following from what was referred to in Chapter 3 (Section 3.5), the Intervention Study applies an educational design research methodology as suggested by van den Akker and Plomp (1993). Findings of the Baseline Survey indicated that, although teachers have regularly been using paper-and-pencil tests - which are good for grasping the basic concepts, performance assessment is important for students to be able to perform real-world tasks and Physics cannot be taught and assessed without practical or laboratory work (see Chapter 5). This argument led to the reinforcement of what is already the Mozambican government policy of adopting the constructivist approach to learning and teaching. In fact, constructivist perspective underpinned the approach applied in this study in improving teacher assessment practices. A number of Physics assessment prototypes were designed in the context of demonstration experiments and they were formatively evaluated in classroom tryouts.

This study was conducted following a process of analysing the problem context, carrying out a Baseline Survey, recommending a type of intervention to be made, and designing and formatively evaluating assessment prototypes. This process reflected a nature of two mixed method approaches namely survey and educational design research.

Three subsections comprise this section. Subsection 4.3.2.1 elaborates on the nature of the educational design research approach. Subsection 4.3.2.2 discusses educational design research as applied in this study, that is, in the context of the Physics assessment materials. Subsection 4.3.2.3 presents the design guidelines for the Intervention Study. This subsection also provides orientation elements for the prototyping process as design specifications for the Physics assessment materials.

4.3.2.1 Educational design research

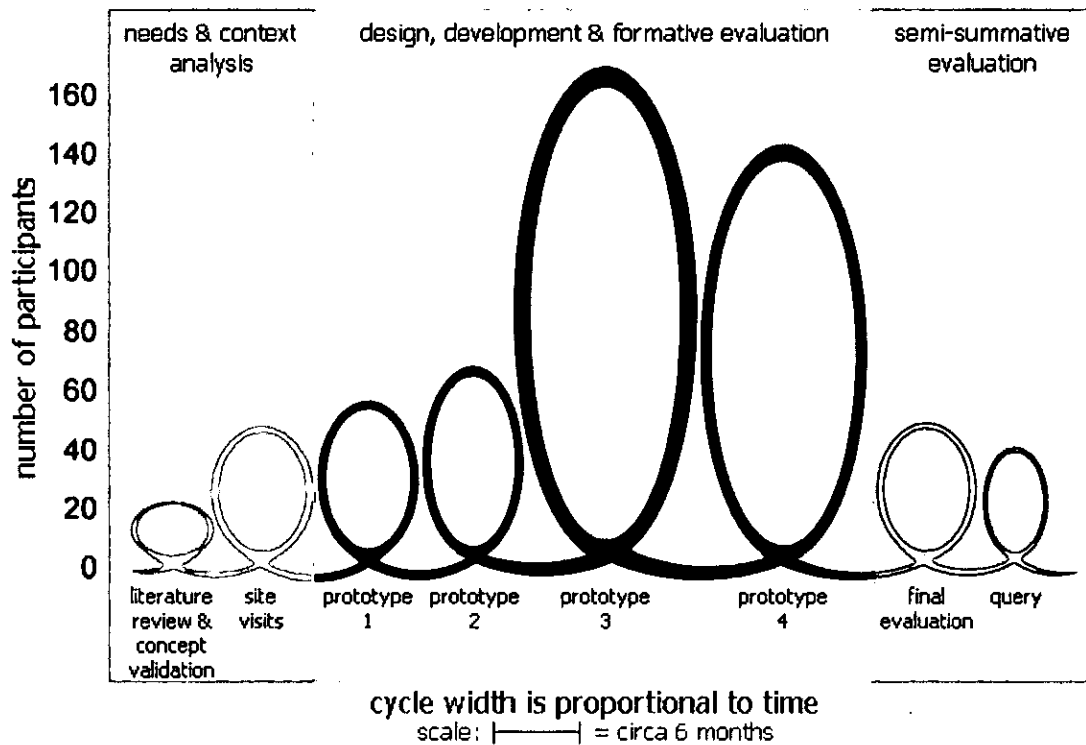
According to van den Akker (1999), educational design research could be an appropriate approach for a complex situation where the appropriateness and effectiveness of an

intervention is unknown beforehand and its success depends on the design and implementation process within a wide variety of the contexts. This is indeed the case for Mozambican schools where many people (particularly teachers, students, and teacher educators) are unfamiliar with the process of designing assessment materials. Furthermore, curriculum materials (including assessment ones) in the country are developed using several theories but rarely utilise findings from research. Therefore, as was referred to in Chapter 2 (subsection 2.2.2), since the Ministry of Education and Culture is in the process of reviewing the curriculum for secondary education (including assessment issues), timely and adequate information is required for the reviewers to make the right choices in such a complex and dynamic situation.

Educational design research is a systematic process of designing, developing and evaluating instructional programs, processes, and products that must meet the criteria of validity, practicality and effectiveness (Seels & Richey, 1994; van den Akker, 1999; van den Akker & Plomp, 1993).

The function of the educational design research, as is the case in any scientific research, is the search for understanding, which results in contributing to the body of knowledge or theory. This search can occur with a very broad purpose through conducting scientific research in a certain domain or at micro level of specific research projects. In education research, the search intends: (i) to contribute to the body of scientific knowledge or theories about education; (ii) to contribute to improving practice; and (iii) to inform decision-making and policy development in the domain of education. Within the context of a research project, various functions of research can be distinguished, namely to describe, to compare, to evaluate, to explain, and to design and develop.

The research process in educational design research (also called development research) comprises educational design processes undertaken in cyclical stages of analysis, design, evaluation and revision activities. The stages could be refined continuously until reaching a satisfying balance between the intended and the realised stage. McKenney (2001) cited by Plomp (2006) gives an illustration of the cyclical process as set out in Figure 4.1.



(Source: McKenney, 2001)

Figure 4.1: The cyclical process of educational design research

According to Plomp (2006) three distinctive phases can be distinguished from this example as set out below.

1. Preliminary research: this involves a needs and content analysis, and a literature review (including site visits) leading to a conceptual framework.
2. Prototyping phase: this requires iterative and cyclical design, development with formative evaluation of the several prototypes as the most important research activity aimed at refining the intervention.
3. The Assessment phase involves semi-summatative and final evaluation to conclude whether the solution or intervention meets the pre-determined specifications.

Throughout all these activities the researcher will do 'systematic reflection and documentation' to produce the theories or design principles as the scientific yield from the research.

As already reported, this study started with the preliminary research on relevant literature about what scholars regard as good practice in classroom assessment, as well as on the context analysis (Baseline Survey) of what assessment practices Grade 12 Physics teachers use in schools.

During the prototyping phase the guiding orientation for employing educational design research in designing exemplary materials and fulfilling the functions mentioned above, is that these prototypes must be of good quality (Nieveen, 1999). Nieveen suggests a framework for making the concept of quality in exemplary materials more transparent, which includes three criteria namely: validity, practicality, and effectiveness. Validity is attained when there is internal consistency between the materials and the state-of-the-art knowledge (content validity), and between the different components of the materials (construct validity). Practicality is attained when the materials are usable by teachers and students in a way that is compatible with the developer's intention. Effectiveness is attained when students appreciate that the desired learning tasks and the learning programme are taking place.

The cyclical character of educational design research does not mean that activities are just undertaken repeatedly. The quality criteria are taken into account and they are given different emphasis in different stages of the research. For example, during the preliminary research where the emphasis is on analysing the problem and reviewing the literature, the criterion of validity is the most dominant, with some attention given to practicality, whilst in that state no note is yet taken of effectiveness. On the other hand, in the prototyping phase much attention is paid in the formative evaluation to the criterion of *practicality*, whilst effectiveness becomes increasingly important in later iterations. Finally, the systematic reflection and documentation is undertaken at the end of each designing cycle, and it is aimed at enhancing design principles and implementation solutions. Table 4.3 depicts the phases of educational design research, the quality criteria emphasised in each phase, and the activities being undertaken.

Table 4.3: Quality criteria related to phases in educational design research

| | Phase | Criteria | Short description of activities |
|---|----------------------------|---|--|
| 1 | Preliminary research phase | More emphasis on validity, less on practicality | Review of the literature and of (past and/or present) projects addressing questions similar to the ones in this study. This results in a framework and first blueprint for the intervention. |
| 2 | Prototyping phase | Initial emphasis on validity and practicality. Later on mainly practicality and gradually shifted to effectiveness | Development of a sequence of prototypes that will be tried out and revised on the basis of formative evaluations. Early prototypes can be just paper-based for which the formative evaluation takes place via expert judgments. Evaluate whether target users can work with intervention (practicality) and are willing to apply it in their teaching (relevance & sustainability), also whether the intervention is effective. |

Various terms are used in the literature for the preliminary research activity, such as 'orientation' (Hadi, 2002), 'needs and context analysis' (McKenney, 2001), 'front-end-analysis' (Ottevanger, 2001), and 'in-depth orientation' (Thijs, 1999). The preliminary research in the Physics Assessment Materials (PAM) for this study is called 'Baseline Survey'. In relation to the prototyping activities, some other authors have used different terms such as 'design and development and evaluation stages' (McKenney, 2001) and 'development and evaluation and semi-summative evaluation stages' (Hadi, 2002), respectively. During the prototyping stage activities consist of tasks to articulate the design ideas into the empirical development stage.

Formative evaluation of the research activities takes place in all phases of educational design research. As illustrated by Table 4.3, it serves different functions in the various development cycles. It also has various layers in design research activities, from the more informal in the early phases of a project (self-evaluation, one-to-one evaluation, expert review) to small group evaluation aimed at testing the practicality and effectiveness, to a full field test (if applicable).

Procedurally, the prototyping phase of this study (the most important research activity) includes: (i) selecting limited exemplary themes or topics; (ii) designing assessment tasks in a standardised fashion; (iii) anticipating teachers' potential difficulties in the implementation process; (iv) providing detailed procedural specifications; and (v) applying a systematic process of formative evaluation of the products. On the basis of these considerations, a careful design of assessment materials is expected to improve the initial implementation process and ultimately the outcomes.

As was referred to in Chapter 3 (Section 3.5), several studies have been conducted using educational design research as an intervention approach, such as those done by Mafumiko (2006), McKenney (2001), Motswiri (2004), Ottevanger (2001) and Teclé (2006). For example, Mafumiko (2006) examined how micro-scale experimentation can serve as a catalyst for improving the chemistry curriculum in secondary schools in Tanzania. McKenney (2001) explored the possibilities of computer-based support for Science education materials developers in Africa. Motswiri (2004) investigated how to support chemistry teachers in implementing formative assessment of investigative practical work in Botswana. Ottevanger (2001) investigated teacher support materials as a catalyst for science curriculum implementation in Namibia. Teclé (2006) explored the potential of a professional development scenario for supporting biology teachers in Eritrea.

The research model by Mafumiko (2006) (Figure 4.2) is used to inform the research model for the Intervention Study of this study due to its similarities with this study. It is a design study aimed at improving a science curriculum in secondary schools in a similar context to that of Mozambique. Mafumiko's model shows a development process of teacher support materials and student worksheets in chemistry in four different versions following a subsequent design, formative evaluation, and revision steps of prototypes.

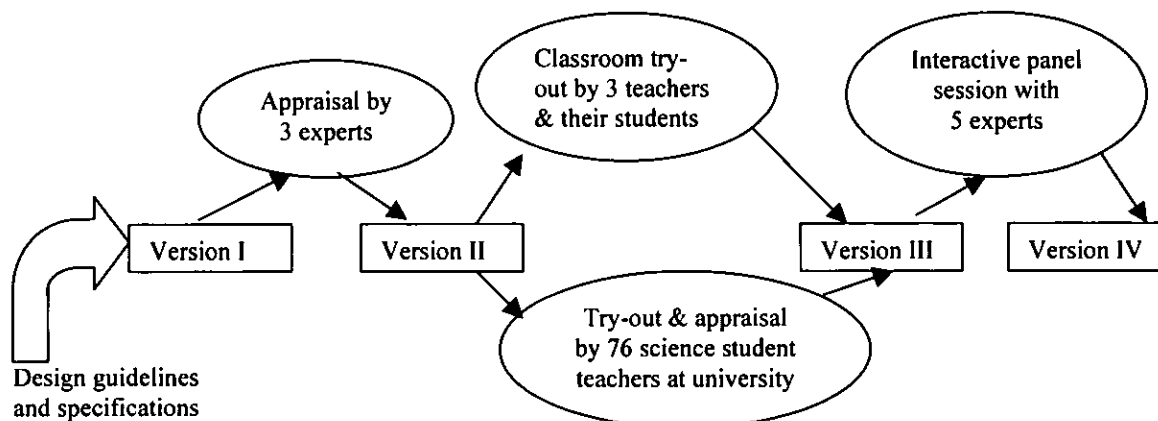


Figure 4.2: The original model by Mafumiko (2006)


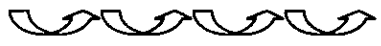
While the first version was developed by the author following design guidelines and specifications of exemplary lesson materials, the subsequent versions (1 to 4) were designed and formatively evaluated by experts and users. The quality of the prototypes was sought through subsequent analysis of the validity, and practicality and effectiveness of the materials. Due to the shortage of time, a full trial of the prototype Version 3 with teachers and students in the classroom did not take place and the intervention was restricted to appraisal by university students and experts. This led to the situation where only the expected practicality and the expected effectiveness of the third version of the prototypes were demonstrated. More evaluation is needed to demonstrate the actual practicality and the actual effectiveness of the intervention.

In the following subsection the educational design research in the context of the Physics assessment materials of this study is further discussed.

4.3.2.2 Design of the Intervention Study

The study on investigating and improving assessment practices in Physics in secondary schools in Mozambique is characterised by a mixture of a survey and educational design research approaches. The exploratory character of the Baseline Survey previously undertaken and the cyclical nature (design, evaluation and revision) of the Intervention

Study are important means of establishing evidence of a good quality within the limitations of this study. In general, the intervention for this study, that is PAM materials, was developed following an educational design research approach (Figure 4.3).

| Baseline Survey | <i>Intervention Study</i> |
|--|---|
| Preliminary research | <i>Prototyping</i> |
|  |  Version 1 Version 2 Version 3 Version 4 |
| Literature review on Physics classroom assessment and Mozambican education policies. Survey on assessment practices. | Design of the prototypes and classroom tryouts (more emphasis on validity and expected practicality with gradual shift to effectiveness). |

- Notes:*
1. Flowchart auto shapes indicate the findings from literature review, Baseline Survey and context analysis.
 2. Block curved arrows indicate cyclical character of educational design research approach.
 3. Increasing gray area means gradual up-scaling of the study.

Figure 4.3: General research design of the study

As already reported in Section 4.3 and in Chapter 3, the preliminary research phase, aimed at providing information about assessment strategies used for Science subjects, such as alternative, authentic, formative, performance assessments, and the role of both teachers and students in assessment, was undertaken. The policies of the Mozambican education system, particularly related to classroom assessment, were also reviewed including the developments around the curriculum review for secondary education (see also Chapter 2). Based upon the findings of this Baseline Survey, informed decisions were taken regarding the topics to be investigated, the type of assessment practice to be used as an example for improvement, the teaching, learning approach to be followed, and the assessment strategy to be adopted during the Intervention Study. The findings also provided orientation on the formulation of design specifications that generated the methodological direction for the design of the prototypes and its tryout in the classroom.

After the preliminary phase, the intervention study consisted mainly of the *prototyping stage* where the prototyping process is presented in terms of a series of subsequent design and formative evaluation and revision steps of versions of prototypes. Four versions of prototypes were developed before the final product was constructed. Validity, expected practicality, and expected effectiveness of the draft prototypes were the focus of this stage with the aim of acquiring clear empirical evidence of the performance of both teachers and students during the classroom tryout. Three experts, four teachers, and three university students (also mathematics or Science teacher) appraised the first version. The second version was produced on the basis of the revision made from the first version and was used in a classroom tryout by two teachers and their 62 students. The third version was developed following suggestions from users (teachers and students) and was appraised by two experts in an interactive discussion. As was referred to earlier, the practicality and effectiveness of the third version of the prototypes were only 'expected' because this version was only appraised by university students and experts and not via empirical testing. This phase resulted in the fourth and final version of the materials. The analysis of the expected effectiveness of this version was done through an evaluation workshop with university students and teachers including final suggestions from experts. Suggestions on the possible incorporation of the material into the new curriculum under review and consequent possible use by teacher training institutions of the PAM materials were given.

During the development of the four versions of the PAM prototypes, the quality was verified and increased in terms of their validity, expected practicality and expected effectiveness (Mafumiko, 2006; Nieveen, 1999; Ottevanger, 2001; van den Akker, 1999).

1. Validity refers to the internal consistency between the materials and the state-of-the-art knowledge (content validity) and to the fact that the various components of the intervention are consistently linked to each other (construct validity). The first phase of formative evaluation occurred with Version 1 of the prototype through appraisal by experts, university students, and schoolteachers and focused on improving more the validity and less the practicality of the prototype.
2. Practicality refers to the usability of the materials by teachers and students (including experts) in ways that are compatible with the developer's intention. In

other words, it means performing the design and tryouts of the activities in the conditions put in place in the learning environment, which are prescribed by the current Grade 12 Physics curriculum, and under the schedules of the Physics teachers. The second phase of formative evaluation took place only with Version 2 by classroom tryout by teachers and students and the emphasis was more on expected practicality of the material with little reference to its effectiveness.

3. Effectiveness refers to the extent to which all users (particularly teachers and students) appreciate the experiences and outcomes of the intervention and the learning task. In general, it reveals the implications of the intervention for both teachers and students in light of the acquired theoretical innovations. The effectiveness of the material was verified in Version 3 of the prototype through appraisal by three experts and constituted the third and last phase of formative evaluation. At the end, Version 4 had improved aspects of validity, expected practicality, and expected effectiveness as intended by the intervention.

Figure 4.4 depicts the research model of the Intervention Study, which includes the preliminary and the prototyping phases.

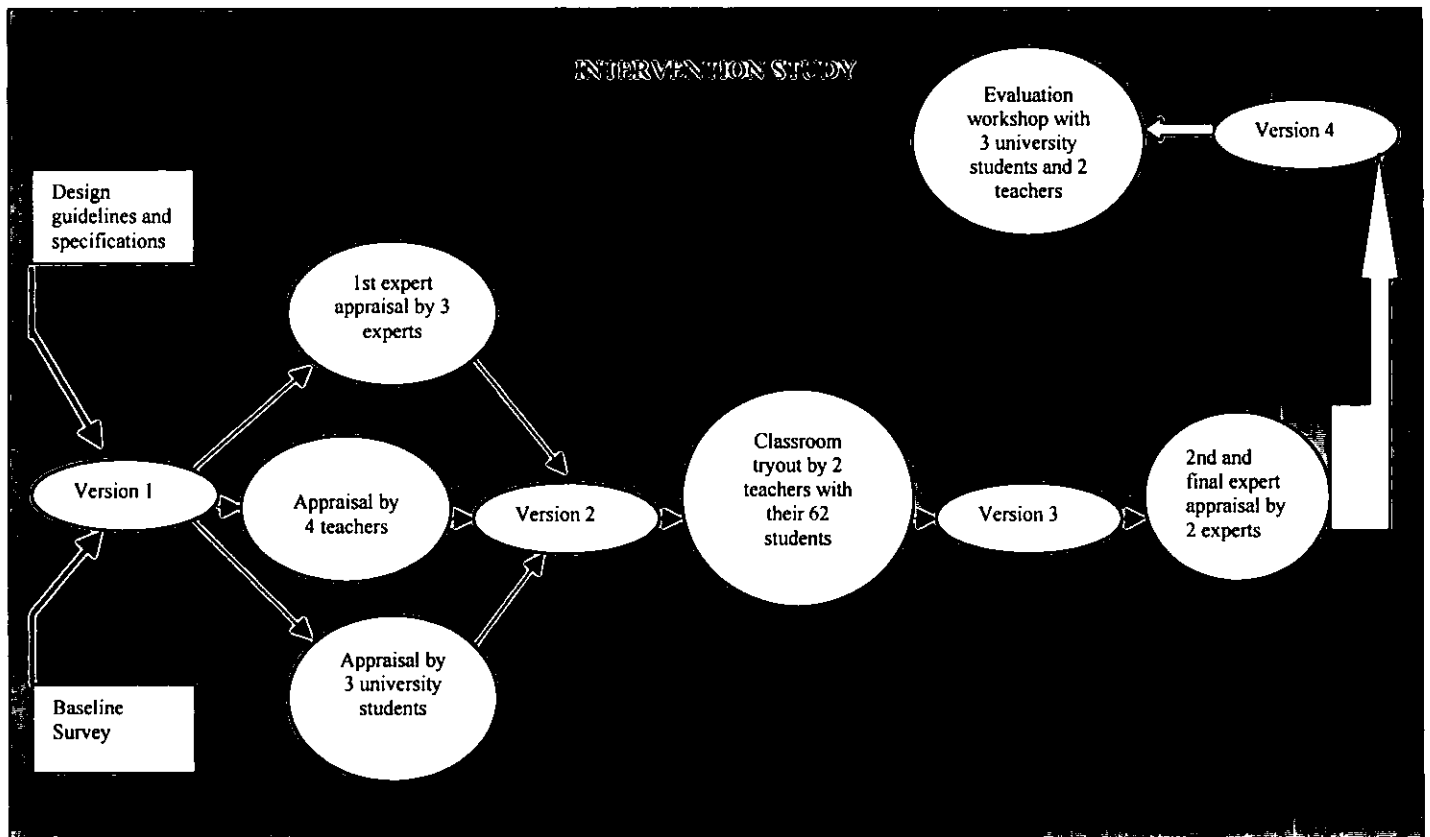


Figure 4.4: Research model of the Intervention Study

As was referred to earlier in this section, the model by Mafumiko (2006) was used to inform the research model for the Intervention Study. Three elements in Mafumiko's original model have been adapted. The first one refers to the use of findings from the Baseline Survey to inform the design process of Version 1. The second element refers to the increased number of appraisers of the Version 1 of the prototypes. The third element is linked to the final appraisal of the prototypes where, due to the shortage of time, it was not possible to try out the final version in the classroom with users to verify its actual practicality and effectiveness.

In summary, the research design for the development of Physics assessment materials for the study on investigating and improving assessment practices in Physics in secondary schools in Mozambique proceeded in the two phases described below.

1. The preliminary research phase consists of a review of the literature about assessment, an analysis of Mozambican education policies, and a Baseline Survey aimed at identifying assessment practices used by teachers in schools.
2. The prototyping phase consists of the development of a number of Physics assessment prototypes to be used by teachers in schools as a way of improving their assessment practices. This phase included:
 - a. the development of the prototypes in cyclical series of design and formative evaluation of the different versions of the prototypes using quality criteria of validity, expected practicality and expected effectiveness of the material in the various prototyping stages; and
 - b. systematic reflection and documentation consisting of analysis of the expected effectiveness of the prototypes and of the sustainability of the study findings.

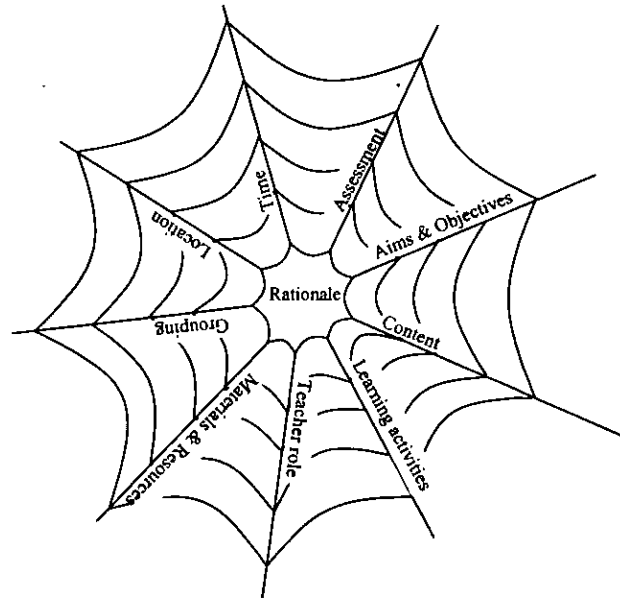
The findings of the preliminary research are discussed in Chapters 2 (Context of the study), 3 (Literature review and conceptualisation of the study) and 5 (Assessment Practices of Mozambican Physics Teachers). Subsections 4.3.2.3 and 4.3.2.4 present the design guidelines and the research procedures for the intervention study. The instrument development for the various formative evaluations and the findings of the Intervention Study are discussed in Chapter 6 (Improving Teacher Assessment Practices in Physics in Mozambique).

4.3.2.3 Guidelines for design of the intervention

Before elaborating on the guidelines for designing the intervention on assessment strategies, it is necessary to discuss how the intervention that is being studied can be looked at from a curriculum perspective. The rationale for presenting this curriculum perspective lies in the fact that assessment is always a component of a curriculum and the intervention will necessarily consist of lessons within which assessment will be the focus of interest. This argument is supported by international literature according to which assessment, instruction and curriculum go hand-in-hand and any attempt to predict a

future direction for assessment must consider the factors that can influence curriculum changes (NRC, 2001; Popham, 2002; van den Akker, 2003). Sadler (1989), quoted by the NRC (2001), provides a conceptual framework that places classroom assessment in the context of curriculum and instruction. Within this framework, three elements are required for formative assessment to promote learning, namely: a clear view of the learning goals derived from the curriculum; information about the present state of the student derived from assessment; and action – taken through instruction – in order to close the gap. Popham (2002) refers to the idea that teachers need to know how to create their classroom assessment devices, interpret, and use statewide test results, and then plan instruction based on instructionally informed assessments. This author also supports the view that teachers must be concerned with their instruction and with the fact that their assessments address appropriate content. Van den Akker (2003) argues that the various curriculum components are a powerful tool in understanding the planning of student learning and the development of accompanying learning materials. He describes, in what he calls a vulnerable curriculum spider web, ten curriculum components to consider in curriculum design and implementation and points to the fact that the crucial challenge for curriculum improvement is to establish balance and consistency between the various curriculum components (see Figure 4.5). With the term “spider web” the author’s intention is to illustrate an existing similarity between a spider web and a chain: the spider web is as strong as its weakest part; similarly, a chain is as strong as its weakest link. In fact, focusing on assessment means that the intervention is focusing on one of these components of the curriculum and any effort to improve assessment should be in a balanced and sustainable manner, taking into account all the components of curriculum.

One relevant lesson that can be learnt from these arguments is that, before putting more emphasis on assessment materials, one should focus on the quality of the lesson materials. For the teachers to be able to conduct effective assessment strategies they need support on preparing good lessons and therefore they need materials of good quality.



(Source: van den Akker, 2003)

Figure 4.5: The vulnerable curriculum spider web

Each of the ten components of the spider web addresses a specific question about the planning of student learning. Table 4.4 shows the curriculum components and their corresponding focus questions.

Table 4.4: Curriculum components

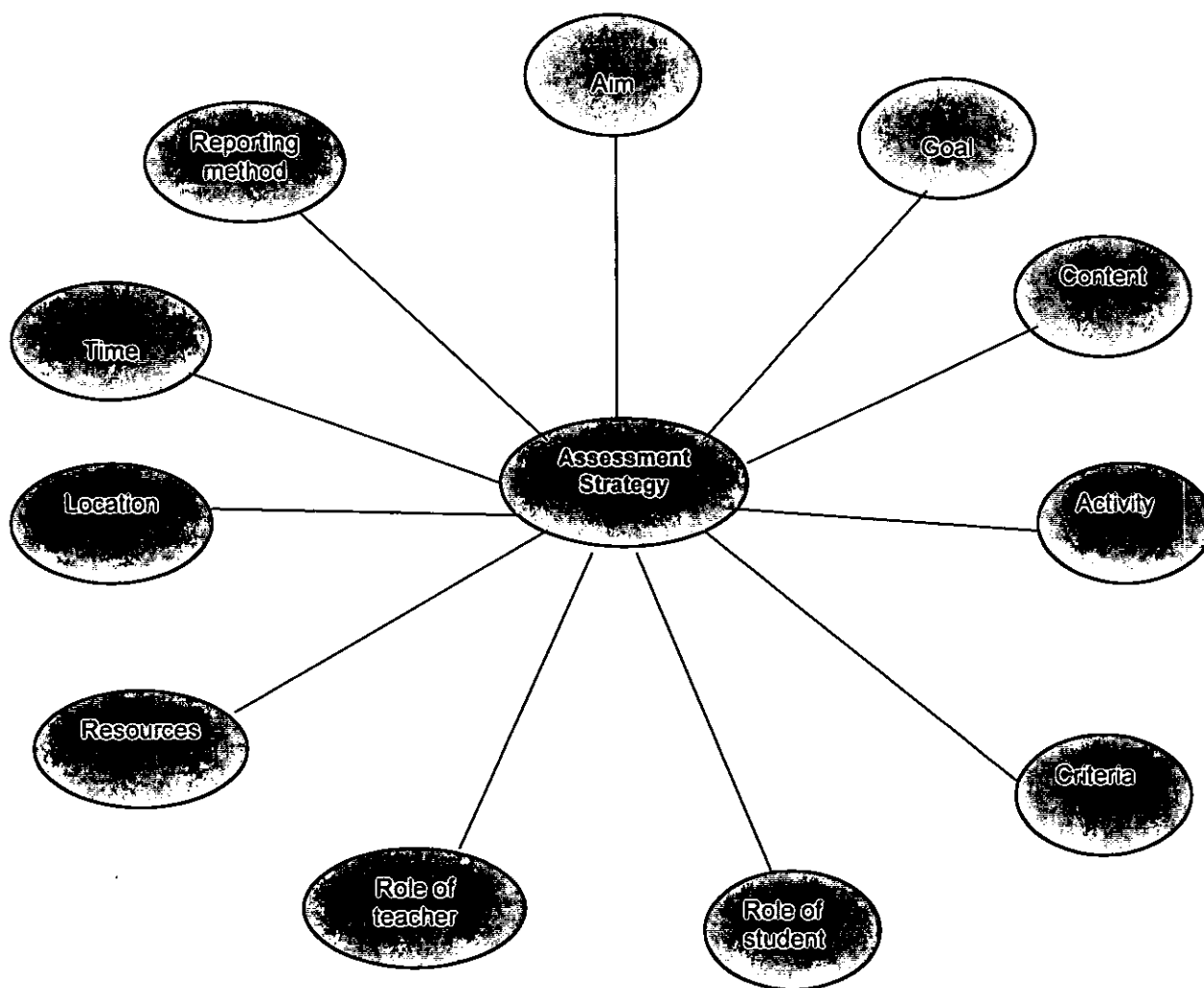
| Curriculum component | Focus question |
|-----------------------|---|
| Rationale | Why are the students learning? |
| Aims & Objectives | Toward which goals are they learning? |
| Content | What are they learning? |
| Learning activities | How are they learning? |
| Teacher role | How is the teacher facilitating learning? |
| Materials & Resources | With what are they learning? |
| Grouping | With whom are they learning? |
| Location | Where are they learning? |
| Time | When are they learning? |
| Assessment | How far has learning progressed? |

(Source: van den Akker, 2003:4)

The van den Akker curriculum spider web (Figure 4.5) shows the dynamic interactions of various components of a curriculum with the rationale at the centre of the spider. It is used in this study to describe the student-centred approach as the focal point to which the other

nine components are linked. The pertinent question under ‘rationale’ is ‘why are the students learning?’ and this means that the answer to this question has implications for the teaching and learning methods followed, as well as the materials and assessment strategies used. Therefore, the ‘rationale’ is the central mission in the learning process.

Similarly, since assessment practices are the focus of this study, for them to be successful, one may look at the various components of which each assessment can be composed, and thus, the assessment strategy is the central aspect of the assessment process. These components are visualised in the assessment wheel shown in Figure 4.6 (adapted from Howie, 2006).



(Source: Howie, 2006)

Figure 4.6: Assessment components

The wheel illustrates that, for assessment strategy to lead to effective learning, several aspects of classroom context must be taken into account and each one must support the other. These aspects are indicated in the wheel in an order (clockwise) which reflects the complete cycle of learning and assessment events advocated by Harlen (2006) and supported by van den Akker (2003). Students must understand what they are supposed to be learning and what is to be assessed and then have a confidence that teachers know these. This understanding includes clarity in aims, goal, content, activities, (assessment) criteria, as well as students' roles and those of the teachers prior, during and after the assessment task. The assessment strategy must also include important planning elements like materials (resources) with which students are supposed to be assessed, where (location) the assessment task will be taking place, and when (time). In the end, the way all these aspects are communicated to students (reporting method) is crucial for effective assessment of learning to take place. Having the assessment strategy at the centre of the wheel implies that the manner in which progress of the student learning can be assessed depends on all these various assessment components. Any changes in the assessment strategy will have to consider changes between its various components as this could affect the student learning.

The quality criteria of the PAM materials discussed in subsection 4.3.2.2 are inherently linked to the van den Akker's typology of curriculum implementation and Howie's interconnections of the assessment component. All these assessment components embody how the curriculum evolves in all its typologies (intended, implemented and achieved) and show the importance of linking assessment (assessing student learning), instruction (what is being taught and how), and curriculum (what should be taught). The validity aspect focuses on the intended curriculum, the practicality aspect focuses on the implemented curriculum, and effectiveness of the materials refers to the achieved curriculum.

In order to reduce the number of assessment components during the classroom tryouts, some adaptations and combinations of the components presented in this wheel have been made. These combinations also ensured that the assessment materials are user friendly for both teachers and students. Thus, in the assessment components of aim, goal, and location

some elements of why the teacher is assessing and in which context the assessment takes place, were added, which resulted in a new component named Rationale and Setting, while content included the student performance expectations and constituted another component called Content and Performance Expectations. Activity, roles of both teachers and students, and time were all embedded in one assessment component named Method, and criteria and reporting method were put into the component of Assessment. The component of resources stood alone as such and was renamed Materials and Resources to imply that these resources include not only books and pencils but also laboratory equipment. This resulted in five assessment components, which are listed below.

1. Rationale and setting: Why is the teacher assessing, toward which goals, and in which context is the assessment component being applied?
2. Content and performance expectations: What content, and on which intended learning outcomes is the assessment focused?
3. Method: (i) What are the activities of the students? (ii) What are the activities of the teacher? (iii) With whom are the students doing the assessment? (iv) At what time in the teaching and learning process is the assessment best applied?
4. Materials and resources: With what materials and resources are the students being assessed?
5. Assessment: How is the quality of the students' final product or task being judged?

Having presented and discussed the arguments in favour of employing certain assessment components in the Intervention Study, the following discussion is about guidelines for designing teacher support assessment materials. Howie (2006) cites Gronlund (1998:18) arguing that during the process of designing assessment materials leading to effective assessment one needs (i) to have clarity on all intended learning outcomes, (ii) a variety of relevant assessment procedures, (iii) fair procedures for all students, (iv) specified criteria for judging students' successful performance, (v) feedback to students that emphasise strengths and weaknesses, and (vi) a comprehensive system of grading and reporting.

Based on these arguments, and as also discussed in Chapter 3 (Section 3.5), it is advisable to design exemplary teacher support assessment materials that focus on four support

levels, namely subject knowledge, lesson preparation, teaching methodology, and assessment and feedback. In the context of this study these support levels are described below.

1. *Subject knowledge*

This level of support is specific to the topics under investigation. It includes making connections with other related topics and a contextualisation in relation to students' prior knowledge focusing on what may support or hinder the students' understanding of the studied topics. It is important that teachers make sure that all these aspects are dealt with before students start to perform the assessment task.

2. *Lesson preparation*

This includes advice to the teacher on the background of the problem that students are expected to solve. It also includes procedural specifications on investigative experimental work, the type of questions to ask while guiding students in the assessment task, and the necessary resources or equipment that students would need.

3. *Teaching methodology*

It includes advice on how to guide students in a student-centred approach for demonstration experiments. This refers to the roles of both the students and the teacher, which includes monitoring how students acquire content knowledge and practical skills.

4. *Assessment and feedback*

This support provides guidance on how to assess the products or characteristics of the students' activities and how to use the results as formative assessment feedback for future planning.

Because Physics support assessment materials are meant to change the teachers' routine and practice by turning their assessment (or even teaching) activities into a more investigative approach, these materials need to be:

- based on the objectives of Curriculum Plan for Secondary Education for Physics, Grade 12;

- developed from materials teachers already use;
- made to reflect clearly stated learning outcomes identified in the core curriculum that students are expected to study;
- designed with adequate support for teaching, and assessment strategies such as lesson preparation, subject knowledge, teaching methodology, and assessment and feedback;
- made to engage students, support curriculum implementation and instruction, improve student teaming, and to report individual student progress; and
- made to help teachers adopt a student-centred approach that includes investigative work and formative assessment.

Turning teacher assessment activities into an investigative approach in the context of this study means having teachers using not only already designed assessment materials but also participate in developing, trying out, evaluating and using their own assessment materials.

Taking into account the assessment components discussed earlier, the following is a discussion of the design specifications for the four support levels as applied to the context of designing the Physics assessment materials.

1. Subject knowledge

Two Physics concepts were the focuses of the prototypes namely force and *inertia* with the aim of helping teachers to assess their students formatively. The two topics were chosen for the following reasons. Firstly, the topics have been identified by the literature as sources of various student alternative conceptions or misinterpretations in many areas of physical Science. Many articles have discussed a number of student alternative conceptions about *force* as related to motion (Champagne et al., 1980; Clement, 1982; Dekkers, 1997; Gunstone, 1987; Thijs, 1987). Dekkers, for instance, lists 19 generalised student statements about situations involving force and the interpretations of their own beliefs expressed in those statements. Due to the fact that in many instances force implies an alteration of state of rest or motion of an object, these student alternative conceptions

constitute difficulties to the understanding of the concept of 'inertia'. Secondly, within the Grade 11 and 12 Physics Syllabus, these two topics have been given a great deal of attention. They are extensively taught in both grades and, added to the concept of energy, they appear to be the most difficult for students.

The teaching (and assessment) strategy proposed by this study to examine the students' understanding of the two concepts is Prediction-Observation-Explanation (POE) suggested by White and Gunstone (1992). This strategy requires students to carry out three different tasks. Firstly, they must predict the outcome of some event, and must justify their prediction. Secondly, they must see or perform a demonstration of the event and must describe what they see. Finally, they must reconcile any imbalance or conflict between what they predicted and what they have actually observed. Details entailing each of these tasks are discussed under teaching methodology's support level.

Having indicated the topics used as examples for the demonstration experiments, the rationale of choosing them, and the proposed teaching strategy, the following subsections discuss how these two topics can be introduced (following the proposed strategy) and presents some activities as examples.

a) Introduction to the concepts of force and inertia

At the beginning of this subsection the introduction of the force concept is discussed. There are various ways of introducing this concept. Some literature recommends starting the teaching of the concept following a cognitive conflict approach whereby the students' prior knowledge and understanding (including their alternative conceptions) on the related subject matter is probed (Clement, 1993; Dekkers, 1997; Mutimucuo, 1998; White & Gunstone, 1992). Firstly, and following the recommendation, the probing can start by giving students some examples of a variety of forces, such as gravity, normal or supporting forces, forces in collisions, and forces of springs. At this teaching stage, students do not yet have a clear notion of interactions between pairs of objects but this can later be given attention at Newton's Third Law of motion. Secondly, students can be helped to understand the concept of force in two conditions: at rest and in motion. When

an object is at rest, the sum of all acting forces is zero. For the condition of motion, the sum of all acting forces can be zero (uniform motion) or different from zero. For this study, the activity describes a laboratory demonstration experiment of a moving object with uniform motion. The aim of the experiment is to identify and compare forces acting on moving objects following the POE strategy. This leads to the formulation of Newton's First Law, according to which, when the sum of all forces is zero, an object at rest or in uniform motion in a straight line will continue in its state, unless it is compelled to change that state by external forces acting upon it.

Two experiments by Galileo can be used to introduce the physical significance of Newton's First Law of motion. In one experiment, and studying the motion of a sphere moving on a horizontal surface, Galileo observed that, if the sphere were pushed with a given force, it would move through a certain distance before stopping (Figure 4.7).

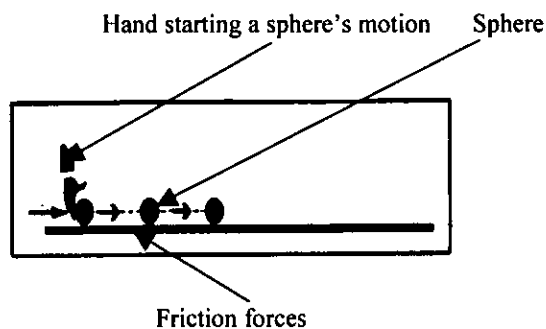


Figure 4.7: A sphere set in motion

When analysing the experiment, it becomes relevant to find out why the sphere comes to rest some time after the pushing force has stopped acting. The reason is that in any moving object (pushed or pulled) there are acting opposing forces, which are impeding the movements. These forces are acting between the moving object and the surface where the object is and they are called 'friction forces' (see Figure 4.8).

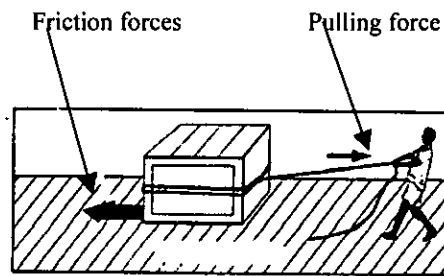


Figure 4.8: Forces on a moving object

Friction forces always act where there is a contact between the moving objects and the surface on which they move and they are forces opposing that movement. In the example of the experiment on Figure 4.7, the friction force acts between the sphere and the horizontal surface and is opposing the movement of the sphere.

In another experiment, Galileo decided to polish the surface on which the sphere was moving to see whether the characteristics of the movement would change. In this case, he realised that the sphere had traversed a bigger distance compared to the previous experiment, before the surface was polished. Then Galileo came to the conclusion that, if it was possible to completely eliminate the force that tends to oppose the movement of the sphere (the friction force), the sphere, after experiencing the action of the initial force (with the push), would continue moving with constant speed in a straight line. From Galileo's conclusions, Newton formulated his First Law of motion, known also as Law of Inertia, as follows:

Every object continues in its state of rest, or of uniform motion in a straight line, unless it is compelled to change that state by external resultant forces exerted upon it.

The key word in this formulation of Newton's First Law is 'continues': an object continues to do whatever it happens to be doing (rest or uniform motion in a straight line) unless an external force is impressed upon it. If it is at rest, it will continue in a state of

rest. If it is moving, it continues to move without turning or changing its speed. Objects at rest tend to stay at rest – objects moving tend to continue moving. This is the physical significance of Newton's First Law of motion and this tendency of objects to resist changes in motion is called 'inertia'.

Having discussed the introduction of the force concept, the following subsection discusses how inertia can be introduced. Drawing from the previous discussion about force and once Newton's First Law is clearly understood, the teacher can easily introduce the concept of inertia. The tendency of objects to remain in its state of rest or of uniform motion in a straight line is called *inertia*. The inertia of objects is observable when (i) an object at rest is suddenly set in motion or (ii) when an object animated with uniform motion in a straight line has the value of its speed or its direction changed. This change is caused by an external influence, which means that the resultant of all forces acting on the object is different from zero. These resulting forces different from zero will cause acceleration. As a conclusion, the inertia of an object will be observable only if the object is accelerated.

Newton's First Law is another way of showing that all matter (objects) has a built-in opposition to being moved if it is at rest or, if it is moving, to having its motion changed. This property of matter is called inertia. The effect of inertia is evident, for instance, on the occupants of a car which stops suddenly. The occupants will be lurched forward in an attempt to continue moving. The larger the mass of an object, the greater is its inertia; that is, the more difficult it is to move, when at rest, and to stop, when in motion. So, the mass of an object is the measure of its inertia.

Some activities on both force and inertia can be suggested for students to carry out following the POE strategy. Below is an example on how to identify and compare forces on moving objects.

b) An introductory activity using the POE strategy (e.g., identification and comparison of forces acting on moving objects)

A prerequisite for the demonstration experiment is that students should previously have used spring balances to measure forces, and that uniform motion from the kinematical perspective has been discussed. As a first experiment, the POE strategy recommends that the lesson can start by asking students to name all forces acting on a soccer ball kicked into the air (prediction). After this exercise they can be asked to kick the ball (as shown in Figure 4.9), observe its behaviour through its entire trajectory, and name all the forces acting on the ball (observation).

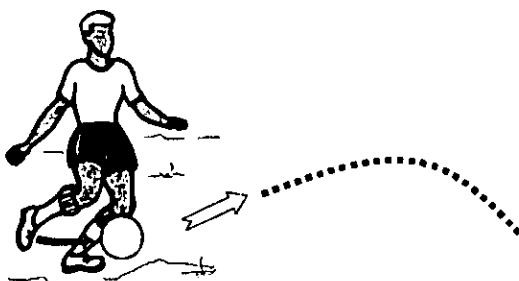


Figure 4.9: A soccer ball kicked into the air

In the end, they must be asked to compare what they have predicted and what they actually observed during the experiment (explanation).

In a second experiment the students can be shown the set-up depicted in Figure 4.10. A trolley is placed on a smooth runway, with spring balances attached to front and back. At the back, a hanging mass is attached to the balance by means of string and pulleys. At the front, the trolley is pulled forward by hand (F_{pull}).

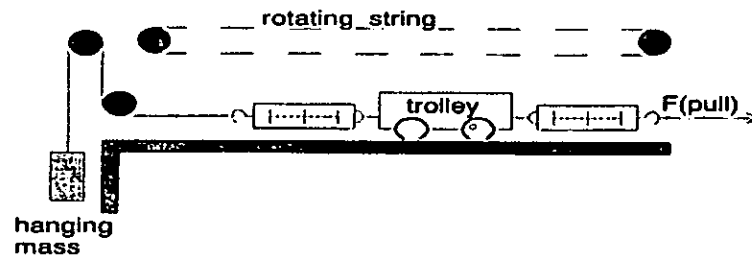


Figure 4.10: Identification and comparison of forces acting on moving objects

Predict: Using the POE strategy, students are asked to predict, individually, how the forward and backward force will compare, if the trolley is pulled forward at constant speed. They are also asked how these forces will compare, if subsequently a bigger constant speed is chosen.

Observe: After the predictions, the experiment is carried out, preferably with the students in small groups. The hanging mass should be big enough for the friction to be negligible. This will ensure that the same forces forward and backward are found at all constant speeds ranging from 0 to 2 m/s (meters per second). A constant speed is obtained pulling the trolley along with a knot in a string revolving above the set-up, which is propelled by an electrical motor from a cassette player. Using various gears, different constant speeds can be obtained.

Explain: The result of the observation is compared with what is found in the textbooks about the behaviour of the forces acting on moving objects and it can therefore be named Newton's First Law of motion. The result of the experiment can then be discussed in the class: 'What did the students think would happen?', 'Which are the acting forces?', 'Why were the backward and forward forces equal?'

It is important to mention that, although these experiments are useful in identifying and comparing forces, they cannot necessarily be used to combat potential students'

alternative conceptions on forces. They only offer students the empirical evidence that their expectations and their alternative conceptions can be deficient. They do not provide explanation about what students already know that can guarantee the shift from their own conceptions to the scientific view.

2. Lesson preparation

Teacher support on the preparation of lessons in which the formative assessment will take place includes advice from the teacher on the characteristics of the lesson and student readiness in terms of the background of the problem that they are expected to solve. General characteristics of the Physics demonstration experiments include content specific knowledge (e.g., description of the intended learning outcomes) and procedural specifications of laboratory work (e.g., materials required for the experiments, the timing of the activities, and suggestions on how to deal with potential problems that may occur). In general, the teacher support for lesson preparation includes two main aspects.

a) General description of the lesson

- Firstly, this involves a description of the main concepts (force and inertia) to be dealt with during the experiments and how they will be formatively assessed. The teacher may start the lesson by asking brief questions to students on what they already know about concepts like mass and speed. A discussion about these concepts may help the teacher to understand and evaluate student predictions and/or their responses during experiments.
- Secondly, there is a description of what constitutes the lesson (for instance, that the students will work in groups of a maximum number of four students each).
- Thirdly, there is a description of the intended learning outcomes based on the Physics Syllabus for Grades 11 and 12 - the aims of the lesson must be clearly formulated by the teacher (emphasising the POE strategy) to help students understand what is expected of them.

b) Lesson preparation

- Explain what is to be done and when (lesson plan and timing).

- Explain the working method: for example, explain that everyone in the classroom must be organised into groups of a maximum number of four students each.
- Anticipate potential difficulties during the experiments associated with the student-centred practice.
- Locate the materials (equipment) required for the experiments. This activity is important to make the teacher aware that the experiments are designed to support student-centred practice using locally available materials.

3. Teaching methodology

As referred to earlier, the teaching methodology used to teach and to investigate student understanding of the two Physics concepts is the *Prediction-Observation-Explanation* (POE) strategy. One of the most powerful contributions of the POE strategy to learning is that it is more direct in revealing students' understanding than the usual style of verbal or paper-and-pencil tests. It focuses on a specific phenomenon of learning. The prediction that is required from students is more likely to imply genuine application of the previous knowledge rather than asking a simple question in the form "explain why...". Furthermore, the students are more likely to evaluate how their knowledge applies to a real situation, because the experiment is directly shown, than the more general thinking implied by a single question. Another key characteristic of the POE strategy is that it allows students to decide what reasoning they must apply in any given situation, whilst their predictions are based on their everyday experiences and beliefs.

Besides the three main tasks that students are required to carry out (predict, observe, explain), some critical steps are to be taken into account when applying this strategy (White & Gunstone, 1992). The first step is that teachers must ensure that all students understand the nature of the situation about which they are supposed to make a prediction. Teachers should also ensure that all students have the same understanding of the situation before proceeding. This can be done by encouraging students to ask questions about the situation under consideration. The second step refers to the importance of having all students indicate, in writing, both the prediction and the reasons supporting the prediction.

This is important because it allows the students to decide what knowledge is appropriate to apply and then apply it. The way students must indicate their prediction and the supporting reasons can be either in open-ended form, i.e., having students writing their predictions on their own words, or on a previously prepared sheet of paper. Recording the reasons for prediction is crucial for the value of the teaching strategy because it shows the link between the concepts involved in the learning situation. The third step occurs during the actual experimentation. It is important that all students write down their individual observations while the experiment takes place. Very often different students will see different things, and if observations are not written down at the time they are made, some students might change their observations as a result of hearing what others claim to have seen. The fourth and last step refers to the students' reconciliation of any discrepancy between what they predicted and what they actually observed. Normally this is difficult for students, but it is advisable because students' explanations at this stage reveal much about their understanding.

4. Assessment and feedback

The ultimate objective of this support level is defined in terms of assessment *of* learning and thus, the emphasis is on all functions of assessment namely diagnostic, formative, and summative. The diagnostic and formative functions of assessment are expressed by a number of design guidelines to facilitate experimental work and of elements of feedback provision that the teacher needs to consider. These design guidelines, as taken from the literature (Dekkers, 1997; Garrett & Roberts, 1982; Gunstone, 1991; Tamir, 1991; van den Berg & Giddings, 1992), are listed below.

- Agreement – having stated the problem to be investigated, the teacher and the students must agree on the procedures to be followed, the evaluation of the explanations given during the experimental work, and the conclusions.
- Learning outcomes – the teacher must be tightly prescriptive about the ideas that the students are supposed to acquire and develop.

- Student participation – In practical work, particularly in laboratory demonstrations, the teacher must produce the event to be investigated according to the purpose to be achieved, while the students attempt to interpret it and make sense of it.
- Type of experiment and aims - Teachers must avoid having too many aims of the experiment to be achieved at once. This may lead to none being pursued.
- Critical thinking and reporting – Teachers are to make sure that students develop a critical attitude towards their actions and interpret the activity's data only in the light of the experimental work pursued and of their own knowledge and experience.

As for providing formative feedback to students, when facilitating demonstration experiments, teachers must consider a number of elements of feedback provision in three main stages of the lesson, namely (i) in lesson preparation, (ii) in the course of the lesson, and (iii) in the end of lesson (Motswiri, 2004). These elements are presented and discussed in Chapter 6 (Section 6.5).

Concerning the summative function of assessment, the use of assessment criteria is crucial to help monitor the performance of students. The criteria to be adopted for assessing student understanding of the inertia concept using laboratory experiment must be such that they provide information about how students performed the task at the end of the experiments. Scoring rubrics are used to assess the student responses to the performance task. These rubrics are observable in nature, and there are specific aspects a student should perform to carry out a performance properly.

Two types of scoring criteria are frequently discussed in the literature, namely analytic and holistic (Moskal, 2003). Analytic scoring rubrics divide a performance into separate facets and each facet is evaluated using a separate scale (see, for example, the different performance levels of the rubric on assembling an electric circuit in Table 3.2). Holistic scoring rubrics use a single scale to evaluate the larger process. In order to develop observable scoring criteria for the POE strategy, analytic scoring rubrics were considered

(refer to Appendix P, Table 2) and the guidelines discussed below were taken into account (Moskal, 2003:15).

- *The criteria set forth within a scoring rubric should be clearly aligned with the requirements of the task and the stated goals and objectives.* A list can be compiled that describes how the elements of the task are into the goals and objectives. This list can be extended to include how the criteria set forth in the scoring rubric map into both the elements of the task and the goals and objectives. Criteria that cannot be mapped directly back to both the task and the goals and objectives should not be included in the scoring rubric.
- *The criteria set forth in scoring rubrics should be expressed in terms of observable behaviours or product characteristics.* A teacher cannot evaluate an internal process unless this process is displayed in an external manner. For example, a teacher cannot look into students' heads and see their reasoning process. Instead, it is necessary for students to explain their reasoning in written or oral form and the scoring criteria should be focused upon evaluating the written or oral display of the reasoning process.
- *Scoring rubrics should be written in specific and clear language that the students understand.* One benefit of using scoring rubrics is that they provide students with a clear description of what is expected *before* they complete the assessment activity. If the language employed in a scoring rubric is too complex for the given students, this benefit is lost. In other words, students should be able to understand the scoring criteria.
- *The number of points that are used in the scoring rubric should make sense.* The points that are assigned should clearly reflect the value of the activity. On an analytic scoring rubric, if various facets are weighted differently from other facets of the rubric, there should be a clear reason for these differences.
- *The statement of the criteria should be fair and free from bias.* The phrasing used in the description of the performance criteria should be carefully constructed in a manner that eliminates gender and ethnic stereotypes. Additionally, the criteria

should not give an unfair advantage to a particular subset of students and which is unrelated to the purpose of the task.

4.3.2.4 Research procedures for the intervention study

This subsection presents the research procedures for the Intervention Study. It describes the number of activities carried out during the design and development phases of the PAM prototypes and the period in which these activities took place. Details about how appraisers of the materials and the participating schools, teachers, and students were prepared for their roles in the intervention are all presented in different sections in Chapter 6 according to their level of involvement (refer to Sections 6.2 to 6.5).

The Intervention Study consisted of a cyclical development of the PAM prototypes in four versions and was undertaken between April 2006 and February 2007. The first version of the PAM prototype was designed by the researcher based on (i) lessons from Baseline Survey findings and (ii) design guidelines and specifications of exemplary assessment materials described earlier (subsection 4.3.2.3), which were adapted from design specifications used by similar intervention studies in Science education (refer to Chapter 3, Section 3.5). This version was appraised by three experts, three university students, and four secondary school teachers. The design and appraisal of this prototype was carried out from April to June 2006. The design and development of the second version – also by the researcher - was undertaken between July and September 2006 and followed suggestions and recommendations from appraisers of the first version. This version was then tried out in the classroom with two teachers and 62 of their students in October 2006. Then the analysis of the tryout findings, which started during the tryout in October, was finalised earlier in November 2006. The findings led to the design - also by the researcher - of the third version of the prototype. It then followed a final appraisal by two experts between November 2006 and January 2007. Finally, the fourth and final version of the prototype was designed and evaluated in a workshop with three university students and two teachers in February 2007.

4.4 Validity and reliability

No research study is perfect. However, controlling the possible threats, which might interfere with the interpretation of the cause-effect relationship, is crucial (Coolican, 1999). Understanding the concepts of validity and reliability helps to analyse the possible weaknesses derived from uncontrolled variables particularly in experimental research. The general definition of validity is that it is a demonstration that a particular instrument does indeed measure what it is supposed to measure (Cohen et al., 2000). Reliability is defined by Cohen et al., (2000) as a synonym for consistency and replicability over time, over instruments and over groups of respondents.

For this particular study, three means were considered to establish validity. The first was face validity where for both research phases (baseline and intervention) the validity was checked in all the corresponding data collection strategies, namely questionnaires, interviews schedules, classroom observation schedules, including the evaluation instruments of the classroom tryouts. The validity was checked by inspecting whether the instruments indeed measure what it is supposed to measure in terms of level and breadth.

The second was content validity (Cohen et al., 2000) of the instruments, which was controlled through consultation with colleagues, teachers, experts, and also by relying on the researcher's experience to ensure the representativeness of the researched area.

The third was external validity (Yin, 1994) of the intervention, which deals with the problem of knowing whether the demonstration experiment results can be generalisable to a broader perspective. This is a particularly difficult notion to achieve for a case study like the one reported in this dissertation. Although no statistical generalisation is possible from the sample involved in the study to the population of Mozambican Grade 12 Physics teachers and students, yet this study strives to generalise the findings to the broader theory underlying the design and development of the intervention. Yin (1994) speaks in this context of analytical generalisation. There will have to be more replications of these findings in more classroom tryouts to determine whether the same results may occur.

In terms of reliability the issues of internal and external reliability were addressed. Internal reliability or consistency was verified to ensure that questionnaire respondents answer related items in similar ways. External reliability or stability was verified by cross checking information on assessment practices used by teachers by comparing the information provided in the questionnaires to that of the interviews and including the classroom observations.

4.5 Ethical issues

The subject of ethics in social research is potentially a wide-ranging and challenging one. Therefore, it was fitting for this study to address the main issues that may confront the researcher in the field. However, before discussing the issue of confrontation with the field, including the researcher's own integrity and transparency, it is important to state that the ethics requirements, as prescribed by the University of Pretoria, were met. Prior to the research, permission were sought from the Faculty of Education regarding ethical considerations involved in the study. The procedures suggested were approved by the Faculty and permission to undertake the study was granted.

Regarding the issue of field confrontation and researcher's integrity, four aspects deserved consideration for the course of the research.

a) *Debriefing and right to non-participation* (Coolican, 1999) – All participants in this research were informed prior to the research about the full nature and rationale of the study they were to be involved in, and there was an effort to avoid any negative influences. The researcher had to emphasise the voluntary nature of the participation, as well as the right of the participants to withdraw at any time should the discomfort be greater than anticipated.

b) *Confidentiality and privacy* - The researcher guaranteed anonymity or requested permission to identify individual participants. For example, when the use of tape recordings appeared to be necessary, permission was sought. Interviewees and teachers observed during classes and those who participated in the tryouts were all asked for their permission to be identified. For the particular aspect of recorded interviews, the

participants were given the right to assume that the records of the interviews would be safeguarded and used as anonymous data only for research purposes.

c) Intervention (Coolican, 1999) - Since the design and development of assessment prototypes was an activity that caused alteration to the teachers' normal routine, there was a need to improve some working conditions. For instance, coffee and snacks were arranged for the teachers in such situations where they had to stay in school much longer than their normal time schedule.

d) The role of the researcher (Plomp, 2006) - This research was conducted in close collaboration with teachers and students who were actively involved, often as members of the research team. The situation led to problems of finding a balance between the role of the researcher as a designer, an evaluator, and an implementer. Making the research open to scrutiny and critique by educational experts, deserving attention to validity and reliability of data and instruments, and having a good quality of research design appeared to be key measures for this potential conflicting role. For instance, the quality of the design was sought through triangulation (of data and its analysis), empirical testing (of the intervention), and systematic documentation, analysis and reflection (of the design, development, and evaluation of the intervention process).

4.6 Conclusion

The research design of the study on the investigation and improvement of assessment practices used by Grade 12 teachers in Physics consists of two main stages. The first stage, the Baseline Survey, involves the identification of assessment practices currently used by the teachers in Mozambique. The survey is based on the context in which Physics teachers are working in schools as well as on the insights of the literature on what is deemed to be good classroom practice. The literature review (as presented in Chapter 3) focuses on the role, current practices, and ways of improving the teacher assessment practices in secondary Science education as surveyed in many educational systems in both developed and developing countries. Expert appraisal and networking with people working in similar fields have also added value to the preliminary research of the overall study. This stage led to the necessary groundwork to the following stage of the research study. The overall findings of this stage are discussed in Chapter 5.

The second stage of the research design is the Intervention Study, which consists of the design and development of prototypes for assessing the performance of Grade 12 Physics students as a way of helping teachers to improve their classroom practice. The prototypes consist of demonstration experiments and a written report and they were designed and developed on the topics of force and inertia, selected from Grade 12 Physics Syllabus for Secondary School in Mozambique. These topics were chosen because of their suitability to apply the POE strategy in inquiring student understanding of Science concepts, particularly for performance assessments (to be discussed in Chapter 5). The development of the prototypes uses a cyclic approach of design and formative evaluation in such a way that successive versions of the material evolve into a final product with empirical evidence of its practicality. The validity, expected practicality and expected effectiveness of the material were verified through appraisal from curriculum, science, and assessment specialists and tryouts by potential users, i.e., teachers and students. The design and appraisal of the subsequent versions of the prototypes, as well as the findings of the classroom tryout, are presented and discussed in detail in Chapter 6.

CHAPTER 5

**ASSESSMENT PRACTICES OF
MOZAMBICAN PHYSICS TEACHERS**

CHAPTER 5

ASSESSMENT PRACTICES OF MOZAMBICAN PHYSICS TEACHERS

This chapter reports on the findings of the Baseline Survey aimed at investigating assessment practices used by Grade 12 Physics teachers in the classroom and designed to inform the Intervention Study. Section 5.1 presents the introduction to the chapter where the aim of the Baseline Survey is formulated and discusses the preliminary literature findings about what is already known on teacher assessment practices. Section 5.2 presents and discusses the main findings of the Baseline Survey with emphasis on the assessment practices used by teachers, their frequencies, their quality, and their relevance for student learning. The section also reveals the reflections drawn from interviews with school directors, pedagogical officers and assessments experts. Finally, the conclusions and recommendations as to what aspects can be taken into account when preparing the intervention phase in order to produce improvements in the teachers' practices are presented in Section 5.3.

5.1 Introduction

The aim of this study is to investigate and improve the assessment practices used by secondary school Physics teachers in Mozambique. As was discussed in earlier chapters (refer to Chapter 1, Section 1.2 and Chapter 4, Section 4.3) in order to address this aim, it was necessary to start by undertaking a preliminary survey aimed at identifying assessment practices currently used by Grade 12 Physics teachers in Mozambican schools. Specifically, a Baseline Survey was carried out to investigate what assessment practices Grade 12 teachers currently apply in Physics at classroom level.

Some studies have investigated the assessment strategies used by Mozambican teachers (INDE, 2005; Lauchande, 2001) and by other teachers elsewhere (Race et al., 2005; Popham, 2002) when assessing science subjects. According to these studies, and as has been referred to in Chapter 1 (Section 1.2), generally the present assessment practice in Mozambican education is fundamentally based on expecting from students the memorisation of concepts, formulas and mechanisation of procedures due to the teachers' weak scientific and pedagogic competence, and to the lack of skills in developing appropriate assessment instruments. This situation is more accentuated in experimental subjects, such as Physics, where teachers do not teach, and hence assess, the student abilities to manipulate, to observe, to generalise, and to establish relationships. According to these studies, this is due, on the one hand, to the teachers' weak preparation for teaching and assessing these abilities, and on the other hand, to the lack of teaching material and equipment such as microscopes and some other electric appliances. Apart from these teachers' difficulties, the studies have also recorded some student weaknesses. Students are seen to have difficulties in answering essay-type questions as their writing skills are inadequately developed. As is the case with other subjects, in the assessment of Physics greater attention is given to marking and grading, instead of providing formative feedback, much of which tends to lower self-esteem of both students and teachers. Not enough advice is provided for the improvement of both learning and teaching. This is the baseline information available from the very few studies conducted in Mozambique in this regard.

Internationally, Race et al., (2005) for instance, report that there is an overemphasis on norm-referenced type of assessment, where students are compared with each other and this leads to demoralisation among the less successful students. Peer-assessment amongst students, i.e., the way teachers are supposed to allow students to assess the work of their colleagues, is also another weak point in Physics classroom assessment. In this regard, Race et al., (2005) point out a number of assessment strategies that can be used by teachers to involve students in peer-assessment. These strategies include (i) student presentations, (ii) reports, (iii) essay plans, (iv) calculations, (v) interviews, (vi) annotated bibliographies, (vii) practical work, (viii) poster displays, (ix) portfolios, (x) performances, and (xi) exhibitions and artefacts. Only a very few of these strategies are

used in schools as a means of allowing students to assess their own work and that of the others. Still according to these authors, teachers' feedback to students often serves managerial and social purposes rather than helping them learn more effectively. For instance, during Physics lessons, teachers tend to give feedback to students individually and not to groups, and this situation does not allow students to learn from the successes and failures of others. This situation is also seen to prevent them from questioning the teacher and challenging his/her comments.

Within this framework, the present Baseline Survey was aimed not only at finding out what assessment strategies are used by teachers in schools but also at reporting on their quality and relevance for assessing Physics. This information served as a platform to start an Intervention Study aimed at realising improvements of the teacher assessment practices. This chapter presents and discusses the fieldwork findings of the following operational research question: *What assessment practices do Grade 12 teachers in Physics in Mozambique apply and what is their quality?*

During the Baseline Survey, both qualitative and quantitative methods were applied to collect and analyse data with the main aim of understanding the characteristics of the current teacher assessment practices, and to produce valuable information needed for designing the Intervention Study. In fact, the above-mentioned operational research question is the starting point for designing and developing an appropriate intervention. This reasoning is in line with Creswell's (2003) research paradigm used to guide the present study and referred to in Chapter 1 (Section 1.3).

5.2 Findings of the Baseline Survey

As has been mentioned in the Survey research design (Chapter 4, Section 4.3) three elements were used as perspectives against which the characteristics of assessment practices used by teachers in schools are described, namely the *types* of assessment practices applied (defined by their frequency), their *quality*, and their *relevance* for learning. These elements were used to guide the formulation of the operational research

questions of the Baseline Survey. This subsection presents and discusses the findings of the Baseline Survey according to these elements and follows the sequence in which the findings answered the corresponding operational research questions.

5.2.1 Assessment practices applied and their frequencies

According to the literature review (Chapter 3), several authors have indicated a number of assessment strategies that teachers normally use in schools. The aim of the Baseline Study was to identify which assessment strategies Mozambican Grade 12 teachers use when assessing their students and what can be said about the quality of these. Findings were derived mainly from the questionnaire administered to twelve teachers including four school directors. The main question addressed to teachers in order to identify the *types* of assessment practices and viewed from the perspective of their frequency was: *How often do you use each of the following assessment practices: portfolios, peer-assessment, verbal tests, homework, paper-and-pencil, projects?* Answers were elicited and the findings were supported by interviews with all twelve teachers, with two pedagogical officers and by eight classroom observations by the researcher. The main findings summarised in Figure 5.1, related to the assessment practices commonly used by teachers in schools and their frequencies indicate that paper-and-pencil tests, verbal tests and homework are the most frequently used assessment practices in schools.

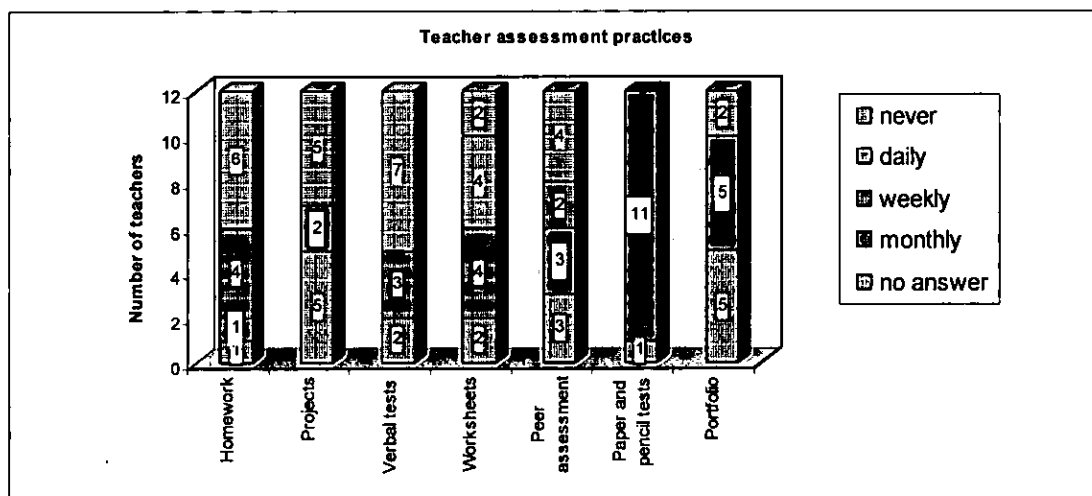


Figure 5.1: Teachers' responses by type of assessment practice (n=12)

Eleven (n=12) teachers said they were using paper-and-pencil tests (one did not respond), and grading homework (six daily, four weekly, and one monthly). Verbal tests were used by ten (n=12) teachers (seven daily and three weekly). Verbal tests are oral questions that teachers ask during the course of the lesson to students and which require immediate answers. The less frequently used assessment practices were projects and peer-assessment. Five teachers said that they never assessed by means of projects and the same number of teachers did not respond to the question. Only the remaining teachers (two) said that they assess through projects monthly. Four teachers never used peer-assessment.

A first analysis of these results raised a concern about the number of missing data in the questionnaire. A large number of teachers did not provide answers to some of the assessment practices: five teachers (n=12) for projects, the same number for portfolios assessment, and three teachers for peer-assessment. This was most likely due to the teachers' poor understanding, sometimes lack of it, of the different assessment practices. Therefore, there was a need to verify what could have been the cause. The findings were triangulated by the interviews with the aim of clarifying the unanswered questions and also eliciting the answers provided by the teachers. During the interviews, teachers were asked to explain their understanding of some of the assessment practices and to say whether they used that type of assessment in their classrooms.

The answers showed that some of the teachers did not use some assessment practices because they were not familiar with that particular practice. In addition, they showed a lack of clarity between what they do in the classroom and what they were actually being asked by the researcher. For example, the following quotation from an interview with a teacher is evidence of the level of teachers' understanding (or lack of it) about portfolios assessment:

[Male, 41 years old, Pemba Secondary School, Cabo Delgado Province]

Interviewer (Ivr): what do you understand by portfolio assessment?

Teacher (Tch): (...) silence...sorry I did not understand the question.

Ivr: what is your understanding of portfolio assessment?

Tch: Huumm,...I do not know about this.

Ivr: Ok. Have you ever determined a student's progress through his/her particular work?

Tch: Ah... yes. We always do this at the end of each semester when we meet as the Physics group of teachers to evaluate the performance of individual students' work and discuss their final scores. We do this twice a year.

Other evidence of teachers' lack of understanding of some of the assessment practices was found in relation to peer-assessment. Some of the teachers expressed the idea that they feel uneasy using peer-assessment because they feel students are not supposed to assess their own work or that of their peers because they think they are all students and that is why they are in school – just to learn. In an interview about the use of peer-assessment one teacher said:

[Male, 28 years old, Chókwè Secondary School, Gaza Province]

I do not think it's a good idea at all (doing peer-assessment in the class). How can you expect somebody to assess other somebody's work without knowing him or herself about it? ...furthermore, I think students... will never feel easy doing this job! You know... they're friends and would not like to give each other's low marks ... just in case if the work was wrongly done. Better I do it myself; I am the teacher, after all.

These answers show that teachers (at least some) were not quite familiar with certain of the assessment practices such as portfolios and peer-assessment, and they lacked insight into the value of using these when designing assessments and administering them in the classroom. Judging from the evidence produced by the interviews, it seemed that this lack of understanding could be the reason for high number of teachers not assessing students using projects or portfolios.

It is worthy mentioning that in Chapter 4, it was referred that a piloting of the all instruments was undertaken (subsection 4.3.3), and that the internal reliability was carefully established in order to ensure that the respondents answered related items in a similar way (refer to Section 4.5). However, this seemingly unfamiliarity of the teachers with some assessment practices was probably not detected during the pilot process because the level of exposure to educational resources (e.g., libraries and bookshops) and

events (seminars, workshops, in-service courses, etc.) decreases as one moves from Maputo to the North. In fact, the first excerpt quoted above is from a teacher from Pemba Secondary School, which is located in the northern region of the country about 2,500 km away from Maputo. However, in order to limit the costs, the pilot phase was undertaken in Maputo schools only, while the Baseline Survey took place in schools countrywide, excluding Maputo. This was a deliberate choice, as one of the aims of the study is to help the most disadvantaged teachers from the provinces. It might, however, be possible that some of the Maputo teachers had also some difficulties in understanding some of the assessment practices but, for some reason or another (not explained by the study findings), the pilot process did not uncover this. In the end, however, it is fair to acknowledge that the decision may have influenced the validity of the instruments at some extent.

Another concern resulting from the analysis of the data is the lack of consistency in teachers' responses from the questionnaire to the interview. For instance, during the questionnaire two teachers said that they were using projects, five of them used the portfolios, and the same number used peer-assessment. However, during the interviews, ten teachers said they had never assessed using projects, eleven said that they had no idea of what a portfolio was, and nine said that they had never used peer-assessment. Table 5.1 shows the difference between teachers' responses to the questionnaire and what emerged from the interviews.

Table 5.1: Consistency level of teachers' responses between questionnaire and interview (n=12)

| | Portfolios | | Projects | | Peer-assessment | |
|----------------------|---|-----------|---|-----------|--|-----------|
| | Questionnaire | Interview | Questionnaire | Interview | Questionnaire | Interview |
| Yes | 5 | 1 | 2 | 2 | 5 | 3 |
| No | 2 | 11 | 5 | 10 | 4 | 9 |
| No answer | 5 | - | 5 | - | 3 | - |
| Level of consistency | Among the eleven teachers who said NO in the interview, five were those who did not respond in the questionnaire, two are the same who said NO in the questionnaire and four changed their minds from the questionnaire to the interview. | | Among the ten teachers who said NO in the interview, five were the same who also said NO in the questionnaire and the other five were those who did not respond in the questionnaire. The two who said YES, stuck to their answers. | | The number of teachers who said NO in the interview (nine) is made up of the same four who have also said NO in the questionnaire, plus three of those who did not respond. Two of them changed their minds from the questionnaire to the interview. | |

The first impression arising from these data reinforces the idea that, most probably, the amount of missing data in the questionnaire reflects the teachers' poor understanding of the different assessment practices. For instance, as far as the portfolios is concerned, all of the five teachers, who did not respond in the questionnaire to the question about using portfolios, quite openly said they never used portfolios, and still four out of the five who said they used portfolios during the questionnaire appeared to change their minds in the interview when they had a better understanding of the concept. Teachers changed their minds after having been informed during the interview what a portfolio was. The same applies to projects and peer-assessment where all the missing data from the questionnaire turned into NO answers.

Although one can conclude, after the interviews, that the majority of teachers seem not to use these assessment practices, the consistency of responses that were given after an explanation from the interviewer still remains questionable. It is not clear whether the

teachers' responses were honest or whether they wanted to please the interviewer. Furthermore, the teachers' unfamiliarity with these assessment practices, which was apparent during the interviews, reinforces the lack of consistency in their responses.

As for the most frequently used assessment practices (paper-and-pencil tests, verbal tests and homework), there were no substantial differences between teachers' responses from the questionnaire to the interview. All of the missing data turned into YES responses, most probably because these are seemingly more popular assessment practices amongst teachers.

So, it would seem that the teachers' most consistent results, in terms of the types of assessment practices, are paper-and-pencil tests, verbal tests, and grading homework where the majority of teachers (more than ten teachers out of twelve) said that they use these assessment practices daily, weekly or monthly.

5.2.2 Quality of assessment practices

One of the ways used to investigate the quality of assessment practices used by teachers in schools was to ask teachers the following question: *How often do you assess the following students' activities: oral communication during lessons, written work, presentations, notebooks, laboratory work, solving problems?* In fact, a frequency in itself is not a good indicator for quality because a teacher can frequently use certain assessment practice of poor quality. However, available literature indicates that classroom assessment in the Mozambican context is exam-driven and constantly using a certain assessment practice can reveal the teacher's interpretation of the quality of that practice as seen by the MEC (Lauchande, 2001; Palme, 1992; Popov, 1994). This information helps to understand how the quality of classroom assessment can be influenced by external examinations as it was later confirmed by the pedagogical officers during interviews. Furthermore, it was also expected that by asking teachers to reflect on these different student activities it would be possible to understand the validity (in terms of content) of the assessments being undertaken by and with students (Chapter 4, Section 4.3).

This is an important starting point for understanding and interpreting the study findings. A 5-point scale was used to characterise the frequency of the use by teachers of the activities: no answer, always (1-2 times per week), frequently (1-2 times per month), sometimes (1-2 times per semester), never. It is relevant to note that, according to the syllabus, Physics lessons for Grade 12 are delivered two days per week with a 90-minute class session on one day, and a 45 minutes session on the other day. For this particular question, it is important to note that, when it comes to establishing comparisons between the findings, one should bear in mind that the frequency level varies from one activity to another. For example, while in a real classroom situation the assessment by means of oral communication is one of the most frequently used activities, the assessment of students' notebooks cannot be done on, for instance, a daily basis. This means that the 5-point scale cannot be literally applied without taking this reality into consideration. Findings from this question, which were also collected through the questionnaire, are summarised by the Figure 5.2 presented below (n=12).

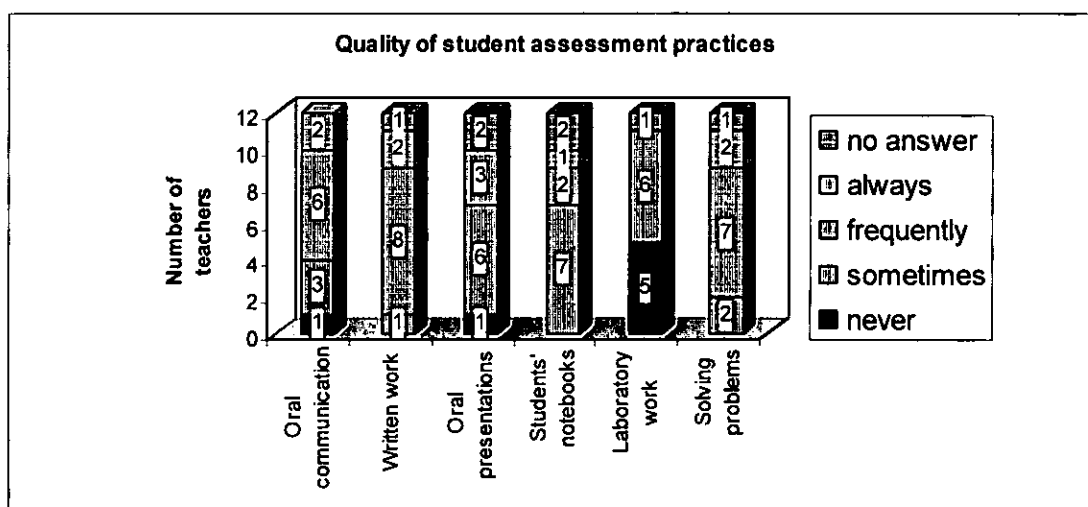


Figure 5.2: Teachers' responses on the assessment of students' activities (n=12)

The most assessed student activities are written work and providing feedback on student ability to solve problems (both used by eleven out of twelve teachers), where only one teacher gave no answer to each of them. Laboratory work is the activity that is never assessed by many of the teachers (five) or is assessed sometimes (six teachers). These findings were also triangulated using interviews. During the interviews, it became apparent that the teachers did not have the same understanding of the meaning of solving

problems. For some of them, solving problems only meant doing exercises in the classroom or at home using calculations. Perhaps a good understanding of the meaning of 'solving problems' could lead to different teachers' responses to this particular activity. In the context of the question asked by the study, 'solving problems' meant not only the student ability to devise solutions based on mathematical calculations, but, more importantly, referred to generating solutions of given real-world problems, for instance, by correctly setting up an electrical appliance.

Bearing in mind that the frequency alone is not enough to verify the quality of assessment practice there was a need to consider another way of investigating quality. Another way of investigating this was the analysis of some paper-and-pencil tests - one of the most frequently used assessment practices - given by teachers to students. A sample of 43 student tests, which had already been corrected by the teachers and handed out to students, was collected and analysed. The tests were randomly collected from six out of the twelve participating teachers – one teacher in each school. Five teachers brought seven tests each and the sixth teacher handed in eight tests. The reason for considering only one teacher in each school is that the tests collected were ACP (Activities of Partial Control given three times per semester) that are prepared by subject groups in schools. It was assumed that the characteristics of the other teachers and of the other teachers' tests would be the same or quite similar. The purpose of this analysis was not only to cross check the information given by the teachers during the questionnaire, but also to investigate the quality of feedback given to students by teachers when they assess the students' written work. Since the written work was identified as the most assessed type, an aspect of interest for the analysis was to verify how the teachers knew their students were making progress on this particular type of assessment practice. Specifically the analysis concentrated on the formative function of assessment through verifying how teachers were monitoring student learning and providing feedback.

Therefore, three aspects of formative assessment guided this analysis, namely the distinction between good and poor feedback (articulation), the time spent by teachers for

providing feedback (time), and the way the feedback was given (personality) (Race et al., 2005).

The following positive aspects of formative feedback were revealed during the analysis of the tests.

- *Articulation* - the feedback of the teachers include clear indication of the quality of work done by students, whether it was good (favourable mention) or poor (critical mention).
- *Timeliness* - most of the teachers gave their feedback timely while moving from one desk to another.
- *Personalisation* - the feedback was mostly given to students in an individualised form, which means that teachers made an effort to give feedback on each student's personal achievement.

One negative aspect in teachers' feedback to students that emerged from the analysis was related to the lack of input given to students in order to empower their learning. Although the teacher feedback was well articulated in terms of favourable or critical comment to the students' work, most times teachers did not provide directions needed to enhance the students' strong aspects (favourable) of learning or to improve their areas of weaknesses (critical). To worsen the situation, the language used by teachers included very strong words particularly for poor works (e.g., 'bad', 'very poor'). This type of feedback can dampen any enthusiasm for learning for most of the students (Race et al., 2005).

Given this situation, it can then be concluded that the assessment practices used by teachers in schools are still not of a good quality. This emerges, firstly, that even though several activities may be used to assess student learning of Physics, written work appears to be the one most used by teachers. Laboratory work, solving problems and oral presentations, deemed relevant for assessing student learning of Physics, are not included in assessment practices. Secondly, the way teachers monitor student learning and provide feedback is poor and consequently, requires substantial improvements.

5.2.3 Relevance of assessment practices

Two specific research questions were asked related to the relevance of assessment practices namely (i) *How do you engage students in the evaluation of their performance?* and (ii) *How often do you use the assessment results in class teaching and assessment?*

In relation to the first question, six alternative options were given to teachers (to select ‘all that applied’) and they are listed below.

- *I do not involve them at all* – students are not involved in the evaluation of their performance in any way.
- *By handing out the results* – students are given the results of their assignments without any relevant feedback.
- *By involving them in self-assessment* – students are given back their assignments and they are allowed to discuss amongst themselves about their own strong and weak points in the assignment.
- *By sharing with them the goals to be achieved* – students discuss and share with the teacher the assessment objectives to be achieved before they engage in the assessment task.
- *By explaining the implications to them* – students are given the opportunity to reflect on the consequences of achieving or not the expected learning outcomes, particularly for graduation purposes (classification of students and setting of standards).
- *By reflecting with them on the assessment data* – students reflect with the teacher about the assessment results (e.g., excellent and poor results) in order to learn from each other’s successes and failures.

The level of student involvement in the evaluation of their own work is seen to be an indicator of how well, or not, the students become aware of the importance of their performance, particularly when they peer assess (Race et al., 2005). Furthermore, Popham (2002) points out that by sharing with students the goals to be achieved and letting them

know what is expected of them helps, for instance, to achieve that goal. Below are the findings of this question (i), which were also mainly collected through the questionnaire.

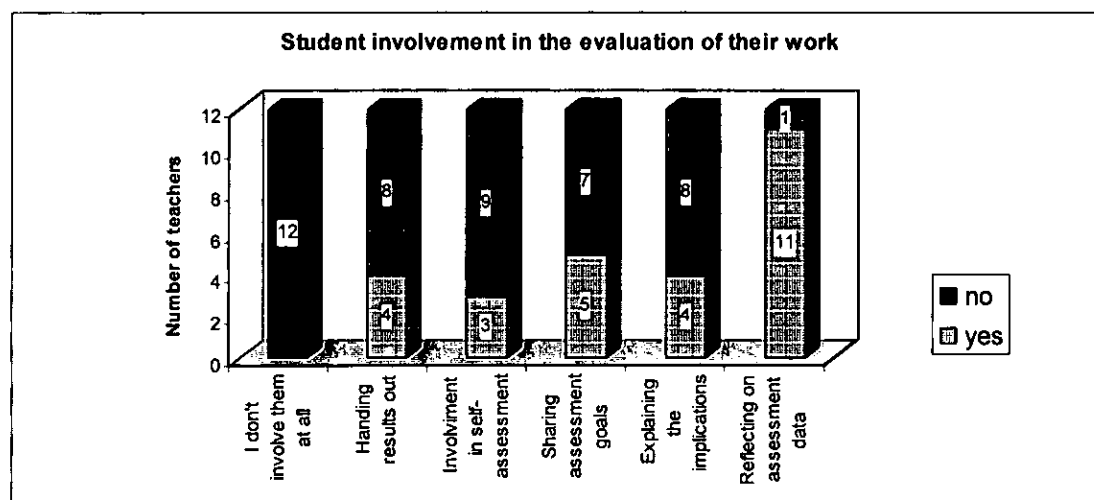


Figure 5.3: Teachers' responses on students' evaluation of their work (n=12)

The most common way used by teachers to get the students involved in the evaluation of their performance is reflections with them about their assessment results in order to motivate them to learn from their successes and failures. This is shown by 'reflecting on assessment data' type of teachers' responses given by eleven teachers (n=12). Self-assessment was less used by teachers to evaluate the student work (three out of twelve teachers).

During the interviews, however, it appeared that some of the teachers' responses to this question were contradictory. For instance, all twelve interviewed teachers referred to the fact that they never allowed students to assess their own work and only rarely allowed them to assess that of their fellow students. By "reflecting on assessment data" they meant summing up how many students got negative marks (0 to 9) and how many got positive marks (10 to 20). The formative element of assessment is not addressed. Interestingly, all teachers (n=12) said that they do not involve students in the evaluation of their work at all. This answer did not come as a surprise because, although the Mozambican curriculum for secondary education to be introduced soon advocates the constructivist approach of teaching and learning, teachers in schools are still employing the traditional teaching style characterised by a strong teacher-centred pedagogical approach. The analysis of the

findings also shows that some teachers' responses were not consistent. This is shown by the number of teachers who engage students by handing out the results and by reflecting with them on the assessment data. In fact, what was seen during classroom observations and also supported by some interviews (eight teachers), these two activities more often occurred simultaneously, i.e., the handing out of the results is always accompanied by the reflection with the students on the assessment data.

Finally, the question on relevance of assessment practices was investigated by asking teachers the question (ii) *How often you use the assessment results in class teaching and assessment?*

Four categories were identified to serve as objectives that the assessment results can be used for namely: (i) to assign grades; (ii) to identify students' strengths and weaknesses; (iii) to help students know and recognise the standards they are aiming for; and (iv) to encourage active involvement of students in their own learning. A 4-point scale was used to show the frequency of the use by teachers of each of these categories: no answer, sometimes (once per semester), frequently (2-3 times per semester), and always (more than 3 times per semester). Figure 5.4 shows the findings.

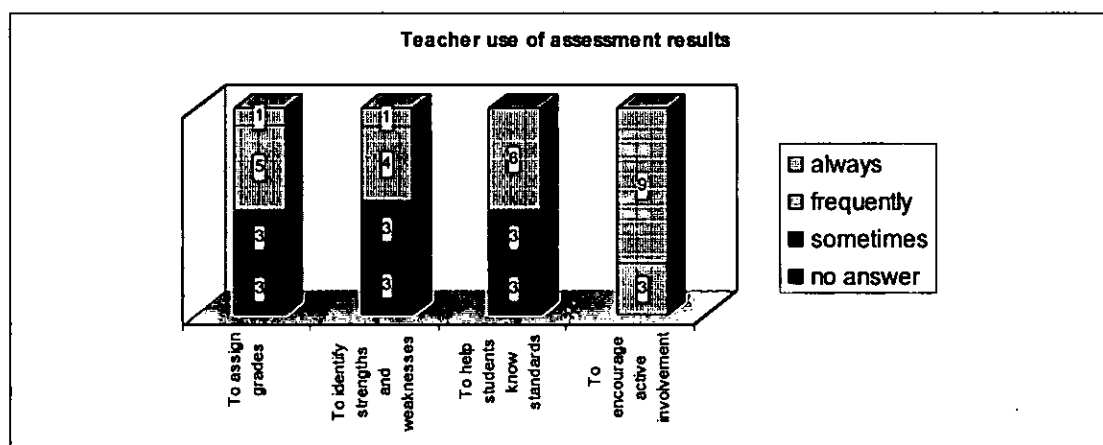


Figure 5.4: Teachers' responses on the use of assessment results (n=12)

Encouraging students to engage in active involvement in their learning is the assessment objective referred to most by the majority of teachers (nine out of twelve). However, an analysis of these findings shows that about one quarter of the teachers did not provide answers to each of the first three categories of assessment objectives and the same number

used the assessment results for these purposes only once per semester (sometimes). This means that very crucial aspects of classroom assessment like diagnosing student strengths and weaknesses, as well as sharing goals with the students have not been taken into consideration by a large proportion of the teachers. Alternatively, the number of teachers who said they *always* used the assessment results, for whatever category, is low: one for assigning grades, another one for identifying strengths and weaknesses, and none for helping students to know the standards. One relevant aspect to consider when analysing these findings is the relatively high number of teachers who said they used the assessment results *sometimes* and *frequently*. As seven teachers referred to it during interviews, these answers refer to the activities in which teachers of the same subject meet to analyse the student performance after an ACP (Activity of Partial Control) – given three times per semester – or when teachers of different subjects analyse the overall performance of the students at the end of the semester. According to the teachers, this analysis is merely statistical and it is aimed to get indications on how students are progressing towards the transition to the following grade level. Activities like item analysis or some other reflections with diagnostic or remedial purposes are rarely taken into consideration during these meetings.

So, in conclusion, if the level of student involvement in the evaluation of their performance is restricted to a mere *reflection on the assessment data* (mainly aimed at inferring pass and fail marks) without using the data for feeding the teaching and learning process, and if the number of teachers who *always* use the assessment results is very low, then it is fair to question the relevance of the teacher assessment practices for student learning of Physics. To worsen the situation, even those teachers who always use the assessment results, this usage is not directed toward the improvement of the teaching and learning of Physics in general. This suggests that, although in the early versions of the prototypes a priority should be given on helping teachers to design good learning materials – as referred to in Chapter 4, subsection 4.3.2.3 – it reinforces the idea that an added attention should be placed at the later versions on quality basic assessment practices and feedback.

5.2.4 Reflections from school directors, pedagogical officers and assessment specialists

The purpose of interviewing school directors and pedagogical officers and of consulting assessment specialists in the MEC was twofold. Firstly, there was a need to cross check the information provided by teachers in the questionnaires and interviews, and the one derived from researcher's classroom observations. Secondly, it was important to know to what extent the objectives of final assessments (examinations) carried out by the Ministry of Education and Culture to Grade 12 students are met as compared to the objectives of classroom assessment carried out by teachers in schools. Specifically, this second purpose is addressed in questions about the *level of teachers' preparedness to conduct formative assessments in schools, the relationship between the information provided by the teacher assessments in schools and that collected by the Ministry's final exam, and the evaluation of the student performance* at the end of the year.

Data from interviews with pedagogical officers indicate that the types of assessment practices carried out by teachers in schools address questions that basically ask students to mention or describe facts or phenomena, but they do not promote the establishment of relationships between these facts, or any comparison and explanations of causes and consequences of such phenomena. Still according to the referred sources, these characteristics of teacher assessments are determined by the teaching strategy, which focuses more on the memorisation of facts. Notes from assessment specialists reinforced this perception. They indicated that, as a result of this situation, the final exam from the Ministry of Education and Culture has also been negatively influenced, firstly, because it is prepared in consultation with teachers and secondly, because it would not be pedagogically desirable to introduce, via national exams, radical changes into the assessment system, which were not dealt with during the teaching and learning process. Both sources are of the opinion that, in comparing the results of the teacher assessments and of the examinations, those from the exams have been relatively low. The sources argue that this situation, however, should not be interpreted as resulting from low levels of achievement demanded by the teacher assessment as compared to the examinations.

Seemingly, the fact that, during examinations, students from different social and regional backgrounds, taught by different teachers, and studying in different learning environments and conditions write the same exam papers and are assessed at the same time, is leading to a counter-productive effect on the student performance. Without any supported evidence, however, one can only argue that there is a need for an in-depth reflection about the quality of both teacher assessments and examinations from the Ministry because several studies have reported teachers' poor skills in preparing acceptable test items, as well as the substandard quality of the exam papers. The Ministry of Education and Culture is conscious of this fact, which it addresses in the revised curriculum of secondary education, by putting more emphasis on supporting teachers in preparing and administering acceptable tests. In light of this complex situation, the interviewees, particularly the pedagogical officers, are of the opinion that one of the solutions to address the problem of quality formative assessment and of the comparability of the information of the two assessments, is to introduce what they call an 'exercise book' containing different types of test items addressing various student skills and competencies to help teachers prepare appropriate classroom assessments. These exercise books should be introduced in teacher training institutions to help the preparation of teachers for classroom assessment because according to the interviewees, the teacher trainers themselves appear to have difficulties in developing appropriate assessment tools.

On the question of the teacher preparedness to conduct formative assessment, both school directors and pedagogical officers pointed out the fact that, more often than not, teachers complain about the large class sizes. Teachers argue that there is no favourable environment to conduct formative assessment in such overcrowded classrooms. This is the reason why, according to the interviewees, the assessment has turned into a mere administrative process by which teachers provide mainly statistics about pass and fail marks and it does not reflect the real situation in terms of the extent to which the teaching and learning objectives have been met.

In relation to the evaluation of the student performance, the pedagogical officers refer that the Ministry's policy is that, when teachers submit the results of the final examinations,

they should include a report in which they indicate the main areas of students' difficulties, the most common errors the students have committed, some identified misconceptions or other alternative conceptions about facts and phenomena. They argue that the reality, however, is that these reports, if submitted, are of very low quality in terms of problem diagnosis. Teachers also complain again about the high number of students per class, which does not allow any in-depth analysis or diagnosis of the most problematic areas.

In summarising the information provided by these sources, it can be concluded that teachers are poorly prepared for conducting effective formative assessment of their students and they seem to face difficulties in preparing appropriate assessment tools. The high number of students per class worsens the situation. Interestingly, the national examinations undertaken by the Ministry are levelled from bottom, i.e., they depend on the quality of the classroom assessments that is reportedly low. The fact is not that the Ministry is in compliance with low levels of performance but, in practice, exam designers seem to be uncomfortable in instilling quality changes into the exams fearing that the standard levels will drop even more. As a result, the assessment process does not feed the teaching and learning process. Teachers are not providing valid reports about important areas of students' strengths and weaknesses. Given the fact that findings indicate that the results of the exams are lower than those of the teachers' assessments and that student difficulties are not diagnosed and monitored, it is fair to question the overall quality of the assessment system in general. The quality of feedback provided by teachers to students is fundamental in any procedure of classroom assessment and be central to the teaching and learning process as a whole. This feedback is effective only if it addresses the identified heterogeneous student needs (Black & William, 1998). But in the context of large classes, this is a challenge that should not be the responsibility of teachers alone but should be addressed in teacher training institutions in terms of how to adopt effective teaching and assessment strategies. The aspect of classroom management, which is a crucial issue in the context of large classes, is another challenge that needs particular attention from all decision makers and educational stakeholders. While effective classroom management and implementation of formative assessment is likely to be successful when the teacher has more teaching experience, resulting in solid pedagogical content knowledge, the most

experienced teachers are the ones who most frequently resist changes. There are many variables involved in collecting and analysing data to provide effective feedback to teaching and learning, which demands both technical and professional knowledge not always present in teachers with so many years of teaching experience. These teachers normally follow their teaching routines, which do not always accommodate suggested innovations. The suggested idea of introducing 'exercise books' seems to be one of the crucial measures for supporting teachers in conducting effective formative evaluation, particularly in the context of the current curriculum review.

5.3 Conclusions

At the outset of this chapter, it was stated that the objective of the Baseline Survey was to identify assessment practices used by Grade 12 teachers in schools as a way to inform the Intervention Study aimed at producing improvements on the teachers' practices. To do so, questionnaires, classroom observations and interviews were administered to a number of teachers, school directors, and pedagogical officers in six schools across the country and within the Ministry of Education and Culture. The overall research question addressed by the data collection instruments was: *What assessment practices do Grade 12 teachers in Physics in Mozambique apply and what is their quality?* Specifically, the instruments investigated aspects related to types of assessment practices used by teachers in schools, their quality, and their relevance for classroom practice. Teachers were asked about the types and quality of their assessments. School directors were queried about the purpose and relevance of assessments, as well as the evaluation of the school performance. Educational officers were interviewed to provide information on the relevance of assessments and the level of preparedness of teachers.

The findings indicated that the most frequently used assessment practices in schools are paper-and-pencil tests, verbal tests, and homework, while projects, portfolios, and peer-assessments are the less frequently used ones. These findings further suggest that some teachers have not used some assessment practices mainly because they were not familiar with them and they lacked preparedness for designing and administering them in the classroom. This raises the issue of the quality of assessment practices used by teachers,

which were then dealt with by investigating the frequency of using certain student activities and by analysing the quality of feedback given by teachers to students in paper-and-pencil tests. The quality issue was investigated in terms of the validity of the content of such activities, which included oral communication during lessons, written work, presentations, notebooks, laboratory work, and the student ability to solve problems. The most frequently assessed student activities were written work and the ability of students to solve problems, with laboratory work being the less frequently assessed. The findings indicate that the sharing of the assessment goals between teachers and students, the provision of feedback, and the facilitation of self- and peer-assessment among students are aspects which are hardly considered. The quality of assessment practices was also investigated through analysis of the teacher feedback to students during written tests. The analysis showed that the teachers not only are assessing limited range of student abilities but also are doing this with poor quality. In addition, feedback is not constructive enough to inform further learning.

One of the concerns about the quality of teacher assessment is that, although in the questionnaires a significant number of teachers said they used portfolios, homework, and assessed student ability to solve problems, during the interviews with school directors (who are also teachers) and with pedagogical officers, it became apparent that the teachers did not have either the same or a consistent understanding of these assessment practices. Another weak point of the teacher assessment quality - and thus of its relevance - is the evaluation of the student performance. Engaging students in a reflection about their assessment results meant, for most of the teachers, to sum up the number of students who got negative marks and of those with positive marks and did not involve any formative element of assessment. This suggests that teachers are in need of professional development training in the design, administration and evaluation of assessment practices.

One relevant conclusion of the Baseline Survey is that, although some literature has argued that there are some assessment practices that teachers use successfully (INDE, 2005; Lauchande, 2001), classroom assessment in most of Mozambican secondary schools is, in general, of poor quality. Teachers appear to experience difficulties in designing and

administering most of the researched types of assessments and in addition, the students are not assessed on their ability to perform real-world tasks. As referred to in Chapter 3 (Section 3.6), a performance type of assessment is seen to be one of the most adequate strategies to help students demonstrate their specific skills in applying the concepts and the knowledge they have acquired into real-world situations. According to the literature, an effective performance assessment is most likely to succeed when it is undertaken in a laboratory context where students can perform real demonstration experiments (Airasian, 2000; Dekkers, 1997; Gronlund, 1998; Tamir, 1991; van den Berg & Giddings, 1992). Thus, without neglecting other relevant assessment practices such as observation methods and oral questioning, a choice has been made for this study to improve teacher assessment practices through demonstration experiments where the students are allowed to demonstrate their ability to do real-world tasks while observing all the procedures involved.

Several aspects related to types, quality and relevance of assessment practices are derived from these findings as far as teachers and students are concerned. The ones listed below were of utmost importance and, therefore, were addressed by the Intervention Study.

1. Teachers and students need support in identifying and applying effective assessment approaches

Assessment of student learning of Physics can be a daunting exercise if effective assessment approach is not considered. Despite the importance of laboratory work for Physics learning, findings from the survey indicated that this activity was less assessed by many of the participating teachers. There were a number of reasons for this. Teachers were reported to find laboratory activities too difficult to assess due to the lack of equipment, because of their poor preparation in training institutions, or due to the time needed to conduct them successfully. The intervention phase of this study should suggest an assessment approach that is capable of achieving desirable assessment objectives in the context of the existing conditions.

2. *Physics teachers need support in designing and using appropriate and relevant assessment practices*

Of the seven investigated assessment practices, two of them (projects and peer-assessment) appeared to be used less by teachers. The reasons for this lack of usage were that either the teachers had no knowledge of that particular assessment strategy or they had a poor understanding about it. Interestingly these two assessments are, according to the literature, some of the most important assessment practices for science education. For instance, Brown et al. (1997) argued that projects have the potential of developing enquiry-based students' skills, which are important for Physics learning. As finished products resulting from a development process, they also bear the potential of developing student abilities for collecting, interpreting and reporting data. In relation to peer-assessment these authors claimed that its power lies on the element of mutuality where students can give and receive feedback from colleagues and teachers, which are important means for student learning. Therefore, the process of improving assessment practices proposed by the Intervention Study should investigate an assessment practice that is associated with these characteristics. This practice, as already argued in Chapter 3 (subsection 3.4.1) is performance assessment. By requiring students to perform real-world tasks, performance assessment, together with the Predict-Observe-Explain strategy of learning and assessment, involves all the elements of enquiry, feedback provision, and assessment of and from peers.

3. *Teachers need support in conducting effective formative assessments, especially in conditions of large class sizes*

Formative assessment is about the use of learning (or not learning) evidence to improve the teaching and learning. If there is no evidence that students are effectively learning, or if the evidence is not used to inform the next steps in the teaching and learning process, then the process cannot proceed. Findings from the Baseline Survey indicated that the teacher preparedness to conduct formative assessment was poor. Important areas of student strengths and weaknesses were not diagnosed and as a result, the assessment process did not inform the teaching and learning process. The intervention phase of this study should address this problem through producing and making available support

assessment materials containing aspects of formative assessment, particularly in the context of Mozambican education system characterized by traditional styles of teaching and overcrowded classrooms. Teacher training institutions are still dominated by a strong teacher-centred pedagogical approach. So, instead of blaming teachers for the lack of implementation of more learner-centred approaches, it is worth improving the education system as a whole through embedding changes in the training centres. The design of support assessment materials for teachers and students should not only go in this direction, but also toward the suggested idea of introducing ‘exercise books’ for supporting teachers in conducting effective formative evaluation, particularly in the current context of curriculum review. These materials should contain suggestions on how teachers can address the issue of conducting effective formative assessment in the context of large classes.

4. Teachers need support in assessing hands-on students’ activities

This study has chosen Physics as a focus subject. This is an experimental subject by nature and cannot be taught and assessed without practical or laboratory work. Findings from Baseline Survey indicated that paper-and-pencil tests, homework, and verbal tests were the most used assessment practices by teachers. These practices are important and useful for grasping basic Physics concepts but not effective for assessing daily tasks of real-world situations, which are essential for Physics learning. Thus, for an improvement of assessment practices in Physics classrooms, the intervention should support teachers in designing and using assessment strategies that teach and assess the student abilities to observe and interpret phenomena occurring to objects, carry out some hands-on activities, report findings, guided by teachers and teaching materials. So an emphasis on those activities requiring students to develop their manipulative skills, scientific investigation, communication and cooperative skills are important aspects to consider for the intervention to be successful.

5. Teachers and students need support in the evaluation of the student learning

One of the most powerful strategies used to motivate students to learn is to get them involved in the evaluation of their performance. This involvement ranges from sharing

with the students the assessment goals to be achieved, letting them assess themselves and their peers, handing the assessment results out, reflecting with them about the assessment data, as well as explaining the implications of their assessment. Data from the Baseline findings, however, showed that the most common way used by teachers to get the students involved in the evaluation of their performance was reflecting with them about their assessment results in the meaning of only summing up the results based on the pass or fail marks. This implies that the intervention should also support teachers on how can they ensure that students become aware of the importance of their performance by establishing a balance between all aspects of learning evaluation, and on how well the teachers can use the assessment results to inform the teaching and learning process.

In summary, the Intervention Study should address aspects of how teachers can identify, design and use effective assessments as well as how to engage students in some kind of performance tasks. Teachers must also be supported on how to evaluate and report the student work. As argued in Chapter 3, conducting performance assessment (e.g., laboratory demonstration experiments), designing and using exemplary materials (portfolios) and asking students to write laboratory reports, are all activities to be undertaken during the intervention and aimed at addressing the Baseline Survey findings.

CHAPTER 6

IMPROVING TEACHER ASSESSMENT PRACTICES IN PHYSICS IN MOZAMBIQUE

CHAPTER 6

IMPROVING TEACHER ASSESSMENT PRACTICES IN PHYSICS IN MOZAMBIQUE

This chapter reports on the design and formative evaluation of the prototypes of exemplary Physics assessment material (PAM) prototypes in Mozambican secondary schools. The materials are meant to assist teachers in providing formative feedback to Grade 12 students on conducting demonstration experiments and writing a report about these experiments. The chapter focuses on how the prototypes were developed and used by teachers and students in a classroom tryout. Guided by the educational design research approach, the prototypes were developed in an evolutionary and subsequent series of design and formative evaluation steps. The purpose of the tryout was to explore the validity, practicality and effectiveness of the materials. Section 6.1 introduces the chapter, with an overview of the design issues and the importance of the Baseline findings for the Intervention Study. Section 6.2 explains how the first prototype was designed and evaluated, while Section 6.3 presents the design, the classroom tryout and the evaluation of the second prototype. The participants, instruments, data collection procedures, and findings of the tryout are all presented and discussed in this section. The findings are discussed from the teachers' perspective, students' experiences, and researcher's point of view. The teachers expressed their views in relation to the quality of PAM materials in general and on the preparation and execution of the demonstration experiments. Students' experiences are illustrated by the aspects of the experiments they liked most, the aspects that they liked least, and how these demonstration experiments differed from the usual Physics experiments they conduct (if any) in the classroom. Aspects of preparation for classroom tryout as well as students and teachers' difficulties during the execution of the experiments are expressed by the researcher's views. The design and evaluation of the third prototype is presented in Section 6.4. Comments and suggestions from experts in relation to the third version resulting in the fourth version are discussed in Section 6.5 as a final expert appraisal of the PAM materials. Section 6.6 presents the conclusions and implications for further development.

6.1 Introduction

This study is aimed at investigating and improving the assessment practices used by Grade 12 Physics teachers in Mozambique. The literature review (Chapter 3) and the Baseline Survey (Chapter 5) provided background information out of which design guidelines and specifications for the Physics Assessment Materials were formulated. The literature review presented a clear picture of the existing need, both in the African context and elsewhere, to use collected evidence of assessment practices to make decisions about the next steps in learning. Lessons learnt from the review indicate that teachers need support in developing and using exemplary support materials on performance assessments and that these materials: (i) should be designed for an ordinary classroom environment to allow users to participate in the process while working in their normal routine; and (ii) should be formatively evaluated to allow that the learning evidence be used to feed the teaching and learning process. Apart from pointing out that teachers have mostly been using paper-and-pencil tests, the findings from the Baseline Survey revealed also that teachers lack preparation in terms of being capable of designing and administering appropriate assessment practices. More importantly, the survey highlighted the need to support teachers in conducting effective formative assessments, particularly for Science subjects like Physics and in conditions of large class sizes.

This chapter focuses on the design and formative evaluation of exemplary materials on performance assessment through demonstration experiments and a written report by students aimed at helping teachers to improve their assessment practices. The materials are designed and evaluated for the topics of *force* and *inertia*, which is part of Mozambican Syllabus for Grades 11 and 12 students.

Building from the previous chapters and recapitulating what has emerged from them, the following aspects appeared to be vital for the Intervention Study.

1. The literature reviewed in Chapter 3 has emphasised the importance of performance assessment as one of the most important assessment practices in Science education. There are two reasons for this. Firstly, because Physics is an experimental subject, which requires students (and certainly students in the final grade of secondary

education) to demonstrate meaningful application of essential skills and knowledge into practice. Secondly, because Grade 12 is the students' gateway to the employment sector or to higher education and thus they should be assessed on their ability to perform real-world tasks.

2. The intervention studies reviewed in Chapter 3 have highlighted the importance of developing exemplary support materials for teachers. These materials not only help teachers in several aspects of teaching and learning (subject knowledge, lesson preparation, teaching methodology, and assessment and feedback) but also allow students to construct their own knowledge.
3. Findings from Baseline Survey reported in Chapter 5, have shown that teachers were experiencing difficulties in designing and administering basic assessment practices. The findings have also confirmed that students were not assessed on their ability to perform tasks. The analysis of these findings indicated that the teachers not only were assessing limited range of student abilities but also that feedback was not constructive enough to inform learning.

In the Intervention Study being reported in this chapter the three aspects mentioned above are brought together. The improvement of assessment practices, being the purpose of the intervention, focuses on a performance type of assessment (through demonstration experiments) aimed at designing and developing a number of exemplary support materials (prototypes), which contain procedural guidelines for monitoring formative assessment and providing feedback to students. A decision had been made to put more emphasis - especially in the early versions of the exemplary materials - on the quality of the lesson materials for teachers and on the strategies for conducting demonstration experiments by students rather than on the assessment strategies and feedback provision (refer to Chapter 4, subsection 4.3.2.3 and Chapter 5, subsection 5.2.3). The reason for the decision is that for the teachers to be able to conduct effective assessment strategies they need support in preparing lesson materials of good quality. Any assessment strategy can only produce good learning results if this learning is taking place with quality learning materials. However, attention has also been paid to assessment strategies and feedback at the later stage of the intervention (e.g., on the final version of the PAM materials) where a number

of guidelines on how to monitor formative assessment and provide feedback is formulated.

6.2 Design and evaluation of the first prototype

As was referred to in Chapter 4 (subsection 4.4.3), the first PAM prototype contains materials on the topics of force and inertia. Although a correct teaching and learning of Physics content is important (concept acquisition skills), the PAM prototype emphasises the assessment component with focus on how to provide teacher' support in improving their formative feedback on student abilities to perform demonstration experiments.

6.2.1 Design of the first prototype

The first version of the prototype on PAM materials was designed by the researcher following design guidelines and specifications of exemplary assessment materials described in Chapter 4 (subsections 4.4.1 and 4.4.2), which were adapted from design specifications used by similar studies done by Mafumiko (2006) in Tanzania, Motswiri (2004) in Botswana, Ottevanger (2001) in Namibia, Nieveen (1997) in The Netherlands, and Teclé (2006) in Eritrea. As has been stated in Chapter 4 (subsection 4.4.1) these studies focused on supporting teachers by providing procedural specifications on lesson preparation, knowledge of subject matter, teaching methodology, and assessment of student learning. The general format of the PAM materials contained most of these features. A very distinctive element that made this study different from the above-mentioned studies is that the focus is on the design of exemplary assessment materials while the referred studies focused on exemplary curriculum materials. Besides, some additions were made which included a *teachers' guide* with description of the components of the assessment strategy and a *glossary of terms* explaining the key concepts used in the prototype. Applying these design guidelines and specifications, the first prototype of the PAM materials consisted of the sections described below.

- Part 1: *Introduction* - presented the place of the concepts of force and inertia in the Cycle 2 (Grades 11 and 12) Physics curriculum and the target student population.

For the teacher's consideration, an explanation about how to teach and assess the concepts using Prediction-Observation-Explanation (POE) strategy was provided in this part. Both teaching and assessment approaches followed the same strategy.

- Part 2: *Teacher's Guide* – provided an explanation of components and functions of assessment to help the teacher develop her/his own assessment strategies and a practice-oriented teacher's guide containing the sequence of content and lesson plan, some logistical aspects, and a plan of how to teach and assess following the POE strategy.
- Part 3: *Demonstration experiments* - started with presenting five demonstration experiments for students to carry out with a set of procedural specifications for teachers to guide the students when performing the demonstration experiments. This part also presented, as a sixth activity, the demonstration experiment report template for students as a summary of the five experiments, and ended with assessment rubrics to be used by the teachers in assessing student performance.
- Part 4: *Glossary of terms* – provided definitions and explanations of a number of terms or concepts used in the PAM prototype and a number of evaluation instruments (Appendices G, H, I, J, K) used by both teachers and students to evaluate the overall quality of the prototype particularly in terms of the practicality of the prototype.
- Part 5: *Worksheets* - presented the worksheets that the students used to carry out the demonstration experiments. The worksheets are composed of two sections. Section 1 (for individual work) corresponds to the first phase of the POE strategy (Prediction) that the students should do before carrying out the experiment, and Section 2 (for group work) is the second phase (Observation and Explanation), which is the actual demonstration experiment and the reconciliation of the data or outcomes.

Table 6.1 provides an overview of the design specifications applied to the development of the first prototype on PAM materials of this study.

Table 6.1: Design specifications of Physics assessment materials

| Area of support | Design specifications |
|-------------------------|--|
| Lesson preparation | <ul style="list-style-type: none"> • General overview of the lesson plan • Description of the intended learning outcomes • Suggestions on lesson plan and timing of the teaching and learning activities • Materials required for the demonstration experiments • Suggestions on how to deal with potential problems that may occur |
| Subject knowledge | <ul style="list-style-type: none"> • Suggestions for preliminary questions on students' prior knowledge to guide students about the experiments they are going to perform • Short explanations or revision of key Physics concepts • Examples of students' questions and answers |
| Teaching methodology | <ul style="list-style-type: none"> • Clear description of the roles of teacher and students • Suggestions on grouping of students during experiments • Suggestion on the need to try out the experiments beforehand • Suggestions on how to apply the POE strategy in its three phases |
| Assessment and feedback | <ul style="list-style-type: none"> • Suggestions on how to assess formatively the different student performance abilities during experiments through using rubrics • Suggestions on how to assess summatively the overall performance of the students using the Demonstration Experiment Report template. |

Based on these specifications, five demonstration experiments were designed for lessons of the first prototype. The sixth lesson was reserved for the introduction of the Demonstration Experiment Report template to students. The lessons are presented below.

- **Lesson 1 (45 min) - Introduction to force concept:** The POE strategy is used to examine student understanding of the concept of force. The experiment intends to introduce the concept of force by comparing some of the students' common-sense beliefs about it, which are seen to be incompatible with the scientific theory.
- **Lesson 2 (45 min) - Investigating the notion of force:** In this lesson, objects are set in motion and the POE strategy is used to explore the aim of the lesson, which consists of remedying the students' alternative conception of impetus.

- **Lesson 3 (45 min) - *Introduction to Newton's Second Law of motion:*** Use of POE strategy to probe students' understanding of the variation of the module of acceleration as a function of varying forces and explanation of the physical significance of Newton's Second Law of motion. The experiment begins with a brainstorm on students' understanding of basic concepts such as force, mass, speed, and acceleration.
- **Lesson 4 (45 min) - *Investigating inertia using a coin on the top of a can:*** The POE strategy is used to introduce the concept of inertia by analysing the behaviour of a coin put on a piece of a card, which is on a milk can.
- **Lesson 5 (45 min) - *Investigating inertia using a bottle put horizontally on a piece of paper:*** The POE strategy is used to demonstrate the effects of inertia. A bottle is put horizontally on a piece of paper, which is on the top of a table.
- **Lesson 6 (45 min) – *Demonstration Experiment Report:*** Explanation of the aim, procedures, methodology, and due date for preparing the Experiment report by students.

Having presented the design specifications, which guided the development of the first prototype, and the experiments to be carried out in the classroom, the following subsection discusses how this prototype was formatively evaluated before classroom tryout in order to improve its validity.

6.2.2 Formative evaluation of the first prototype

The formative evaluation of the first prototype of the PAM materials was undertaken mainly through appraisal by experts and to a lesser extent by users. The reason for this is that the focus of the appraisal was to further explore the validity of the material, i.e., the internal consistency between the prototype and the state-of-art-knowledge on Physics. Some aspects of practicality of the material were looked at by verifying, for instance, whether the usability of the prototype was compatible with the researcher's intention.

The appraisal was made through individual consultation with three groups of education professionals. Firstly, this was achieved through gathering opinions from three experts, namely a curriculum expert from the University of Twente, The Netherlands, a Science education expert from Eduardo Mondlane University, Mozambique, and an assessment expert from the University of Pretoria, South Africa. All experts have ample experience in Physics pedagogy, design and development of curriculum materials, and assessment. Secondly, the prototype was appraised by three postgraduate students from Eduardo Mondlane University who are also Physics teachers at high schools in Maputo. Two of them were involved in developing teacher support materials and in curriculum review for secondary education within the Ministry of Education and Culture. Thirdly, the appraisal was done by four Physics teachers who have more than ten years of teaching experience and who are teaching Physics in schools around Maputo City.

All experts were asked to comment on the overall content, the POE teaching and assessment strategy and on what they thought about the quality of the material in general. The curriculum and assessment specialists looked at a number of the curriculum components and functions referred to by van den Akker (1999) seen from the perspective of assessment. The science education expert was asked to review the quality of the prototype from the subject matter's point of view, i.e., the content accuracy of the two concepts under investigation. To guide their overall appraisal experts were requested to consider the following two questions (see also Appendix F):

a) Are the various items of the prototype accurate and specific enough to convey the intentions of the developer of establishing validity?

b) Is there consistency between the evaluation instruments (Appendices G to K) and the prototype?

Appendices G to K are the evaluation instruments used by university students, teachers and students to evaluate formatively the validity and practicality of the prototype before and after the classroom tryout. The university students and the Physics teachers were

asked to provide their views on the validity and usability of the materials as users. Specifically they were asked not only to provide feedback about the POE teaching and assessment strategy, but also on the teachers and students' potential problems during the activities, and other practical problems linked to carrying out the experiment. Appendices G and H are the evaluation instruments used by university students to guide their appraisal to the prototype.

In general, all appraisers (experts, teachers and university students) were positive about the idea of the PAM prototype. They commented that the content and the level of difficulty of the activities were suitable for the level of the students. However, they observed a number of aspects that needed attention in order to improve the quality of the prototype and made the following suggestions.

- In the wording of the demonstration experiments there is a mismatch between investigating physical concepts, and students' alternative conceptions. For instance, the experiment nr. 1 (*Introduction to force concept*) investigates a 'concept', while the experiment nr. 2 (*Investigating the notion of force*) is about students' (mis)conceptions. There is, therefore, a need to research specific knowledge and address the tasks in a coherent manner.
- An illustration of how the lessons will be conducted and experiments carried out needs to be provided. The teacher should have a practice-oriented lesson plan with all steps to be followed during each experiment.
- All demonstration experiments must be tried out in a real laboratory class environment to ensure practicality of the materials under the existing conditions.
- The demonstration experiments should be planned to fit within the time allocated for Physics lessons in the teachers' timetable (two double periods of 90 minutes, per week).
- With the PAM materials in general, a clear distinction and sound instructions should be given to distinguish between the teachers' guide, meant to help teachers develop their own assessments and the students' worksheet, for the demonstration experiments to be carried out by students.

- The teachers should be made aware that the POE strategy is not only an assessment strategy but is also a teaching strategy.
- The goals of the students' tasks must always be accompanied by corresponding assessment criteria.
- The prediction phase should be made individually to prevent a situation where relatively bright students take the lead in the discussions ahead of the low achievers.
- The assessment components and the glossary of terms sections should be made short enough to allow teachers to use them in a normal classroom environment.

6.2.3 Conclusions and recommendations for improvement

Overall, the expert and user appraisal was instrumental in improving the validity and some aspects of practicality of the Physics assessment materials by generating valuable suggestions for improvement. In general, all relevant suggestions were incorporated into the second prototype. Furthermore, in terms of the duration of lessons, the teachers have suggested that it would be more realistic and practical to have two time slots of 90 minutes each instead of the suggested three. This reduction in time did not raise any problems because, as a result of the suggestion given by the appraisers, the experiment nr. 2 (*Investigating the notion of force*) was removed and not included in the tryout. In the end, the number of experiments was reduced from five to four. Teachers also suggested that the introduction of the template for the experiment report could easily be given at the end of the last experiment and therefore does not need a separate time slot.

6.3 Design and evaluation of the second prototype

This section presents the design of the second prototype and the evaluation focus and questions, the selection of participating teachers and students, and the evaluation instruments for the classroom tryout. The design of the second prototype is presented in subsection 6.3.1. The description of the participants, the characteristics of the evaluation instruments and the procedures are all presented in subsection 6.3.3. Subsection 6.3.4

provides the findings of the classroom tryout, while the findings from the formative evaluation of the second prototype are provided in subsection 6.3.4.

6.3.1 Design of the second prototype

The review of the first prototype was the first cycle of formative evaluation and the comments and suggestions were taken into account in the design of the second version of the prototype, which was used in the classroom tryout. The second version of the prototype on PAM materials was also designed by the researcher following suggestions and comments from experts, teachers and university students. Adaptations and additions made to accommodate the suggestions from appraisers included those listed below.

Length of sections: The explanations of components and functions of assessment (Teacher's guide section), which initially included ten assessment components, and the definition of concepts or terms (Glossary of terms) were made short in order to be user-friendly for teachers. As a result, the teacher's guide in the second prototype had five assessment components (see Chapter 4, subsection 4.4.3). A practice-oriented lesson plan was included to help teachers guide the demonstration experiments.

Content: The wording of the experiments was improved to avoid confusion between investigating physical concepts and discussing students' alternative conceptions (see Table 6.2). As a result, experiments nr. 3, nr. 4, and nr. 5 were reformulated and the section about the demonstration experiment report was embedded in the second time slot of 90 minutes.

Execution of the demonstration experiments: As a result of the suggestions made particularly by teachers, handouts of the initial phase of the POE strategy (prediction) were prepared and distributed to students in separate sheets from the other two phases (observation and reconciliation) to avoid giving away the reasoning and the answers to the questions that students were supposed to provide after performing the demonstration experiments.

In summary, the section on demonstration experiments for the second prototype was composed of four experiments distributed into five lesson periods. Table 6.2 summarises the lessons, the content, and the periods of time for each lesson.

Table 6.2: Summary of lessons for the second prototype

| Lesson | Summary of the content | Time and periods |
|--------|--|------------------|
| 1 | <i>Introduction to the force concept:</i> The objective of this experiment is to introduce the concept of force. In order to help students understand this concept the POE strategy is suggested, which allows them to compare their common-sense beliefs with the experimental results (scientific theory). | 90 minutes |
| 2 | <i>Newton's Second Law:</i> The objective of this experiment is to understand the concept of force. The main part of the experiment includes the student understanding of variation of the module of acceleration as a function of force. The relevance of the experiment is that it creates a situation of a moving object with a constant speed. Students are asked to read the forces measured by spring balances and calculate the resultant forces. | |
| 3 | <i>Introduction of the concept of inertia 1 - A coin on top of a can:</i> The objective of this experiment is to understand how inertia can be realized through analysing the behavior of a coin put on a piece of a card, which is on a can. By using the POE strategy, you are firstly required to predict what will happen to the coin if the card is flicked quickly, then to perform the experiment yourselves (in groups) and finally to draw a reconciliation between the prediction and observation. | 90 minutes |
| 4 | <i>Introduction of the concept of inertia 2 - A bottle on a paper:</i> This experiment is about demonstrating the effect of inertia. The bottle is put horizontally at rest on a piece of paper, which is on the top of a table. You are required to realise that after flicking quickly the paper, the bottle tends to stay at rest. Again, the POE strategy is used to assess your understanding of inertia. | |
| 5 | <i>Demonstration Experiment Report:</i> Explanation of the aim, procedures, methodology, and due date for preparing the lab report by students. | 45 minutes |

Lesson three of this prototype is given here as an example of a plan for demonstration experiments following the POE teaching and assessment strategy. With the exception of the lesson on Demonstration Experiment Report, all other lessons of this and other subsequent prototypes followed similar planning structure. In order to illustrate the

characteristics of the PAM materials, the execution of this lesson by the teacher and the corresponding student worksheet are presented in Figures 6.1 and 6.2, respectively.

Figure 6.1: Execution of the lesson: a practice-oriented lesson plan for the teacher

| Lesson 3: Introduction of the concept of inertia 1 - A coin on top of a can | |
|--|--------------------------------------|
| Objective of the lesson | |
| The objective of the lesson is to introduce the concept of inertia using the experiment of a coin put on a piece of a card. The lesson intends to illustrate that a coin put on a piece of a card, which is on the top of a can, remains on the top of the can even after the card is flicked away. | |
| Intended learning outcomes | |
| During the experiment students should be able to understand (by observing and explaining) that, because of inertia, the coin remains on the top of the can, when the card is flicked quickly. Students should be assessed towards: | |
| -realising the difference between flicking the card quickly and flicking it slowly as this is important for the explanation of the concept of inertia. | |
| -explaining the reasons behind the behaviour of the coin. | |
| The teaching and assessment strategy to achieve this is characterised by (i) allowing students to predict what will happen to the coin put on a piece of a card if the card is flicked quickly (individually); (ii) letting them perform the experiment themselves and observe the behaviour of the coin (in groups); and (iii) asking them to draw the reconciliation between their predictions and what they actually observed during the experiment (individually). | |
| Lesson plan and timing | |
| Activity | Approximate time (in minutes) |
| Start of the lesson | 5 |
| Activity: Experiment | |
| -Prediction | 10 |
| -Experiment | 15 |
| -Reconciliation | 10 |
| Assessment and feedback | - |
| -Monitoring | |
| -Providing feedback | |

| | |
|------------------------------|-----------|
| Conclusion and end of lesson | 5 |
| Total | 45 |

1. Start of the lesson (maximum 5 min)

You may start the lesson by asking students brief questions on what they already know about concepts like force and mass. A small discussion about these concepts may help you to understand and evaluate student predictions and/or their responses during experiments. Examples of the questions to get the discussion started are:

- What do you understand about force?
- What do you understand about mass?
- (Other questions...)

Having had a number of responses to these and other questions and having facilitated the discussion:

- State the objectives of the lesson (emphasising the POE strategy) and clarify what is intended to be achieved at the end of the experiments.
- Explain the working method: everyone must be in the classroom and they must be organized into groups of a maximum number of four students each.

2. Activity: demonstration experiment

(i) Prediction (maximum 10 min)

Start the experiments by forming the groups and its distribution in their places. First distribute the first part of page 1 of the students' worksheet (about prediction) to each student. Guide the students in doing part 1 of the POE strategy (prediction) individually.

- While the students are looking for answers to the questions posed on the prediction section, help them to do the reasoning, but only as a moderator!

(ii) Experiment (maximum 15 min)

- Distribute the necessary material for the demonstration experiment: A coin, a piece of a card, and a can. Assign students in groups of two or three and give each group the remaining part of the students' worksheet. Ask students to have pencils and sheets of paper for calculations.
- Ask students to perform the experiment in groups following the steps indicated on the Student Worksheet.
- While they are doing the experiment and answering the questions posed in the observation section, keep helping them doing the reasoning, but only as a moderator!

(iii) Reconciliation (maximum 10 min)

Guide students on how to compare and explain consistencies or lack of them between results from the prediction and from the observation, but only as a moderator!

3. Assessment and feedback (to be considered throughout the lesson)

You must undertake a formative evaluation of the students' work through:

- starting the lesson by asking brief questions to students on what they already know about force, speed, mass, or inertia;
- verifying whether the students understood the outcome to be achieved at the end of the experiment, and procedure (teaching and assessment methodology) to be followed to achieve this;
- observing what students do (individually and in groups). During the experiment you should ensure that the coin is put on the top of the can adequately and that the card is flicked away quickly. Whenever possible, you must also ask probing questions (e.g., why should the card be flicked quickly?);
- encouraging students to discuss amongst themselves (during the experiment) several aspects of the experiment. For instance, ask them to explain why the coin does not remain on the top of the can when the card is not flicked quickly;
- allowing students (during the reconciliation phase) to reflect on differences or similarities of their predictions and on those observed during the experiments and allow comparisons between their ideas with those of their colleagues.

Remember, your role is to facilitate the students' work, and should act only as a moderator.

4. Conclusion and end of the lesson (maximum 5 min)

You may round off the lesson by recapitulating its objective, that is, to understand why the coin put on a piece of a card, which is on the top of a can, remains on the top of the can when the card is flicked away quickly. Whenever time allows, ask some individual revision questions to verify the extent to which the intended learning outcomes have been achieved. Give students immediate congratulatory or critical feedback (preferably individual). Explain also the students that they will have to evaluate and summarise their answers and observations about the laboratory experiment. Tell them that this will be done by the means listed below.

- Writing Demonstration Experiment Report - distribute the Report template and urge them to meet the deadline for submission.
- Filling in an Evaluation Experiment Questionnaire - this is to be given immediately after performing the last experiment.

The type of information to be asked in the Demonstration Experiment Report and in the evaluation questionnaire and how to assess it is provided in the PAM assessment materials (Appendix P).

Round off the lesson by asking students to clean up and return the materials used in the experiment.

Figure 6.2: Worksheet for students

Experiment nr. 3: Introduction of the concept of inertia 1 - A coin on top of a can

The objective of this experiment is to teach inertia through analysing the behavior of a coin put on a piece of a card, which is on a can. By using the POE strategy, you are firstly required to predict what will happen to the coin if the card is flicked quickly, then to perform the experiment yourselves (in groups) and finally to draw a reconciliation between the prediction and observation.

3.1 Equipment required:

- (a) A coin
- (b) A piece of a card
- (c) A can

3.2 Prediction:

A coin is placed on a piece of a card, which is on a can. If the card is flicked quickly the coin:

- (a) will be dragged off with the card
- (b) will stay where it was (on the top of the can)
- (c) other: _____

Give reasons for your prediction:

3.3 Observation (experiment):

In groups of three students each, perform the following experiment:

Place a coin on a piece of a card on the top of a can, as shown in Figure 2. Flick the card quickly.

Describe your observation.

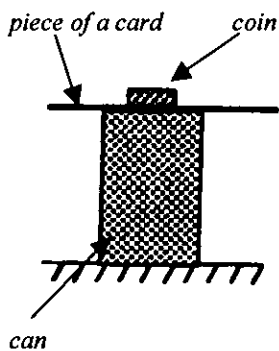


Figure 2: A coin on top of a can

Repeat the experiment twice. Describe what you observe regarding what is happening with the coin.

Explain why _____

3.4 Reconciliation between prediction and observation:

(a) Compare the results of the prediction and those of the experiment.

(b) Are the results of the experiment the same as those of your prediction? (Yes)____/(No)_____.

(c) Justify your answer _____

Subsection 6.3.2 discusses how the second prototype of the PAM materials was evaluated before the classroom tryout.

6.3.2 Evaluation design of the second prototype

The main objective of the tryout of the second prototype was to find out whether the PAM materials developed, appraised, and tried out with and by teachers in the Mozambican context were already practical and effective for the Grade 12 level Physics classes. Based on the findings from the Baseline Survey and with this objective in mind, the evaluation of the second prototype was designed. The initial step consisted of the identification of a number of topics which was used to guide the formulation of the main evaluation questions. This was followed by the formulation of the main evaluation questions which were the basis for developing the evaluation instruments for both teachers and students. The main evaluation question for the teachers was formulated as follows:

- *What are the teachers' opinions about the use of PAM materials and their experiences with student performance assessment in the demonstration experiments?*

For the students the evaluation question was formulated as follows:

- *What did the students like and dislike about the demonstration experiments?*

These questions were then used for elaborating the sub-questions in the data collection instruments for teachers and students. Some examples of the aspects dealt with in the

evaluation sub-questions included the clarity of language, the description of the experiments, the practicality of the POE strategy, the role of teachers and students during formative assessment, and the time required to carry out the experiments following the POE strategy. Table 6.3 provides an overview of the sub-questions posed to each of the users.

Table 6.3: Overview of the sub-questions for the evaluation design

| | Elements of typology | Sub-questions |
|----------|--|--|
| Teachers | General impressions about the material | <ul style="list-style-type: none"> • Is the language clear and understandable for students? If not, explain what the problems were. • Was the description of the demonstration experiments and pictures clear for the students or did they have many questions? If not, what are the needed improvements? • Do you feel the prototype as a whole needs any changes or additions? If yes, what changes? |
| | Opinions about the demonstration experiments | <ul style="list-style-type: none"> • Was the POE strategy of teaching and assessment practical for students' reasoning? <ul style="list-style-type: none"> ○ Which parts were useful and why? ○ Which parts need improvements and why? • What was your role as a teacher during the experiment? • Do you feel that the main objective of these demonstration experiments will be achieved? If not, which particular aspects will not be met and why? |
| Students | Opinions about the demonstration experiments | <ul style="list-style-type: none"> • What is your impression about the structure of the material? • What did you like and dislike about the demonstration experiments? |
| | Experiences concerning differences between demonstration experiments and the usual Physics experiments | <ul style="list-style-type: none"> • Were the demonstration experiments different from the usual Physics experiments you are used to in your class? • What was your role as a student during the experiment? • Did you face any problems during the execution of the experiments? If yes, what were the problems? |

Apart from any practical problems that teachers and students could identify during the classroom tryout, the evaluation also addressed two aspects from the researcher's perspective. Firstly, the researcher's focus was on those aspects where teachers seemed to experience difficulties during preparation of the demonstration experiments. Secondly, the emphasis was on those aspects that could be regarded as problematic for both teachers and students during the execution of the experiments. This information was collected by the researcher using a logbook.

In summary the sub-questions shown in Table 6.3 (first column from right) were used as questions to guide the designing of evaluation instruments. The sub-questions addressed the two main evaluation questions from four perspectives, namely: (i) teachers' general impressions about the Physics assessment materials; (ii) teachers' and students' opinions about demonstration experiments following the POE teaching and assessment strategy; (iii) students' experiences concerning differences between demonstration experiments and the usual Physics experiments; and (iv) researcher's general impressions, opinions and experiences about the course of the demonstration experiments.

Subsection 6.3.3 discusses how the second prototype of the PAM materials was tried out in the classroom.

6.3.3 Classroom tryout

The second version of the prototype was tried out in the classroom by two teachers and their students (n=62) from Joaquim Chissano Secondary School in Gaza Province and from Matola Secondary School in Maputo Province. Four demonstration experiments were tried out in the classroom. This version had taken into account comments and suggestions from the appraisal of the first version done by experts and the postgraduate students. The main objective of the tryout was to identify initial problems with the *practicality* of the PAM materials in the context of Mozambican Grade 12 level Physics classes. More specifically, the tryout was intended to find out whether the materials have

the potential of being successfully used by teachers and students following the POE teaching strategy.

The following subsections report on the characteristics of participants, the type and characteristics of the data collection instruments, and the procedures followed during the classroom tryout.

Participants

The tryout was undertaken in two public and urban secondary schools from Gaza and Maputo Provinces. Participants were two Physics teachers and 62 Grade 12 students from the science stream. The schools are here referred to as school 1 and school 2, while teachers are indicated as T1 and T2. Table 6.4 provides a summary of the characteristics of participants in the classroom tryout of the second prototype.

Table 6.4: Characteristics of participants of the classroom tryout

| Participants | School 1 | School 2 |
|------------------------------|-----------------|---------------|
| <i>Teachers</i> | T1 | T2 |
| -gender | -Male | -Female |
| -professional qualification | -Licenciatura | -Licenciatura |
| -teaching experience (years) | -10 | -12 |
| -other responsibilities | -Subject leader | -None |
| <i>Students</i> | Class A | Class B |
| -number | 21 | 41 |
| -age-range | 17-23 | 16-21 |

In total, about 67 participants tried out the materials. Sixty-two (13 girls and 49 boys) Grade 12 Physics students from the two schools, their two teachers (one from each school), and three other teachers from the Francisco Manyanga Secondary School who were invited to participate in a one-day evaluation workshop together with the teachers from school 1 and school 2. The workshop was meant to train the participating teachers in using the assessment materials, with an emphasis on their application and utilisation (functions of assessment relevant for Physics, sequence and content of lesson periods,

preparation and execution of lessons, etc.). The two teachers who took part in the tryout had *licenciatura* degrees, a five-year teaching course in Physics and mathematics and each had more than 10 years of Physics teaching experience. Teacher T1 was also responsible for coordinating the subject for the grade in his school.

Instruments

The instruments used for the tryout were designed on the basis of the analysis of the evaluation design (refer to Table 6.3) and thereafter adapted from the instruments used in similar studies by Mafumiko (2006) and Tecle (2006). They included an evaluation questionnaire and interview schedule for teachers (Appendices I and J), a student questionnaire (Appendix K) and a researcher's logbook. Characteristics of these instruments are presented below.

Teacher questionnaire and interview schedule

Both teachers individually completed a questionnaire at the end of all sessions with Physics demonstration experiments. The questions in the questionnaire focused on the teachers' general opinions about the experiments in terms of relevance of the topic, content, structure of the teaching strategy, and presentation of the experiments. The questions also included aspects related to potential problems with the availability of time to carry out the experiments, other practical problems linked to experiment execution, and teachers' suggestions on how to improve and adapt the teaching strategy in their classes. Following on from the questionnaire and based on their results, a follow-up open-ended interview with the two teachers was conducted to elicit information and provide clarification on some aspects from the questionnaire. The teachers' questionnaire and interview schedule can be found in Appendices I and J.

Student questionnaire

All 62 students from the two schools completed a semi-structured questionnaire at the end of all demonstration experiments. The questionnaire was meant to discover the opinions and experiences of students of the demonstration experiments they conducted. The focus was on what the students liked or disliked most, the specific learning activities, and how

the lessons compared to their regular Physics demonstration classes. For more information about this questionnaire see Appendix K.

Researcher's logbook

As referred to at the beginning of this subsection, the researcher used a logbook to keep a record of activities and observation notes linked to the use, by both teachers and students, of demonstration experiments and Physics assessment materials in general. The focus was on those areas where teachers seemed to experience difficulties during the preparation of the experiments. For the execution of the experiments, the researcher kept running notes on how teachers introduced the POE strategy, and how they monitored the student activities and the assessment process. Notes were also kept on the observed responses of students to performing Physics experiments and any difficulties they experienced.

Procedures

As a preparation for the classroom tryout, the researcher distributed copies of the PAM materials (the second version of the prototype) to the teachers a week before the tryout (after the evaluation workshop) for preliminary reading. Four demonstration experiments were conducted in the tryouts in each school. The tryout process consisted of an interactive preparation session (a half-day workshop) with teachers and the researcher at each school to introduce the materials, the purpose of the study, and clarification of potential difficulties in understanding. During this session the demonstration experiments were also tried out in advance with the teachers and copies of the student worksheets were given to the teachers for preliminary reading. The following subsection describes how the demonstration experiments were undertaken in each school.

School 1

Three days in a row were spent in this school during the tryout. The first day was devoted to the interactive and preparation session for introducing the materials to the teacher. The second day was for the tryouts, i.e., for conducting the lessons with the students. The four demonstration experiments were conducted during one block of 90 minutes (consisting of two lessons of 45 minutes) by the teacher, with the researcher as participant observer. Two

experiments were conducted in each lesson of 45 minutes. The teacher conducted the four experiments in one of his classes. The third day was initially dedicated to the completion of the evaluation questionnaire by the teacher and by students. Then a general discussion with students about the experimental results and the difficulties they encountered in carrying out experiments using the POE strategy, was held. Thirdly, and following the discussion, the students were given the laboratory experiment template to take home and to use for writing the report. They were urged to submit the report within one week. Finally and by the end of the day, the teacher was interviewed to find out his views about the experiments in general and of the practicality of the materials overall. In particular, the teacher was asked to express his views on students' formative assessment, the impact of the laboratory report on students' assessment, and the practicality of the POE strategy.

School 2

At this school the same number of days as in school 1 was spent following the same procedure. The preparation session took place during the first day between the teacher and the researcher. Teacher T2 conducted the four experiments during two lessons. The experiments were all co-supervised by the researcher as a participant observer. The last activity involved administering the evaluation questionnaires to students and to the teacher, as well as interviewing her with the aim of gathering her experiences and opinions about the materials.

As referred to earlier on, the purpose of the second tryout was to increase the practicality of the PAM materials from users' feedback in the context of classroom environment. Teachers and students were asked to conduct the experiments and to use the PAM materials as prototypes, i.e., not as final products, but as products which might require improvements. Therefore, flexibility in adapting the materials and encouragement to provide alternative ideas whenever needed were allowed, always paying attention to the POE strategy, the performance assessment task, and practicality of the materials.

Subsection 6.3.4 reports findings of the classroom tryout.

6.3.4 Findings from the classroom tryout

The classroom tryout focused on the *practicality* of the PAM materials, particularly the use of POE strategy. The emphasis was on: (i) how the demonstration experiments were conducted in class; (ii) how teachers felt about using the materials, following the strategy and monitoring the experiments; (iii) what the students experienced about the experiments following the POE strategy; (iv) what teachers and students' conceptual difficulties were with the materials and experiments as seen by the researcher; and (v) what lessons could be learned from demonstration experiment reports written by students and evaluated by the teachers. The following subsections present and discuss results of the classroom tryout.

Teachers' impression about the PAM materials and their opinions on the experiments

Teachers' opinions about the PAM materials and demonstration experiments are discussed from two perspectives. Firstly, the focus is on how practical and effective the materials can be as teaching and assessment tools. Secondly, the emphasis is on the contribution of these materials – particularly the POE strategy – to the students' understanding of Physics concepts.

The general impression of the two participating teachers about the practicality and effectiveness of the PAM materials was positive. They indicated that materials of this nature would be very useful in helping them, firstly, to enhance their professional experience and secondly, to prepare their own assessments for students. For instance, they liked the presentation and structure of the materials following the POE strategy, and felt that this helped students to construct their own understanding of the concepts as advocated by the constructivist approach and recommended by the Grades 11 and 12 Syllabus. Teacher T1 for instance, referred to the fact that, in his role as subject leader, he had been facing demands from colleagues to provide support in lesson preparations particularly with guidelines on how to design and facilitate appropriate demonstration experiments. This teacher commented that, although his school had some laboratory equipment, most of

his colleagues seemed to be ill-prepared to carry out laboratory lessons apparently due to their poor preparation. This situation “seems to be more acute for those teachers who were trained many years ago and who, despite their long teaching experience, appear to be in more need of support than their younger and newly trained colleagues”, said this teacher. Besides this positive impression about the materials, the teachers also raised some concerns. For instance, the fact that the POE strategy is principally more suitable for the introduction of new concepts and motivational phases of teaching rather than for lessons aimed at consolidating already acquired concepts or knowledge, was referred to by teachers as one aspect that could be taken into account in future. Table 6.5 summarises teachers’ general impression of, and comments about, the improvement on the practicality and effectiveness of PAM materials.

Table 6.5: Teachers' general impression about the second prototype of the PAM materials

| Aspects of the material | Teacher 1 | Teacher 2 |
|--|---|---|
| Language | Understandable for the level of the students. | Very clear, provided that some related concepts be introduced to students well in advance. |
| Description of the demonstration experiments | Very well structured, fall under prescribed G12-level, clear instructions and easy to follow. Materials available locally. But in experiment nr. 2 better talk about <i>average speed</i> . | Content very well presented, experiments well explained but some related concepts need to be explained in advance (e.g., measurement error). |
| Pictures | Clear and neat, except the picture of experiment nr. 2 | Picture of experiment nr. 2 need to be improved |
| Teachers' guide | Useful but too long to follow. It is difficult to get time to read the entire guide prior to designing of each task. | Very relevant for Physics teacher, well structured and understandable provided the glossary of terms. |
| Time management | Difficult with so many students in the class. Ninety minutes would be enough for only four experiments. | Time could be a serious problem given our overcrowded classrooms. Maybe it would be desirable to conduct experiments during extra time periods. |
| Relevance and usability | Very useful and relevant for G12 students. The experiments provide students with opportunity to think, observe, perform tasks, and discuss the results. | Very helpful, especially for introducing new concepts. |
| Practicality of the strategy | Very practical: Materials available or obtainable at a reasonable price. Only time would be problematic. | Good strategy, easy to follow but time consuming. It requires students to think and reason about the events. |
| Role of the teacher | The strategy is new for the students and my role was mainly to explain students how to follow the steps of the strategy. | My role was to try out the strategy in advance and to guide students during the experiments. |
| Students' involvement | Students were very fascinated with the experiments and this was because I think they thought it was a preparation for the upcoming final exams. | The students seemed to like the demonstration experiments especially because we rarely do experiments like these in our school |
| Achievement of the main objective | The objective was achieved. Students seem to understand these concepts very well now. | The main objective was achieved provided that the suggested improvements will be taken into account, especially with the experiment nr. 2. |
| Suggested improvements | The time issue must be taken care of. The teachers' guide also seems to be too long. | Improvements have to be made on experiment nr. 2 (pictures and wording), and the number of experiments should be reduced to fit the timetable. |
| POE strategy | | |

The table indicates that both teachers considered the PAM materials very useful for Grade 12 students. The materials were found to be helpful for students because they provided them with the opportunity to merge what they have acquired theoretically with experimental practice. Teachers stated that the POE strategy provided a solid foundation for how to design and conduct demonstration experiments. Predicting the results of the events, carrying out the experiments, and observing and explaining the outcomes of the experiment would help students to identify and minimise their own misconceptions about force and inertia. On the practicality of the strategy, both teachers expressed satisfaction with the availability of materials required for the experiments as illustrated by the following response given by teacher 1:

“When I received the copy [of the PAM materials] I did not expect good results because students are not used to conducting demonstration experiments... I was anticipating problems especially in reconciliation of theory and practice... However, this POE strategy is so simple, clear, and well structured that it was easy to follow. But, as I have said before, time can be a problem.”

From this statement it seems that teachers regard the POE strategy as practical for Physics teaching and learning in the classroom and that it can lead to good student results, although it is not explicitly indicated how. The fact that this teacher did not mention the problems that he had anticipated does not allow one to draw relevant conclusions about the potential of the strategy seen from this perspective. Apparently, the problems are linked to the unfamiliarity of students with demonstration experiments or to the students' inability to provide sound arguments about similarities or differences between their predictions and their observations.

As for their roles as teachers, the findings indicate that they spent most of the time trying to help students in following the steps of the strategy because this was also new for them.

When teachers were asked how they perceive the contribution of the materials to the improvement of student learning, they referred to the pedagogical element of the POE strategy. According to them, the strategy helps to sharpen the students' levels of comprehension of the concepts of force and inertia. They mentioned the example of the experiment nr. 1 (*Introduction to the force concept*) as most illustrative. While

some of the students predicted that three forces (force of gravity, force of the 'kick' and force of air resistance) were acting on the ball during its flight, in the reconciliation phase, where they were involved in discussions amongst colleagues, they were able to amend their thoughts. During the sharing of ideas, it emerged (the students remembered each other) that the experiment was undertaken in an idealised system where the force of air resistance is neglected. It was also during the reconciliation phase where the students realised that during its flight, the ball is no longer under influence of the force of the 'kick'. In fact, and according to teachers, in this specific demonstration experiment the observation phase was not as effective because students were unable to observe the kind of forces acting on the ball directly. This phase had a rather more instructional effect in the other experiments where events were more directly observable. So, the potential of the POE strategy lies on the possibility given to students to contribute their ideas and enhance the building of knowledge. If students do not engage in such discussions, they do not get the possibility to develop their critical thinking through learning by their own mistakes and those of their colleagues.

Responding to the question of whether there were students who were not active during the experiments, the two teachers indicated that all students were very active. However, the teachers gave different reasons for the active students' involvement. While for teacher 1 the students were very active because the experiments took place closer to the examination period (and therefore they thought would give them hints for the examinations), teacher 2 felt that this could be because they were experimenting with something new in their school life. Demonstration experiments had not been carried out in school 2 for the whole year.

The two teachers also indicated some aspects that they did not like about the POE strategy and the experiments. The most worrisome of them all is the time spent in performing the experiments. They suggested reducing the number of experiments per lesson or extending the time allocated for the lesson to fit the time needed for each experiment. To solve the problem of the time, the teachers were of the opinion that students should be given the worksheets well in advance to read before they actually arrived in the class. Still according to the teachers, this problem of shortage of time did not allow them to realise at what extent the POE strategy would have facilitated the

students' understanding of force and inertia concepts because the students lacked sufficient time to discuss their views amongst each other and to build informed consensus. Teacher 2 indicated that she did not like experiment nr. 2 because the pictures were misleading and the wording used to explain the several steps was confusing.

Teachers' experiences on conducting demonstration experiments

The teachers' experiences, related to how demonstration experiments were conducted, are expressed by findings obtained through teacher evaluation questionnaire (Appendix I), follow-up interviews (Appendix J), and the researcher's logbook. These data collection instruments focused on two distinctive characteristics, namely the teachers' general impression concerning several aspects of the materials and their opinions about the POE strategy.

Before elaborating on teachers' experiences related to the experiments, it is important to refer that, as was discussed earlier, as a preparation for the classroom tryouts a half-day workshop took place with each teacher as a preparation session prior to the experiments. During this workshop each teacher had an opportunity to clarify potential difficulties with the PAM materials in general (refer to Appendix P), with emphasis on the components and functions of assessment considered relevant when assessing Physics concepts, the sequence and content of lesson periods, and the preparation and execution of the lesson (Part 2, subsections 1 to 3). Teachers' impressions about the quality of the PAM materials in general were presented in the previous subsection. The present subsection discusses the teachers' experiences related to the preparation for and execution of the demonstration experiments.

Teacher 1 felt that the preparation of these experiments took a large amount of time, which is not always available in the normal time schedule. The teacher is of the opinion that teachers' commitment is crucial, if good results are to be achieved from the experiments. In fact, during the preliminary tryout, this teacher spent about 35 minutes in lesson preparation: ten minutes looking for the equipment, fifteen minutes trying out each experiment, five minutes grouping the students, and another five minutes introducing the task to the students. This amount of time is considered to be

large if taking into account that the lessons are placed one after another on the teaching timetable.

Teacher 2, although she also raised concerns about the time needed for preparation, said that this problem could still be overcome with more practice. She argued that preliminary preparation and practising by the teacher could contribute in saving the time. For this teacher the most worrisome problem was the large number of students in the class. She is of the opinion that the POE strategy required a step-by-step guiding of students that was not easy for her with eight groups of about five students to be monitored at the same time.

Effectively both teachers agree on two aspects that need to be considered for the improvement of the demonstration experiments: the time spent during preparation and the management of the class during the execution of the experiments. In relation to the time, the classroom tryout revealed that at least 10 to 15 minutes were used to set up the experiments before the lesson actually starts. However, the normal timetable, characterised by lessons conducted one after another after or following lessons of some other subjects, does not allow such time flexibility. Concerning class management, the problem lies on the large number of students per class. In order to get all students involved in the experiments, class A in school 1, which was comprised of 21 students, had to be grouped in four groups of four students each and a fifth group with five students. The number of groups and that of the students per group was even higher in school 2, where the class consisted of 41 students. Seven groups of five students and one group of six were formed in this school. It was, therefore, difficult for the two teachers to monitor all the groups at the same time in terms of providing relevant guidance and feedback for all students. As a result, teachers became mere explainers to all the students, rather than concentrating on guiding those with difficulties. Some of the most active students were indeed leading the discussions during the experiments but a significant number of other members of the groups, especially the low achievers, were passive and very dependent on the others.

Students' opinions about the demonstration experiments

Students' experiences and views about classroom tryout conducted by way of demonstration experiments were obtained through the students' questionnaire

(Appendix K) and the researcher's logbook. Two themes are used to present the findings on students' opinions about the experiments, namely aspects of the experiments students liked most, and aspects that they liked least.

(a) Aspects of the demonstration experiments students liked most

The overall results obtained from the questionnaire indicated that the students' experiences with the demonstration experiments, like those of the teachers, were positive. The majority of students liked the demonstration experiments, especially the three steps of the POE strategy, namely prediction, observation and explanation. They indicated that they were very fascinated by the way they were able to realise their misconceptions during the experiments and, based on the observation, to develop their own explanations of the discrepancy between prediction and observation of the events, and to draw informed conclusions. The following comments illustrate some students' reactions to the experiments. The comments are grouped in terms of (i) how students felt about learning Physics and (ii) how they perceived the usefulness of the POE strategy including their sense of enjoyment towards the demonstration experiments.

Learning of Physics

Under the aspect of how students perceived that the demonstration experiments can help them learning Physics better the following extracts can be quoted.

- *"I liked most the experiment nr. 2 because I was able to personally analyse in depth the experiment and to draw the conclusions on my own".*
- *"I felt like a great physicist when I realised the differences in accelerations of two blocks made of the same material and in which are exerted two equal forces".*
- *"On the experiment about the ball [experiment nr. 1], I was very surprised to realise that there are many forces acting during the entire flight of a ball kicked down. I did not know why the ball was taking long to come down when kicked into the air".*

From one perspective these extracts show that the students' understanding of Physics concepts was enhanced in the sense that it was possible for students to develop an understanding and to realise the existing interconnections between force, mass, displacement, speed and acceleration. For instance, it is evident from the second extract that, at least, this student has grasped the main idea of the experiment which

consisted of showing the variation of the module of acceleration as a function of force and the possibility to introduce the concept of inertia via the Newton's Second Law. However, from another perspective, these students' comments indicate that the demonstration experiments not only contributed to enhancing the students' understanding of Physics but also to uncover some common-sense beliefs about force and other related concepts, which seems to exist amongst students. For instance, analysing the last extract (about experiment nr. 1) it can be realised that students did not attend to the fact that the experiment occurs in an idealised system (as prescribed in the Grade 11 and 12 Physics Syllabus) where the force of air resistance is neglected. Rather, they are considering a normal system and they are interpreting the interaction of forces as a struggle between many (not necessarily opposing) forces. It appears for students that the ball takes a while to come to rest because of the number of forces acting upon it. According to the literature this could be seen as some kind of students' common-sense beliefs following from a metaphor that 'a win go to the stronger' (Hestenes et al., 1992). Still according to these authors, for the students, the ball is taking long in the air because forces are still 'fighting' and it will only come at rest when the stronger wins. They do not see the time spent by the ball during the entire flight as a result of the magnitude of the kick that can produce a smaller or bigger reach of the ball (X_{min} or X_{max}). This, in turn, can lead to some kind of 'dominance principle' according to which the most active agents produce the greatest forces. If this students' interpretation prevails it can hinder the students' understanding of the Newton's Third Law of motion based on action/reaction pairs.

Usefulness of the POE strategy and enjoyment of the experiments

The extracts below can be quoted on this respect.

- *"Starting any experiment by making your own prediction of the event is so fascinating that you never forget the results of the comparison [reconciliation] ... even if your thoughts are not equal to the outcomes of the experiment. This is what I liked most about the experiments".*
- *"I liked most the reconciliation between my own thinking [prediction] and the practice [observing the actual experiment]. After the experiment I was able to evaluate my own mental ability".*

- *“I enjoyed doing practical by myself and using locally available material. For instance, the experiment about ‘the piece of a card put on the top of a can’ was so real and in my group we had a heated debate because it was about something happening in our real life”.*
- *“On the experiment about the bottle [experiment nr. 4], it was so fascinating to see the bottle remaining at rest on the top of the table even with the paper being flicked away”.*

These extracts reveal that students felt the usefulness (including practicality) of the POE strategy, particularly the importance of making predictions. They seemed to realise that predicting means drawing upon relevant observations and data and saying something about a future of an event, even if that prediction does not necessarily coincide with the actual observation. This is indeed crucial for the POE strategy because it raises the students’ awareness and understanding of the process itself rather than the outcome.

The second extract shows also how students can improve their own reasoning by discussing and comparing (during reconciliation) their ideas with those of their colleagues. These comments also reveal the students’ positive attitude towards Physics demonstration experiments, with emphasis on the element of the student self involvement and manipulation of equipment. They enjoyed being involved in the experiments themselves and developing their own understanding of the physical phenomena. The fact that students had extensively discussed the experiment about the piece of a card put on the top of a can (refer to the second extract from bottom) shows how relevant it was for them to use locally available material based on their real-life situations. They appeared to be more motivated to discuss and share ideas if the events under investigation are about facts they experience in their own life and environment, which is in line with the recommended constructivist approach. An analysis of these extracts, however, shows a lack of students’ ideas about how they perceived the transition from their predictions to the reconciliation of the ideas, i.e., it was not clear on which basis they have made the predictions.

In general, however, the experiments seemed to have created or increased the student motivation to study Physics. The following statements from some students from school 2 are illustrative of this fact:

- *“I think we should have this kind of experiments more frequently because they help us to grasp the subject matter and boosts our interest in studying Physics”.*
- *“Why don’t you give us more experiments like these? I would like to practise more, particularly on my own at home”.*

Besides the positive comments that the students have given about the POE strategy, the analysis of the students’ responses to the question of what problems they may have faced with the use of the POE strategy shows another promising aspect of the strategy, which is the high level of student involvement in the experiments. This involvement is seen to be an important factor for raising the student critical thinking. Many students said that during the three steps of the POE strategy everybody in the class had to say something about what he or she thinks of the experiment being undertaken. According to students, the fact that at the end of each of the three steps of the POE strategy (prediction, observation and reconciliation) each student was required to justify his or her answer was important because each student had to explain the rationale behind his or her own reasoning. Nobody could keep quiet. This is a crucial element in enhancing student critical thinking in the sense that it requires students to reflect about the questions posed and to think about the events occurred before formulating an answer or taking any action. Still, according to students, the POE strategy enhances the support element amongst students in the class as illustrated by the following statement of one of the student:

- *‘...it [the POE strategy] allowed us to help each other’.*

By helping each other, students can develop a sense of collegiality which is important for raising the students’ levels of confidence in everything they can do. When students collectively agree on what and how to do to find a solution to any posed problem, they can develop a number of arguments against or in favour of each of their actions. Besides, the clarity (or lack of it) of these arguments is essential for the teacher to understand the level of students’ difficulties and be able to provide students a timely and relevant feedback about their actions.

However, as referred earlier, students’ experiences with the POE strategy were not only positive. One negative aspect mentioned was the students’ unfamiliarity with the strategy. Demonstration experiments in most of schools in Mozambican classrooms, as

already mentioned in Chapter 3, are not being carried out either due to lack of equipment or poor teacher preparation. Even in those schools where demonstration experiments are performed, they do not follow this strategy. This fact posed a challenge to most of the students during the tryouts because this was their first experience. The fact was evident in some of their answers to the question about what problems they had faced with the POE strategy. The answers were:

- *“It was a problem, particularly in the first experiment, because it is not used in our lessons”*; [Student 1]
- *“It was good although my colleagues did not approve some of my ideas”*. [Student 2]

Emerging from the answer of the student 1, it seems that student difficulties may have been linked more to the lack of familiarity with the strategy rather than its application in concrete situations but with more practice and exposure to the strategy, it can help them to do better. The answer from the student 2 shows that although the students are positive about the strategy, they still need to understand that what matters most is the strategy leading to the knowledge construction rather than the learning trajectory being followed. They need to understand that their ideas do not necessarily need to be the same. Rather, the most important step of the strategy is the reconciliation of the students' ideas where consensus is built around the intended learning outcomes prescribed for the experiment.

b) Aspects of the demonstration experiments students did not like

Some students from both schools listed a number of aspects, which they did not like. Fourteen out of 62 students (from both schools) indicated that they did not like the fact that, in the experiment nr. 1, they were not asked to represent graphically all the forces acting on the ball during its flight, namely the force of the 'kick', the force of air resistance and the force of gravity. They thought that a visual representation could have improved their understanding of the nature of these forces and their impact on the movement of the ball.

Few students from school 2 (six out of 41) indicated that they did not like doing the experiments in groups. They argued that they would prefer to do the experiments

individually because the sharing of ideas or answers, particularly during the reconciliation phase of the POE strategy, made the drawing of conclusions very confusing and time consuming; besides it was also difficult to describe the observed events of the experiments in groups of more than three students.

Students' experiences in connection with differences between demonstration experiments on PAM materials and regular Physics laboratory lessons

Students from school 2 indicated that they had never been involved in Physics demonstration experiments before. Specifically they stated that their normal classes were theoretically orientated with the teacher explaining how things would happen in an experimental situation. When informally asked by the researcher whether they had ever done some practical work before, they indicated that this was done through exercises normally given as homework. One student said:

"At the end of each unit we are often given some reviewing exercises to do at home. Some of these are corrected by the teacher in the next lesson".

The students stated that they were very fascinated at being involved in demonstration experiments of this nature because the worksheets were very well structured, practical and easy to follow during the course of the laboratory class.

Many students from school 1 (16 out of 21) indicated that the difference between the demonstration experiments based on PAM materials and their regular Physics laboratory lessons was that the lessons in the tryout, were closely facilitated by the teacher and were accompanied by worksheets with detailed instructions on how to do things. During the tryout, the teacher was more involved with explaining what to do and how to carry out the activity. In this regard, one student said:

"In our regular laboratory lessons the teacher would simply tell us what needs to be done and wait to see whether or not we managed to reach the desired outcome. Today's lessons were scientific, modern, and more interactive; nobody could keep his/her mouth shut."

Moreover students indicated that in the laboratory demonstration lessons on PAM materials they liked to discuss, interact, and construct their own knowledge, and they were encouraged to reflect on the experiments. They would first think about the

situation (during the prediction phase of the strategy), observe the course of the phenomenon (during the experiment), and compare their reflections with what actually happened (on reconciliation). This was very helpful for them in identifying and explaining the differences or similarities between their thoughts and the actual outcome. Students described the POE strategy as accurate and explanatory but they acknowledged the fact that it is somewhat time consuming. Students in general felt that the major difference between their regular Physics laboratory lessons and the tryout experiments was the guidance of the teacher and the time given to them to think about the phenomena and carry out the experiment before they were asked to draw conclusions from the experiments. It is worth mentioning that two students, who arrived late in the class, indicated that they found the worksheets difficult to follow even though they were immediately asked to join in the working groups. Apparently they had difficulties in joining in with the group particularly as they had no foundational knowledge on which to build their understanding. Three students did not respond to this question.

On the question on how they describe their participation in the experiments, the majority of the students (53 out of 62) indicated their participation as *active and interested*, because *“I was always contributing my own ideas”* and *“Because each of us had to justify his/her answers”*. The remaining nine students did not respond to this item. There were missing data in this item because it seems that some of the students are not familiar with the type of the question characterised by circling one or more options. Similarly, when asked what problems they might have faced during the execution of the experiments, they referred to the fact of that they had never previously participated in these laboratory lessons. As a result, students explained that: *“Some of the experiments were not easy to make a prediction (experiment number not specified)”* and *“The experiment about comparing the accelerations was difficult to understand”*.

Finally, when the students were asked for final comments or suggestions for improvement in the way the lessons were conducted, they indicated, among other things: the need for more time for them to be able to reflect about each item in the worksheet without rushing to another item; the inclusion of more topics in the experiment; the need for more frequent demonstration experiments of this nature; and

the need for experiments to be carried out individually to allow adequate visualisation. Apparently, the number of students per group had negatively influenced the participation of all students as some students actually had difficulty in seeing what was going on in the experiments. This is, however, a point of concern given the large class sizes of the majority of Mozambican classrooms.

Researcher's observations on the classroom tryout

The researcher's observations are presented in two parts corresponding to two observation phases. The first phase comprised aspects of teacher preparation for the classroom tryout in which the focus was on those aspects where teachers seemed to experience difficulties during preparation of the experiments. The second phase consisted of the aspects of both students and teacher difficulties during the execution of the experiments. More specifically, during this second phase, the researcher kept notes on how the teacher introduced the POE strategy and how s/he monitored the students' activities. Therefore, the researcher's observations are summarised under the headings of lesson preparation and execution of experiments.

(a) Lesson preparation

A week before the demonstration experiments, the teachers at both schools were provided with copies of the PAM materials for preliminary reading. They used the materials to prepare the experiments with some success and managed to try out most of the experiments in advance without too many difficulties. It seems that the teachers' reaction to the materials was positive. There were, however, some problematic aspects that were raised by the teacher from school 1 (the first school to conduct the tryout) during the meeting held a day before the experiments, which needed to be addressed in the next version of the prototype.

- There were concerns about the length of the PAM materials. The teacher felt there was too much to read and suggested that a shortened version, for instance, of the description of the assessment components (Part 2, Teacher's Guide) would be adequate. The teacher also indicated that the subsection on "Start of the lesson" (Part 3, Demonstration Experiments), consisting of brief questions for students to determine what they already knew about force and inertia

related concepts, might steal quality time from the experiments and could therefore be excluded from the Guide.

- In Experiment Nr. 2 (*Newton's Second Law*), the aspects such as (i) the reduction of the values of force (the value of the force was 6N) and of masses ($m_1=1\text{Kg}$ and $m_2=2\text{Kg}$), and (ii) the need to consider an average speed (rather than speed) when determining the value of speed of the masses m_1 and m_2 , were raised. All these problematic aspects were discussed between the teacher 1 and the researcher and the necessary amendments were made before carrying out experiments in school 2. The values of both forces and masses were found to be too high to be manageable in a demonstration experiment. As a result, the values of force were reduced to 2N, the m_1 and m_2 masses to 100g and 200g respectively and the students had to calculate average speed.

In school 2 the experiments were carried out with relative success compared to school 1.

b) Execution of experiments

The execution of the experiments in both schools started with the teachers organising students into groups. In school 2, due to the high number of students ($n=41$), the teacher did not follow the format suggested by the PAM materials (a maximum number of four students per group). Instead, seven groups of five students each and the eighth with six, were formed. After forming the groups, the teachers in both schools introduced the students to the objectives of the experiments, explained the working methodology with emphasis on the POE strategy, and asked each group to appoint a chairperson. When all groups had been formed, the teachers distributed Worksheet 1 (about the prediction phase) and allowed five minutes for student prior reading but did not ask students to pose any questions.

As the students were conducting the experiments, teachers walked around providing guidance and support where appropriate. Some students had difficulty in using correct equations for determining the values of acceleration and average speed on the Experiment Nr 2 (*Newton's Second Law*). As this phase of the experiment was carried out in groups and with the teacher support, the difficulties were alleviated. This

experiment, however, appeared to be the most laborious and time consuming of all. Discussions about getting the experiment well set up and applying the right equations for calculations continued beyond the estimated time and teachers, sometimes, seemed to have difficulties in handling the working environment. At a certain stage the researcher intervened and assisted the students. As a result, very little time was left for drawing conclusions and rounding off the lesson, particularly in school 2.

During the experiments, the practical activities seemed to be a motivation for the students and facilitated their learning. Some students argued that experiments like these could help them become good physicists, if these were undertaken more regularly. As already indicated, some of the students said that individual work is more productive than group work because of the value of discussions. However, from the researcher's point of view, this is an issue for concern because classes are large and individual practical work could prove difficult to conduct in terms of classroom management, unless teachers are capable of arranging students into small groups or pairs and then manage and monitor the discussions. Like their students, teacher involvement and motivation was also good despite the difficulties in time management. However, in some cases, groups did manage to finish their experiments within the estimated time.

An analysis of the demonstration experiment report

At the end of the demonstration experiments, students in both schools were given a demonstration experiment report template (see Appendix B of the PAM materials). The template contained an outline which would guide the students when writing their reports. It was suggested that the report was divided into a number of titled sections, such as purpose, procedure and theoretical background, as the required length of each section was indicated. To facilitate understanding, each section title was explained in detail. Specifically, the report was meant to assess the student ability to design and conduct demonstration experiments as well as to communicate experimental results. Student success in this task was evaluated by the teachers using 0-20 point scale. In general, the report was designed not only to assess the students' awareness of what should be included in a demonstration experiment report and how it should be put together, but also to support teachers in evaluating the level of student performance in this type of assessment practice.

Initially, the report was meant to be individual but, due to the large number of students per class, which implied that students be organised into groups, it was decided between the teachers and the researcher that reports be written in groups. The groups were made up of the students who worked together during the experiments. Reports were submitted in a one-week period of time as planned.

All five groups of school 1 submitted their reports while, from school 2, reports were submitted by six out of eight existing groups, making up a total of 11 reports. As referred to earlier, classroom tryouts were conducted in a school period close to the final examinations. According to the teacher T2, this seemed to be the reason why not all groups managed to return their reports; apparently students were busy preparing themselves for the examinations.

An analysis of the reports from the students' perspective indicated that they produced descriptive rather than substantive reports, which would illustrate a deeper understanding of the content. Sections about procedural information such as *purpose* and *procedure* were more accurate and explanatory of the tasks undertaken than those which required an ability to argue how they have developed their reasoning during the prediction and explanation stages of the POE strategy (*theoretical background* and *data* sections). Although the *conclusions* section also included the student ability to describe the focus of each experiment, the identification of potential sources of error and the discussion about how such errors impacted on the results proved to be difficult for students.

From the teachers' perspective, although the report template had explicit guidelines on what to include in each section, it seemed that the teachers focused their attention more on students' manipulative skills during the experiments. Their comments on student reports were more on controlling the students' description of adequate execution of the experiments (manipulative skills) dealt with in the *procedure* section, rather than other investigative skills. Aspects about (i) concept building or concept acquisition, where students would present well-described experiments and discuss the assessed concepts, and their connections with other related concepts (*data* session), (ii) the assessment of theoretical knowledge base (*theoretical background*), and (iii) the student ability to

conduct and communicate the results of the experiments (*conclusion*), were not adequately dealt with by the teacher feedback to students. This fact, however, is not surprising because teachers, like their students, lacked investigative skills needed to write reports.

Overall, the combination of both perspectives (teacher and students') indicates that the PAM materials should also pay attention to providing support for student ability to communicate experimental results based on an integrated instructional process, investigative and manipulative skills, and the drawing of informed conclusions. Regarding teachers, the materials should provide support aimed at enhancing their critical thinking not only on procedural aspects of the experiments, but also on those aspects that require students to argue about experimental situations on the basis of sound theoretical foundations.

Having presented and discussed the findings of the second prototype, the following subsection discusses how this prototype was formatively evaluated to include suggestions given by teachers and students.

6.3.5 Formative evaluation of the second prototype: The way forward

Suggestions from the experiences with the tryout were then used to design the third version of the prototype of the PAM materials. The findings of the various stages of the classroom tryout show that, drawing on the teachers' experiences of the demonstration experiments and use of the PAM materials, the advice was to find a balance between the number of experiments to be conducted per lesson and the time allocated to perform the experiments.

The teachers felt that, although the POE strategy seemed to be very effective in raising students' motivation and willingness to perform demonstration experiments, it is time consuming because it involves discussions, building of consensus amongst students, and close facilitation from teachers. This balance should not necessarily be sought by adding more time to the timetable but by utilising other practices such as distributing worksheets to students before the actual demonstration experiment starts to allow for prior reading. The teachers noted also that they needed to read the Teachers' Guide

well in advance and prepare the experiments before the normal time schedule. This forward planning and preparation illustrates the importance of teachers' commitment, which is seen as a successful factor in obtaining good results from the use of experiments in Physics learning.

Three sections of the PAM materials were considered by the teachers as the most helpful for their teaching activities, namely the Practice-Oriented Lesson Plan (Appendix P, Part 2, subsection 4), the Glossary of Terms and the Demonstration Experiment Report Template for Students (Part 5). However, they considered subsection 1.1 of the Teachers' Guide too long and suggested that it be reduced to allow a friendly use.

Students' comments about the second PAM prototype suggest that the next step would be for students to develop graphical representations on those experiments involving forces acting on a ball. A visual representation would allow the students to improve their understanding of the nature of these forces and the impact of these forces on the movement of the ball. Some of the students expressed dissatisfaction about doing the experiments in groups because the sharing of ideas during the reconciliation phase of the POE strategy made the drawing of conclusions confusing and time consuming. Students' experiences about the differences between demonstration experiments on PAM materials and regular Physics laboratory lessons highlighted the role of the teacher during the tryout as crucial to the successful running of the experiments. More comments from students about their experiences referred to the apparent unfamiliarity with question items characterised by circling one or more options and, for some of the experiments, it was not easy to make a prediction.

In writing the demonstration experiment report, students appeared to experience difficulties with those sections that required the ability to argue a point to demonstrate their reasoning (*theoretical background section*), which would show how potential errors could impact on the experimental results and consequently lead to drawing wrongly informed conclusions (*data and conclusions sections*). Interestingly these are the same sections where teachers also revealed difficulties in providing effective formative feedback. This suggests that support is needed for both students and teachers in developing in-depth thinking and reasoning skills.

As a conclusion, students' opinions and experiences about classroom tryout reveal that students appeared to like demonstration experiments. The experiments allowed the students to verify theory for themselves, and increase their understanding and awareness that Physics is related to real-world events. In addition, demonstration experiments gave them a sense of control over their own learning. Within the research literature this corresponds to affective aims of demonstration experiments in Science teaching where the use of Science demonstration experiments in a constructivist approach can effect conceptual change (Dekkers & Thijs, 1995). Another aim of the experiments is linked to clarification of the scientific method, or enhancing problem solving skills (cognitive aim). This aim, however, does not seem to have been fully met during the experiments. Students seemed unable to describe and explain the events as accurately as possible. For instance, they were unable to provide explicit reasons for their predictions of the events and how these differed from their observations. Finally, there is the aim of development of investigative skills, such as applying research methods, formulating hypotheses and drawing conclusions from data. This aim was intended to be addressed through the preparation of demonstration experiment report. In this perspective, although the report template contained substantive information about *purpose, procedure, theoretical background* and *conclusions* from data collected, students' reports were more descriptive than explanatory of the student ability to argue and explain their reasoning during the prediction and explanation stages of the POE strategy. Without completely neglecting the first aim (affective aim), the third prototype will attempt to address the last two aims namely cognitive and investigative skills aims.

All these comments and suggestions were taken into account and informed the design of the third prototype of the PAM materials.

6.4 Design of the third prototype

The design of the third prototype was based on the comments and suggestions from the classroom tryout. The main focus of the revision was to gather information on the practicality of the PAM materials in terms of layout, structure, and content presentation so that the next version would have strong elements on *effectiveness*. In

relation to the layout and structure, the prototype had the characteristics described below.

- *Presentation:* The third version of the prototype contained outlined numbering of headings to differentiate various types of information. The main headings are 1.0 Introduction; 2.0 Objectives of the lesson; 3.0 Teaching strategy; 4.0 Assessment criteria; 5.0 Demonstration experiments; 6.0 Assessment rubrics; and 7.0 Demonstration Experiment Report template.
- *Target groups:* Initially the PAM materials included information for teachers and students without separating sections for each target group. The new version showed a clear distinction between the Teachers' Guide and student worksheets.
- *The prototype:* Apart from the distinction between the support materials for the teacher and worksheets for students, these worksheets were included in the prototype and the length of the subsection 1.1 of the Teachers' Guide was also reduced as suggested by both experts and teachers.
- *The teacher's practice-oriented plan:* The subsection on preliminary discussion about force and inertia and their related concepts was removed from the teacher's practice-oriented plan in order to make the materials user-friendly.

The content presentation was also reviewed. The third prototype highlighted the main characteristics of the POE strategy and its relevance in investigating student understanding of Physics concepts. The presentation of content was specifically reviewed in the items listed below.

- *Teaching phase:* Greater detail on how the POE strategy can serve both introductory and consolidation phases of teaching and learning concepts were added in the Teachers' Guide. This was seen as a need to assist teachers understand the shift from assessing *of* learning to assessing *for* learning.
- *Lesson plan and timing:* The lesson plan and timing of the experiments was revised to accommodate the four experiments plus the assessment related activities that were not included in the previous versions.
- *Demonstration experiments:* The wording of the experiment nr. 1 (*Introduction to the force concept*), which required the students to name '...the force (s) acting on the ball' was reformulated to specify '...the force (s) acting on the

ball *during its entire flight*' and emphasised that the experiment occurs under an idealised system where the force of air resistance is neglected. A picture showing the kicking of the ball and its trajectory was also included. The experiment nr. 2 (*Newton's Second Law*), which proved to be time consuming and in which students showed difficulties, was replaced by another demonstration experiment on comparing different forces acting on moving objects. This new demonstration experiment was intended to help students identify and compare forward and backward forces exerted on a moving object at constant speed (refer to demonstration experiment nr. 2 in Appendix P).

- *Demonstration Experiment Report template*: The content of the report template was also reviewed on the basis of the analysis of the reports received from students and the teachers' feedback to students. Improvements were made by giving additional explanations for the section of 'theoretical background' to allow students a better understanding and to assist in developing their reasoning skills. The improvements included asking students to focus explicitly on the two investigated concepts (force and inertia) and not on any other related concepts that they may have dealt with during the experiments. The new wording no longer named the related concepts to be included in the discussion with the intention to allow that students themselves be able to identify those which are worthy of discussion. It was also noted that the initial wording required excessive theoretical information (e.g., a discussion about diagrams, graphs, tables, and other visuals that the students had seen during their lessons), and this made the section lengthy and less specific. As a result, the number of pages for this section was reduced from four to three.
- *Assessment rubrics*: Students must be made aware of the levels of performance expected of them. Therefore, teachers were now given about 20 minutes to explain to students the assessment rubrics being used to assess their performance in the experiments.
- *Correction guide*: Correct responses to the experiments as well as clarifications, where deemed appropriate, were to be provided in the third version of the prototype.

Following this revision, the third prototype was produced and appraised by two experts in an interactive discussion. The results of the experts' interactive appraisal are presented in Section 6.5.

6.5 Final appraisal resulting in the fourth prototype

The appraisal of the third prototype that resulted in the fourth version was undertaken in two stages. The first stage consisted of expert appraisal and the second of an evaluation workshop with teachers. Both stages were used as a systematic reflection and documentation of the materials in terms of their expected *effectiveness* and *sustainability* of the study.

For the expert appraisal, two experts – out of the three who appraised the first prototype - were asked to focus their attention less on the practicality and more on the effectiveness of the PAM prototype including the effectiveness of the intervention as a whole. These experts were selected for their ample experience in designing curriculum materials, and in assessment and evaluation. Other areas of interest also included projects evaluations and the relationship between language (particularly English) proficiency and (mathematics) achievement. Although the experts involved in the appraisal of this version gave positive comments, indicating that they perceived the material as useful in supporting teachers in the improvement of their assessment practices, they felt that the focus of the intervention – improvement of teacher assessment practices – was not effectively addressed. They also expressed concerns about the structure of the material overall (refer to the four support levels presented in Chapter 4, subsection 4.4.3). Regarding the focus of the intervention, they suggested that this could be done by explicitly addressing the aspect of teachers' provision of feedback to students based on the assessment strategy applied.

The experts also suggested revision of the structure of the material, taking into account the four support levels (subject knowledge, lesson preparation, teaching methodology and assessment and feedback), and their advice focused on the need to present the material in the PAM prototype according to the way teachers would use it in the classroom. More specifically the experts' advice emphasised improving the specificity of the material to make sure that there is clarity on how teachers are to collect evidence

(i.e., written? per group discussions? whole class discussions?), when conducting formative assessment, and then how to use the evidence to develop a formative assessment.

For the evaluation workshop, two groups of material users were invited. Three university students who had appraised the prototypes in the early stage of the intervention, and the two teachers who had tried out the materials in the classroom, participated in the workshop. The university students are also Physics teachers teaching in high schools. All participants were asked to evaluate the materials and reflect on their *effectiveness* before implementation on a large scale. Aspects of effectiveness included, amongst others, the use of the different functions of assessment for designing classroom assessments, the use of the practice-oriented lesson plan for preparing and conducting lessons, the application of the design guidelines for monitoring demonstration experiments and providing feedback, and the evaluation of the proposed demonstration experiments in terms of the practicality of the POE strategy.

In order to capture the reflections of the participants in the workshop, Guskey's model of evaluation (2000) was employed during the workshop. According to this model, five levels of evaluation can be taken into account from the users' initial concerns about an innovation to what might happen in practice. These levels are (i) teachers' first reactions, (ii) teachers' learning, (iii) nature of school support, (iv) teachers' new knowledge and skills, and (v) student learning outcomes. Table 6.6 shows the evaluation levels of the Guskey's model with the main questions linked to each level.

Table 6.6: The five levels of the Guskey (2000) model as applied to this study

| | | |
|---|------------------------------------|--|
| 1 | Teacher reactions | <ul style="list-style-type: none"> Did the Physics demonstrations experiments meet the teachers' expectations? Did they consider the content, process and context of the experiments useful and relevant? |
| 2 | Teachers' learning | <ul style="list-style-type: none"> Did teachers acquire the intended knowledge, skills and attitudes towards demonstration experiments? |
| 3 | Nature of school support | <ul style="list-style-type: none"> Did the respective schools support the implementation of the Physics demonstration experiments? |
| 4 | Teachers' new knowledge and skills | <ul style="list-style-type: none"> Did what teachers learn affect their classroom assessment practices? |
| 5 | Student learning outcomes | <ul style="list-style-type: none"> What would be the impact of the Physics demonstration experiments on the student learning outcomes? |

During the evaluation workshop, the emphasis was given to Guskey's first, second, and fourth levels of evaluation, which focus on teachers' reactions to the intervention, teachers' lessons from the demonstration experiments, and the effect that the new knowledge and skills that the teachers have gained might have on their classroom practices. The reason for not considering the third (nature of school support) and fifth (student learning outcomes) levels is that these are long-term indicators, which could not be verified at the time of the study. Assessment materials need to be empirically tested with the target population and fully implemented on a large scale to be able to effectively evaluate the impact of the intervention on the student learning outcomes as well as the level of school support to the innovation.

Workshop findings about teachers' reactions (first level) and their learning about demonstration experiments (second level) indicate that the content addressed by the experiments is relevant for student learning and the context in which the demonstration experiments are carried out, is also appropriate. However, the teachers appeared to be concerned about the process of implementation of the intervention. The concerns are at students' management level: The step-by-step POE strategy requiring close monitoring by the teacher might be problematic if suitable class management strategies are not adopted. These concerns were shared by both teachers and university students. They were all of the opinion that it takes time to effect a change and as such, they will need to implement the innovation more regularly in their classrooms to be able to evaluate the impact of its outcomes. In fact, they attested that, after participating in the tryouts

and studying the exemplary PAM materials, they were willing to try out the materials in their respective classrooms.

Guskey's fourth level highlights the effect of teacher learning on changing their classroom assessment practices. The findings show that teachers have gained some additional knowledge input and experience that may improve their skills regarding the design and monitoring of appropriate assessments (particularly performance-based assessments) following the POE strategy. Although it was also difficult to get a clear picture of how teacher learning could affect their practice, it became apparent that the participants as a group had improved content wise and that there was a willingness to engage in new practices which is, in its own, a sign of success. There were, however, some concerns at a personal level. A few of the teachers seemed to be unfamiliar with some of the assessment terms or concepts used in the PAM materials. According to the participants, support from their school Boards and from the MEC via professional development workshops could help to deepen their knowledge and develop and sharpen their skills. Without this kind of support, a shift from the actual teaching and learning routine might prove to be difficult due also to its implications on schools' timetables. Once more, to realise the effective improvement of the actual assessment practice requires time.

All these questions, suggestions and comments were taken into account in designing the fourth and final prototype of the PAM materials (Appendix P). The final prototype consisted of five parts namely Part 1 (Introduction for the teacher), Part 2 (Components and functions of assessment), Part 3 (Design guidelines and feedback provision), Part 4 (Demonstration experiments), and Part 5 (Student worksheets). As a result of the suggestions from the classroom tryout, appraisal, and evaluation workshop, substantive changes were made to Part 2 (Components and functions of assessment) and Part 4 (Demonstration experiments). The glossary of terms used in the prototype that constituted a section in its own in the first prototype, was embedded in Part 5, and a new Part 3 (Design guidelines and feedback provision) was added to explicitly address the issue of designing, monitoring and assessing demonstration experiments, as well as providing feedback to students. In summary, relevant changes resulting in the fourth prototype included the contents of Parts 2 and 4 as well as the creation of a new Part 3. These changes are discussed below.

Part 2: Components and functions of assessment

The core of this research is to improve teacher assessment practices. To achieve this purpose, explicit attention was given to support of teachers in designing and applying appropriate assessment strategies for student learning. This section provided firstly an explanation of *components and functions of assessment* to help teachers develop their own assessments. Secondly, a practice-oriented Teachers' Guide containing the *sequence of content and lesson plan, some logistical aspects*, and a *plan* on how to teach and assess following the POE strategy, was given. The Teachers' Guide (practice-oriented lesson plan) was structured in such a way that it follows the four support levels discussed in Chapter 4 (refer to subsection 4.3.2.3) with emphasis on assessment and feedback for the teacher on how to conduct the experiments, what to assess, and what type of feedback should be given to students.

Part 3: Design guidelines and feedback provision

This section provides guidelines on how to design and monitor demonstration experiments in the classroom and to facilitate student learning through provision of formative feedback. Five design elements, as taken from literature, (refer to Chapter 4, subsection 4.3.2.3) are discussed in this section and they are listed below.

a) Agreement - *the teacher and the students must agree on the relevance of the problem to be investigated, the procedures to be followed, and the conclusions of the evaluation of the explanations given during the experimental work.* This means that the students must understand the relevance of the problem being investigated. As Dekkers (1997) and Tamir (1991) indicate, (i) the teacher's intended purpose should become the students' own purpose, (ii) the activity designed to achieve this purpose is to be understood and accepted by the students, and (iii) the students' conclusions are to be discussed, valued and related to the teacher's hoped-for conclusion. By getting involved in all these stages, students would then understand the intended purpose of the practical work and perceive the task as relevant.

b) Intended learning outcomes – *the teacher must be prescriptive about the ideas that the students are supposed to acquire and develop* (Dekkers, 1997). The students must understand the procedures to be followed in order to achieve the proposed ideas.

This is important as a criterion to verify the effectiveness of the experimental work. It is, therefore, crucial to be aware that these ideas may not be acquired at the end of the task. If this happens, however, it means that the teaching process or the design and implementation of the experiment may have not been successful and a revision of these aspects is needed.

c) Student participation – *In experimental work, particularly in demonstration experiments, the teacher must produce the event to be investigated according to the purpose to be achieved, while the students attempt to interpret it and make sense of it.* In so doing, the teacher may find a balance between his/her expository approach (which has its own educational value) and the student-centred exploratory approach. An intended conceptual development can equally take place when having each student handling the equipment her or himself or when some students observe others handling the equipment but acquiring information about the event under investigation (Garrett & Roberts, 1982). One of the most important aspects in students' participation in the experiments is that they should be able to relate the proposed task to previous activities, and the understanding of the relationship between the proposed procedure and the purpose to be achieved is crucial, even if they do not handle the equipment directly.

d) Type of experiments and aims - *Teachers must avoid having too many experiment aims to be achieved at once. This may lead to none being pursued (Hodson, 1993). Rather, they must select the proper experiment for the chosen aim and match the written instructions.* Students should not be involved in activities that may distract their attention from the aim of the experiment. Several authors point out that side-issues such as getting the equipment to work, data acquisition, extensive data manipulation, and complex measuring procedures are all 'noise' that can lead the students to perceive each of them as the main purpose of the practical (Gunstone, 1991; Johnstone & Wham, 1982; van den Berg & Giddings, 1992). It is suggested that experimental work must rather concentrate on qualitative observations (e.g., in concept acquisition) and finding a balance between a too complex and a very trivial task.

e) Critical thinking and reporting – *Teachers are to make sure that students develop a critical attitude towards their actions and interpret the activity data only in the light*

of the experimental work pursued and of their own knowledge and experience. Students should also be able to summarize and report the main aspects of the experiment including the central aim and outcome, the basic methods applied, and the underlying theory of the demonstration experiments.

In relation to feedback provision, a list of aspects aimed at supporting teachers on how to provide formative feedback to students, particularly during the course of the experiments, is provided (see also Chapter 4, subsection 4.3.2.3). The list contains procedural information on how to support teachers from lesson preparation to lesson evaluation (summative assessment). The reason for selecting these aspects is the recommendation by the literature as being effective for monitoring the student learning during performance assessments and were successfully used by a similar study about practical work (refer to study by Motswiri, 2004). It is, however, important to note that this list does not intend to suggest reinforcement of rather traditional (i.e., teacher-centred) curriculum implementation of Mozambican teachers but it deliberately contains statements on what is perceived to be relevant teacher actions for the context of demonstration experiments. These elements of feedback provision indicate that when facilitating demonstration experiments, teachers must be able to:

a) Lesson preparation

- Take time to read the support materials and reflect on the experiments well in advance. It helps clarify ideas about the outcomes being pursued.
- Assemble and try out each experiment before the actually lesson starts. It is crucial for detecting potential problems (e.g., shortage of equipment, time constraints for conducting the experiment, inappropriate set-ups and procedures).

b) Course of the lesson

- Start the lesson by asking brief introductory questions to students on what they already know about concepts or events to be investigated.
- State the objectives of the lesson, clarify the outcome to be achieved at the end of the experiment(s), and explain the teaching and assessment methodology to be followed (including the procedures).

- Observe what students do and ask probing questions to help them reflect on their activities. This is important to focus students' attention on important elements of the experiment.
- Encourage students to discuss amongst each other. It helps to develop their own models of learning and the capacity of the class to function as a community of learners.
- Give opportunity to students to reflect on their own tasks and on those of their colleagues in a critical way.
- Keep in mind that the teachers' role in the experiments is to help students develop their reasoning, and act mainly as a moderator.

c) End of lesson

- Provide immediate feedback to students (when asking probing questions) so that they understand to what extent they are achieving the intended purposes. The feedback should preferably be individual and either congratulatory or critical.
- Round off the lesson by providing a summary of the main conclusions of the demonstration experiment. Give students homework and ask them to prepare a short report about the experiment(s). Due to large classes and time constraints, a follow-up to the homework and experiment reports can be given during following lessons.

Part 4: Demonstration experiments

The fourth prototype was comprised of four demonstration experiments following the POE strategy. Substantive changes refer to the fact that in this section each of the experiments had applied the design guidelines and the specific elements of feedback provision presented and discussed in Section 3 of the PAM materials. All aspects of lesson objectives, intended learning outcomes, activities during the experiments, assessment and feedback (on monitoring and provision of feedback), and summative evaluation of the experiments (using assessment rubrics) were taken into consideration in this version. At the end of the section, a glossary of terms used in this prototype and a report template to guide students in preparing a report to summarise the experiments are given. The final version of the fourth prototype is presented in Appendix P.

Section 6.6 summarises the conclusions drawn from both the classroom tryout and the evaluation workshop, and discusses the design principles, which were used to address, in the PAM materials (refer to Appendix P, Part 3), the issue of designing and monitoring demonstration experiments, as well as the teachers' provision of formative feedback to students.

6.6 Conclusions and implications for further development

The central research question of this study, which is addressed by the intervention study, investigates how teacher assessment practices could be improved. Based on preliminary information obtained from the Baseline Survey and on what emerged from the literature as good practice, a choice was made to design and try out assessment materials on performance type of assessment in the context of demonstration experiments. It was assumed that, by supporting teachers in designing and trying out assessment prototypes in one type of assessment practice and in a given context, the teachers would be able to transfer their knowledge and skills to other assessment practices. A number of prototypes were designed and subsequently evaluated to examine the quality of the lesson and assessment materials in terms of validity, expected practicality and expected effectiveness. While the concepts of validity and practicality were given more emphasis during the prototyping phase of the intervention, expected effectiveness was more emphasised during the systematic reflection and documentation which were carried out through the evaluation workshop and final expert appraisal.

Teachers' general impression about the PAM prototypes and their opinions about the experiments show that they perceived the materials to have the potential to improve their assessment practices in the context of demonstration experiments in their classes, because the suggested POE strategy allows students to be actively involved in meaningful individual and group tasks. On the practicality of the materials, teachers were seen to have problems in terms of its length – particularly the Teachers' Guide. Teachers were also observed to find problems in developing a meaningful formative assessment orientation during the experiments and drawing lesson conclusions. This fact was due to the reported lack of explicit advice on feedback provision from one

side and on the lack of sufficient time to conduct the experiments and guide final discussions leading to final conclusions. Teachers also perceived the exemplary materials as being inconsistent with their ordinary classroom practice, because they involve additional time and costs, which teachers perceive as hindering the successful implementation if not taken into account by all educational stakeholders, namely students, teachers, school leadership, Parents' Councils and the MEC.

Students' comments on PAM prototypes and on the experiments showed that, although they are unfamiliar with the POE strategy, they enjoyed the materials and the activities and would like more practice and exposure to the strategy. They perceived the worksheets as being easy to follow and the strategy as wonderful in terms of the potential to raise their chances of being engaged in addressing the problems and generating solutions by voicing their own ideas. Students appreciated the role of the teachers as facilitators during classroom tryouts and their own roles as active students.

Experts have also indicated that the PAM materials have great potential in improving the teacher assessments. However, they emphasised that explicit attention needs to be paid in providing guidelines to teachers on how to apply assessment practices for learning and on provision of effective formative feedback to students.

Drawing from the reflections of the final experts' appraisal, the evaluation workshop findings, and in combination with relevant literature reviewed on this respect (Chevane, 2002; Dekkers, 1997), a number of design principles for successful student assessment, in the context of demonstration experiments, are to be taken into consideration. These principles are listed below.

- Clarity - students should understand the intended purpose of the experiment.
- Relevance - students must perceive the laboratory experiment as valuable to assist in learning.
- Understanding - students must understand the procedure to be followed during the experiments and be able to execute it.
- Relationship - relating the proposed task to previous activities and understanding the relationship between the proposed procedure and the purpose to be achieved is crucial.

- Critical thinking - students must be challenged to use their knowledge and experience to develop their critical thinking which would assist them in the conducting of the experiments, the interpretation of evidence and the drawing of conclusions.
- Reporting - students should be able to summarise the main aspects of the experiment including the central aim and outcome, the basic methods applied, and the underlying theory of the experiment.

Many examples in the literature report positive experiences in students learning but they are not necessarily suitable for enhancing the assessment process. For instance, Tamir (1991) refers to some principles which bring about positive effects for instruction in general, and are based on a learning-cycle approach. This approach stimulates students to formulate questions and the teacher to introduce and explain new concepts, as well as apply results in new situations. The principles listed above were selected for being suitable for the Predict-Observe-Explain approach suggested by this study. They intend to support teachers in designing and applying meaningful assessments. 'How-to-do' advice on monitoring the process of assessment and providing feedback was given in Section 6.5. For instance, on the issue of reporting, students found it difficult to summarise meaningfully the important aspects of a small demonstration experiment they had just undertaken but rather recall some of their manipulations in the laboratory. In this case, the teacher should be able to provide students with the opportunity to reflect on their own tasks, develop critical thinking (e.g., by asking pertinent questions), and then understand to what extent the students have achieved the intended purpose. This can be achieved by following the design guidelines and elements of feedback provision discussed in Section 6.5.

It was against the findings from the classroom tryout, the suggestions from experts, as well as in connection with what the literature says are good practices for classroom assessment that the current version of the PAM materials was designed. Although the potential of the materials is acknowledged, it is important to note that the fourth version of the PAM materials cannot be seen as final because, as started in earlier chapters (see, for instance, Chapter 4, subsections 4.3.2.1 and 4.3.2.2), the study has only capitalised on the potential of the material for improving assessment practices for teachers. This potential was sought through investigation on expected practicality and

expected effectiveness. The materials can only be accepted as finished products when its final version has been tried out repeatedly with a number of users in schools. It is also noteworthy to state that there are still areas of improvement both in terms of the quality of the material itself – the process of teachers' feedback provision – and of the implementation of innovation within the current classroom context which is characterised by overcrowded classrooms, syllabus time constraints and limited school budgets. For instance, the issue of overcrowded classrooms is a reality for the majority of African schools. In the Mozambican context, if this is not adequately addressed, it could make the proposed POE strategy in the context of demonstration experiments a difficult approach to implement. Literature (Chevane, 2002; Cossa, 2007; Dekkers, 1997; Hodson, 1993) shows that the strategy can be successful if students are working in small groups. These authors argue that there are several factors to consider when designing and implementing the strategy particularly in the context of experimental work. These factors, which serve as guidelines for the design of successful experimental work, are summarised below.

- *Closed or open-ended*

A question about practical work (e.g., demonstration experiments) that is very often discussed in the literature is whether formulating the problem, designing an execution strategy, and drawing the conclusions should or should not be left to students alone (Dekkers, 1997; Hodson, 1993; Tamir, 1991). Tamir (1991) presents the degree of openness on students' inquiry skills at four levels (Levels 0 to 3). At Level 0, problem, strategy (or procedure), and conclusion are all given to students and they only need to collect data and check whether these data are 'correct', while at Level 3 all activities are carried out by the students. The literature, however, argues that what matters is not whether all inquiry skills are predetermined or not, but whether the teacher and the students agree on the relevance of the problem, the adequacy of the strategy, and the conclusion of the evaluation of explanations given.

- *Prescription of the learning outcomes*

An evaluation of the effectiveness of the demonstration experiments is a very crucial step. Before guidance and support are provided to allow teachers and students to reach the intended outcomes, it is necessary to prescribe the ideas that need to be acquired and understood. If, however, the prescribed ideas are not met, this would mean that

some decisions were not well taken care of (e.g., the procedure followed was inadequate, the designer's arguments for reaching the ideas were different from those of the students), and the teaching and assessment strategy will need to be reviewed.

- *Course of the demonstration*

A decision is needed on whether the students should handle the equipment or material themselves or whether the intended outcome can be reached by only having them observing other students (or the teacher) doing the experiments. This decision, however, should consider that demonstrations may have the value of student-centredness where the teacher (or other students) may produce the event, while the students interpret and make sense of it. But an important factor, as illustrated by some student statements during laboratory demonstration, is that having each student handling the equipment is better and gives them the opportunity to 'play'.

- *Relationship between the type and complexity of the experiment and the aims*

Too many experimental aims to be met at once can lead to none of them being attained (Hodson, 1993). It is important to avoid issues that can distract students from the actual purpose of the experiment. As Johnstone and Wham (1982) point it, this is 'noise' that can produce counter-productive effects. For instance, if the teacher does not clearly define the problem to be investigated, does not indicate the procedure to be followed, or does not get the apparatus to work in time, each of these could act as the main purpose of the experiments for students. Furthermore, besides clarity in defining the aim, it is also important to select adequate experiments for the chosen aim and match the written instructions with these (van den Berg & Giddings, 1992).

Some 'open ends'

As experts have indicated, the PAM materials need to have clearly formulated design principles for their successful use by teachers and students in the context of demonstration experiments. It was argued by teachers and supported by experts that the POE strategy is an approach that can optimise student learning if it is accompanied by continuous provision of feedback. This argument is valid when taking into account that more emphasis was put on the lesson materials in the early stages of the intervention and only later, on assessment strategies and feedback provision. This reinforces the idea that the POE strategy combines instruction with assessment.

The teachers who participated in the classroom tryouts and in the evaluation of the intervention, also perceived the materials to have the potential to improve their assessment practices. Whether its successful implementation would actually take place depends on the involvement of all other educational stakeholders (e.g., the school leaderships and the MEC) particularly in the way they might address the involved costs and additional time required. Relevant literature on experimental work has provided promising design guidelines that can support teachers and students in successfully implementing the changes but it has been argued that the guidelines are mostly effective when working in small groups of students. The fact is that with or without overcrowded classrooms, a successful implementation of the innovation, within the current Mozambican classroom context characterised by syllabus time constraints and limited school budgets will need time to be achieved. But one thing is certain: a lack of support for teachers in designing and implementing alternative assessment strategies that contain procedural specifications on how to provide effective feedback to students can be detrimental for the future of student learning.

The final version of the PAM materials designed in this study contains specific design principles for designing and implementing successful demonstration experiments. Further evaluation research on the impact of the implementation of the suggested assessment practices and innovations is required, given the fact that, although the classroom tryouts were adequately developed and conducted, time constraints did not permit an opportunity to try out the fourth version, evaluate it, and draw definite conclusions about the long-term impact of the intervention.

CHAPTER 7

**CONCLUSIONS AND
RECOMMENDATIONS**

CHAPTER 7

CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the conclusions of the study drawn from the two phases, and the recommendations in relation to three different perspectives. It starts by summarising the context in which the study was undertaken, the research questions and approach (Section 7.1). The main findings emerging from the Baseline Survey and from the Intervention Study are also discussed in this section. The discussion is presented along the two stages of the intervention phase and their impact on teachers and students. Section 7.2 further discusses methodological, substantive, and scientific reflections of the study. Section 7.3 summarises the main conclusions of the study. The chapter ends with Section 7.4 which presents recommendations for policy and practice, for further research, and for further development work.

7.1 Summary of research questions and approach

In order to be able to say anything about the improvement of assessment practices, it is important to consider the methods used to collect the data and whether the data collection instruments were fair to all participants (Pelgrum, 1989). The aim of this research was to investigate assessment practices used by Grade 12 Physics teachers in Mozambique and how can they be improved upon. What follows in this chapter is a consideration of the findings of this research, which leads to the conclusions, and a modest attempt to draw some recommendations on the use of research conclusions for policy and practice. The first subsection of this section provides the research context and research questions of the study (7.1.1), the findings of the Baseline Survey (7.1.2) and of the Intervention Study (7.1.3), and their impact on both teachers and students.

7.1.1 Research context and questions

The major education issues and challenges for Mozambique are, among others, the limited access to education, and the inefficiency (high repeat and dropout rates) of the

education system (UNDP, 2000). Concerning access, statistics available at the Ministry of Education and Culture (MinEd, 2003) indicate that at the Lower Primary School, Cycle 1 (EP1), the average enrolment net rate in 2003 was of 69.4%, consisting of 66.4% for girls and 72.4% for boys. As a result, every year, around 30% of children at school age become potentially illiterate, as they miss the chance of entering the school system with severe consequences for their access to the subsequent levels of education system. The devastating civil war, particularly in the 1980s, imposed severe restrictions on the Government efforts of promoting mass education. By 1992, when the war finally came to an end, about 3.530 schools (corresponding to 58% of those that existed by 1983) had been closed or destroyed, affecting more than 1.5 million pupils (MinEd, 1994). Although post-war education statistics (MinEd, 1994, 1999) show that both school enrolments and the number of schools has been increasing since the end of the war, the illiteracy rate remains high. In 2002 the average illiteracy rate was 53.5%, being even higher (68.6%) for females (Sitoe, 2006). Despite the fact that community, private, and church schools offer alternative opportunities to some students (particularly for secondary education), their contribution is still insignificant in regard to the improvement of access to education. The majority of students from the poor social groups, particularly from rural areas, do not have access to secondary education and there are significant geographic (North, Centre, South, and rural vs urban) and gender disparities.

The efficiency of the Mozambique education system and its value, as is the case elsewhere, is judged by the quality of the outputs. Common sense judgments, normally drawn from the performance of the students leaving high school, are that their general knowledge is below average. Although there have not been clear indicators to assess the quality of the outputs of the education system, it is widely acknowledged that 'nowadays students know less and less' (Sitoe, 2006:45). The concern is more acute in Science subjects and at secondary education level where, according to the UNDP Human Development Report, the quality and efficiency of the education system are the lowest (UNDP, 2000). One of the reasons for this situation is that the ESG2 curriculum is highly academic and demands a high level of theoretical knowledge, without promoting practical skills that would facilitate the integration of graduates into the labour market. Life skills do not feature sufficiently in the curriculum, and as a result, there have been high repetition and dropout rates. This weakens the possibility of

developing a critical mass of Mathematics and Science students, a fact which is subsequently reflected in the number of new entries to university. In fact, Entrance Examinations to the UEM have also been used as indicators of the quality of high school leavers. The pass rate in these examinations is below 20%. Mathematics, Portuguese and Sciences (Biology, Chemistry and Physics) are the subjects in which high fail rates are registered.

The Ministry of Education and Culture, in recognition of the need to achieve effective results in educational provision, is undertaking a curriculum review for secondary education and a coordinated curriculum and assessment strategy has been designed (MEC & INDE, 2007). This strategy indicates that there is a need to produce secondary level graduates to ensure an adequate supply of teachers and other public servants, as the future job market is uncertain and subject to rapid technological change.

It is within the context of the poor performance of the students leaving high school - particularly in science subjects - and the need to explore the potential of the new curriculum being proposed by MEC that this study had been undertaken. The central research question of the study was formulated as follows:

What assessment practices do Grade 12 teachers in Physics in Mozambique apply and how can they be improved?

To tackle this research question, the study adopted a twofold research approach. A Baseline Survey aimed at gaining an overall impression of the assessment practices used by secondary school Physics teachers in schools and an Intervention Study aimed at producing improvements on teacher assessment practices. While the Baseline Survey was basically a preliminary research following a survey research method, the Intervention Study is the main study. This followed an educational design research approach whose overall design comprised of two stages namely (i) prototyping and (ii) assessment, systematic reflection and documentation.

Specifically, the aforementioned main research question has been guided by the operational research questions set out below.

- For the Baseline Survey these are:
 - *What assessments practices do Grade 12 Physics teachers apply?*
 - *What is the quality of the assessment practices?*
 - *How relevant can the assessment practices be for students learning?*

- For the Intervention Study this is:
 - *How can the teacher assessment practices be improved?*

7.1.2 A summary of findings emerging from the Baseline Survey

As indicated earlier in this and previous chapters (see also Chapters 3 and 4), a preliminary research aimed at reviewing assessment practices used by secondary school Physics teachers in Mozambique. As a first step towards a good preliminary research to inform the study, a decision was made to conduct the Baseline Survey in carefully selected Mozambican secondary schools from different provinces. In order to address the three operational research questions of the survey, a purposive sample of 12 Physics teachers, four school directors, two pedagogical officers, and three assessment specialists was selected in order to obtain a representative picture of instruction, management and inspectorate perspectives.

The information needed to acquire answers for the operational research questions was obtained via interviews, questionnaires, classroom observations and written notes. A triangulation of both sources and data collection instruments made it possible to validate the information collected. The following are the findings of the Baseline Survey.

- Concerning the types of assessment practices, it is relevant to start by mentioning that the most frequently used assessment practices in schools are paper-and-pencil tests, verbal tests, and homework, while projects, portfolios, and peer-assessments are the less frequently used ones. In general terms, however, it can be concluded that that teachers appeared to be unfamiliar with some assessment practices or they had different understanding of the concepts

used to name some of them (e.g., portfolios and peer-assessment). This was illustrated by lack of consistency in teacher responses from the questionnaire to the interview, and the relatively large amount of missing data.

- As for the quality of assessments, several student activities were used as quality criteria namely oral communication during lessons, written work, presentations, notebooks, laboratory work, and solving problems. The most frequently assessed student activity is written work, followed by ability of students to solve problems. Laboratory work is the activity that was never assessed by many of the researched teachers. Another criterion used to verify the quality of the teacher assessments was the feedback given by teachers to students. The findings indicate that teachers were giving expressed (both favourable and critical), personal (the feedback was given individually to students) and timely feedback (teachers gave their feedback promptly while moving from one student's desk to another).
- Regarding relevance the level of student involvement in the evaluation of their own work was seen to be an indicator of the relevance of the assessment for student learning. The most common way used by teachers to get the students involved in the evaluation of their performance was to reflect with them about their assessment results in order to get them learn from their successes and failures. The use by teachers of assessment results for the evaluation of the student work was another indicator of relevance. The majority of teachers (nine out of twelve) indicated the encouragement of students to engage in active learning as the most used indicator of relevance. A point of concern, however, in relation to the relevance of teacher assessments, is that about one quarter of the teachers (3 out of 12) did not provide answers to the majority of relevance indicators namely (i) to assign grades, (ii) to identify student strengths and weaknesses, and (iii) to help students know and recognise the standards they are aiming for. This suggests that very crucial aspects of classroom assessment such as certifying mastery for assigning grades, diagnosing student strengths and weaknesses, as well as sharing goals with the students have not been taken into consideration by a relatively large proportion of the teachers.

In general, the baseline findings provided a clear direction for the intervention to be undertaken in the following phase of the study. These suggested that, although teachers have mostly used on-paper written assessments, they rarely involve students in performance type assessment. Even those teachers who conduct performance assessment still lack preparation in terms of designing and administering them properly.

During and after the Baseline Survey, an in-depth literature review was undertaken. The review was not only meant to discuss the arguments of several scholars on the topic but also to provide a platform from which to better conceptualise the intervention phase of the study. Five lessons emerged from the review as shown below:

1. Constructivism is one of the most relevant theories in student knowledge construction and it represents a powerful theoretical resource that may maximise student learning.
2. Besides the importance of other classroom assessment practices, performance assessment plays a crucial role in assessing Physics learning. It calls upon the students to demonstrate specific skills and competencies and requires them to perform real-world tasks that demonstrate meaningful application of essential skills and knowledge.
3. Teachers must be supported to contextualise assessment. In this respect, criterion-referenced assessment must be the basis of judging student evidence, the feedback must be judged and used by both students and teachers, and the assessment in general should be directed *for* learning.
4. In undertaking assessment *for* learning, teachers must consider completing the entire cycle of assessment events. This, however, should be done taking into consideration that classroom assessment occurs at the intersection of instruction, classroom management, and assessment, i.e., the broader curricular context has to be dealt with.
5. If assessment is to be effective for learning an entire cycle of goals-evidence-judgment of achievement-next steps in learning-goals has to be completed. In this respect, formative assessment is at the heart of student learning, where feedback for students is only effective when it is used to guide improvement.

Taking into account the findings from the Baseline Survey and considering what the literature outlines as good practice, it was concluded that somehow there was neither sufficient research dealing with the extent to which assessment strategies were used for Physics as a subject, nor a reported professional support for teachers to assist them in the development of performance assessment materials for use in an ordinary classroom environment. This study addresses these shortcomings as follows: firstly, by providing support to teachers in developing and using exemplary support materials on performance assessments; secondly, by ensuring that the development of such materials is done in an ordinary classroom environment to allow teachers and students to participate in the process while working in their own environment; and thirdly, by formatively evaluating all development stages of the intervention, so that the learning evidence is used to feed the teaching and learning process.

7.1.3 A summary of findings emerging from the Intervention Study

This stage of the study was directed towards designing and developing Physics assessment materials aimed at helping teachers to improve their assessment practices in schools. The intervention dealt primarily with the design and formative evaluation of a series of Physics prototypes in the context of demonstration experiments. The validity and practicality of the prototypes were verified using appraisal by *experts, university students, teachers* (for validity aspects), and *teachers and students* in a classroom tryout (for practicality). Secondly, the study addressed the issue of assessment and systematic reflection and documentation of the intervention as a whole where the effectiveness of the material was also subject to appraisal by experts. So, the overall discussion of the main findings from the Intervention Study is presented along the structure of (i) formative evaluation of the prototypes, (ii) systematic reflection and documentation of the materials, and the impact on both (iii) teachers and (iv) students from the perspective of practicality and usefulness of the intervention.

Formative evaluation of the PAM materials

The study selected the topics of force and inertia for the Grade 12 Physics curriculum for exemplifying the demonstration experiments. Exemplary PAM prototypes were developed, appraised and tested in a classroom tryout. The trajectory of the

prototyping process of the PAM materials in force and inertia was presented and discussed in Chapter 4 (Figure 4.4).

The results from the appraisal by experts, teachers and university students provided indications about the validity of the first version of the materials in terms of internal consistency between the materials and the state-of-art knowledge, and the consistency of the various components of the intervention. This stage of appraisal focused *more on improving the validity and less the practicality* of the prototype and it culminated in concrete suggestions for the revision of the first version. The revision suggestions were incorporated in the subsequent version of the prototype (Version 2), which was tried out in the classroom. The trial focused *more on the practicality and less on the effectiveness* i.e. the emphasis was on the usability of the materials by teachers and students in ways that are compatible with the developer's intention.

The results from the classroom tryout showed that the overall impressions of both teachers and students were positive. The two participating teachers indicated that they liked the presentation and structure of the materials following the POE strategy, and regarded the material as being very useful in helping to enhance their professional experience and to prepare their own assessments. As for the conduct of experiments, they referred to the aspect of their commitment as crucial for achieving the desired experiment results. Preparation of experiments takes time and effort from teachers in such overcrowded classrooms. These are not always available within the curriculum schedule. Effectively, both teachers agreed on the need to improve managerial strategies related to time and class size. Students enjoyed the POE strategy (the sequence of the three steps) because it allowed them to realise their misconceptions during the experiments and to develop their own explanations of the observed differences between prediction and observation of the events. Doing experiments in groups - particularly during the observation phase - was the aspect that they did not like. They argued that working individually permitted better visualisation of the events and it helped in developing their own ideas, which could later be shared in groups during the reconciliation phase of the POE strategy. When asked about the differences between demonstration experiments on PAM materials and their regular Physics laboratory lessons, they highlighted the important role played by the teacher during the

tryout. They said that the teacher acted more as a guide to students than as a transmitter of knowledge.

Overall, the revision decisions from the classroom tryout were used during the second phase of the intervention - systematic reflection and documentation – to improve the effectiveness of the material and the sustainability of the intervention as a whole.

Systematic reflection and documentation

The first stage of the systematic reflection and documentation consisted of final appraisal by two experts who assessed the material in terms of their *effectiveness* and *sustainability* of the study. The outcome of this reflection led to the production of the last version of the prototype (Version 4). During the second stage, an evaluation workshop with teachers and university students was conducted. Although all participants attested that they were convinced about the potential of the materials to improve classroom assessments, the teachers appeared to be concerned about the process of implementation of the intervention. Effective management strategies to deal with overcrowded classrooms and support from school leadership were mentioned as some of enabling factors which could contribute to the success of the innovation. The following subsections present reflections on the impact of the intervention on teachers and students. The reflections are presented in connection with findings from the classroom tryout and with teachers' suggestions from the evaluation workshop.

Impact of the intervention on teachers

At the outset of this study (Chapter 1, Section 1.1), it was stated that Mozambican secondary school teachers strive to meet the same learning outcomes for all students. An assumption underlying this research was that good assessment practices used by teachers in their classrooms could well serve as one of the vehicles to realise these learning outcomes. It was assumed that if all teachers, regardless of their different levels of qualifications, are made familiar with the characteristics and usefulness of formative assessment, they can generally improve their assessment practices and particularly help students learn Physics better. Two reasons support this assumption: firstly, the Mozambican Government, through the MEC, decided to consider the revision of assessment practices as one of the strategic priorities to be taken into account during the curriculum review of the secondary education curriculum (MEC &

INDE, 2007). Secondly, arguments from various authors about classroom assessment (Chapter 3, Section 3.6) indicate that teachers need support in developing and using exemplary support materials on performance assessments, and that, more importantly, these materials should be evaluated formatively to allow the learning evidence to be used to feed the teaching and learning process.

The central research question of this study dealt with how to improve assessment practices used by Grade 12 Physics teachers in schools. After investigating the assessment practices currently used in the classroom, it was decided to undertake an intervention aimed at improving such practices. However, measuring the impact of the intervention on teachers' familiarity with formative assessment materials, and on the use of the learning evidence to feed the teaching process, has been daunting. This is because the overall success of the intervention is not limited to the teachers' ability to conduct demonstration experiments but also depends on their subject knowledge – worrisome teachers' under- or non- qualifications levels are reported – on methodological aspects within the new curriculum, and on the supervision process undertaken by the MEC on school follow-ups. This is in line with Joyce and Showers (1988, 1995) who indicate that a systematic combination of five training components (theory, demonstration, practice, feedback, and coaching) leads to effective teacher learning about an intended change and its transfer into classroom practice. For this reason, a decision was made to explore the indications of teacher learning by triangulating self-reported statements during the evaluation of the intervention and classroom observations by the researcher.

The participating teachers reported that, after the demonstration experiments, their awareness and skills in designing and conducting experiments were enhanced. Effectively, there are indications that demonstrate teacher learning from these experiments. Classroom observations conducted after the intervention, and results from the evaluation workshop, indicated that teacher lesson preparation (e.g., lesson plans and organisation of the materials and equipment) had followed the specific suggestions from the exemplary PAM materials. However, as it was emphasised by the two teachers who participated in the tryouts, an improvement of teacher assessment practices will have to depend (i) on their personal commitment and (ii) on the flexibility of the curriculum regarding time allocation because the POE strategy

requires a step-by-step guiding of students which is not easy in the context of overcrowded classrooms. Furthermore, based on classroom observations, it was evident that in a school where the teacher was more active and committed, it was possible to organise students to conduct the experiments with relatively less time than in the other, even with the involvement of far more students. This shows that the findings between the teachers' perceptions of the demonstration experiments and lessons learnt as measured by the workshop results and classroom observations, are concurrent.

The impact of the findings on teacher learning concurs with Garet et al., (2001). These authors argue that professional development interventions aimed at improving teacher content knowledge, pedagogical content knowledge, and active learning opportunities, when integrated into school daily life are related to teachers' perceptions of enhanced knowledge and skills. However, as was suggested earlier in Chapter 6 (Section 6.6), the overall success of the intervention will ultimately depend, amongst other factors, on the support teachers receive from various educational stakeholders. This is illustrated by the teachers' answers during the evaluation workshop of whether the respective schools have provided support to the implementation of the Physics demonstration experiments (refer to Level 3 of Guskey's model in Table 6.6). Although they said it was too early to make an evaluation, their general impression was characterised by certain scepticism about the potential success. They argued that they were sceptical because in similar initiatives, schools were not very supportive, particularly regarding the financial costs. They are of the opinion that, unless changes are embedded in teacher training programmes and gain support from the MEC right from the initial phase of the teacher training programmes, it will be difficult to predict positive changes at the school level alone.

Impact of the intervention on students

One of the criteria for measuring the impact of demonstration experiments on student learning of Physics is the learning outcomes. It was assumed that changes in participant teacher skills and knowledge in conducting Physics demonstration experiments and formative evaluation of their students would, therefore, improve student learning. The assumption was based on the fact that an improvement in teacher assessment practices would ultimately lead to better results. However, as was the case

with teachers, connecting the dots between intervention findings and student learning has been a difficult task for this study. Firstly, because, as referred by Sykes (1999), demonstration experiments, like the ones reported in this study, alone cannot produce student learning outcomes unless the educational system in which they are embedded is supportive enough to produce positive changes (e.g., adequate teacher training system, conducive classroom environment, government willingness to support changes). Secondly, learning outcomes are also influenced by student activities outside the school environment – peers, parents and the media (Guskey, 2000; Ogunniyi, 1986). The focus of the study did not allow any investigation into the influence of these external factors. One way to address this complexity of measuring the impact of demonstration experiments on student learning would be to evaluate the students through an achievement test or by an interview. Because of time constraints, the researcher was not able to do such an evaluation.

Another criterion to measure the impact of demonstration experiments on students is their attitude towards the experiments. The students' general attitude was positive. In the evaluation questionnaires they reported that they enjoyed their participation in the experiments particularly the opportunity to develop their own explanations of the events. They felt that the three steps of the POE strategy increased their motivation to study Physics. It allowed them to start any experiment by voicing their own predictions and observing the actual course of the event before drawing any conclusion about it. Comparing their own thinking with that of the others, evaluating their own mental ability, and drawing informed conclusions were some of the aspects referred to by students as the most valuable insights gained from the demonstration experiments.

The last criterion used to evaluate the impact of the demonstration experiments on students was the demonstration experiment report, which was prepared and submitted by the students after the experiments. This report helped to give an idea of the quality of student learning in terms of research skills and reasoning abilities. By writing the report, the students were able to express their level of understanding of force and inertia, as well as their ability to design, conduct, and communicate the results of the experiments.

7.2 Reflections of the study

Section 7.1 provided a summary of the context in which the research was carried out and the research questions that guided the investigation. The section has also reflected on the main findings of the two components and provided an in-depth discussion of the impact of the demonstration experiments on both teachers and students. The present Section (7.2) discusses the main reflections about the study in relation to three perspectives namely methodological, substantive, and scientific.

7.2.1 Methodological reflection

As referred to in Chapters 1 and 4, the study concentrated on investigating and improving assessment practices used by secondary school teachers in Mozambique situated within the pragmatic knowledge claim, making use of a mixed mode methodological approach. It employed a survey approach to investigate the assessment practices currently used by Grade 12 Physics teachers in schools prior to an educational design research approach used in the Intervention Study and aimed at improving such practices. As argued in Chapters 1 (Section 1.2) and 2 (Section 2.3), the importance of the Baseline Survey for the intervention lies on the prior understanding of the assessment practices taking place in schools and classrooms before designing an effective intervention aimed at improving these practices. This understanding implied looking at the school and assessment practices into three perspectives: firstly, the characteristics of the curriculum in place, the situation of the teachers working in school, the existing infrastructures and the whether the school culture is conducive to the design and implementation of the intervention (input); secondly, the understanding of the existing classroom practices in terms of how the instruction is related to assessment (process) and thirdly, the analysis of student achievement and of their attitudes towards the learning of Physics (output). All this baseline information on the conditions in place was important for designing and implementing an effective intervention for improved classroom practices.

The survey approach permitted the collection of data in cross-sectional studies using a variety of data collection instruments such as questionnaires or semi-structured interviews, with the intention of obtaining a maximum variation sample. Both

qualitative and quantitative descriptive methods were used to analyse the data so that an overall picture of the survey could be captured. The triangulation of sources and data collection instruments also made it possible to explore the views of teachers, school directors and educational officers about the types, quality, and relevance of assessment practices used by teachers in schools. Due to the potential of this approach, it was possible to obtain a good picture of the types assessment practices used in such a huge and wider population across the country as well as considering assessment from various perspectives (instruction, management and inspectorate).

This mixed method approach, however, also involved some constraints during the course of the research. During the Baseline Survey the problem was linked to the discrepancies between answers given by teachers to questions in the questionnaires and in the interviews, possibly due to the fact that data collection instruments were piloted with teachers in Maputo, whilst all schools in the sample were from outside Maputo. The decision to pilot in Maputo was deliberate because of the costs and time constraints involved in doing a countrywide pilot study, although it meant that the possibility for diverse responses increased (refer to considerations about assessment practices used by teachers under 5.2.1). A greater unfamiliarity with the concepts seems to increase as one move away from Maputo area due to the lack of libraries and information dissemination generally, and poor teaching conditions that the teachers face countrywide. As a result, some teacher limitations in dealing with data collection instruments went unnoticed during the pilot phase. The lesson that can be learnt from this is that, irrespective of any logistical problems, additional efforts must be made to minimise problems resulting from differences between subject characteristics of the pilot process and those of the actual research.

In the Intervention Study, the applied educational design research approach provided flexibility in developing an intervention step-by-step within the problem. Van den Akker (1999) argues that the approach enables the realisation of small-scale examples of interventions, and generates methodological guidelines for the design and evaluation of such interventions. Therefore, the educational design research approach was considered appropriate and useful for the Mozambican context because of the opportunity given to teachers to design, develop and evaluate the intervention with a local relevance, while working in their own environment. Teacher participation in the

process advocated by the approach, was influential in understanding their potential difficulties in the implementation process as well as the local conditions, which were crucial to the future improvement of the intervention. However, one of the methodological problems with the educational design research approach, which is linked to its formative research character of working in the natural setting, is the generalisability of findings (Walker, 1992; Yin, 1994). The concern is the extent to which the research findings are transferable from the situation being studied to situations not being studied. In fact, this was the case with this study because, during the prototyping process, results from the formative evaluation activities were incorporated into the design process of the subsequent prototypes. This makes the generalisation of the findings of this study to a wider population a daunting exercise, due to the lack of statistical representation – a very small number of teachers took part in the formative evaluation of the prototypes – and to the lack of replications of the evaluation findings into different users.

In light of these considerations, the readers of this dissertation are encouraged to consider some of the analytical forms of generalisation suggested by Miles and Huberman (1994), Teclé (2006), and Yin (1994). These are as follows:

- an in-depth description of the research context, research design, and formative results;
- the rationale for choosing the quality criteria and research design;
- the researcher's role in the research activities;
- a detailed description of the data collection instruments, their administration procedures, the number and characteristics of respondents, and the methods of data analysis; and
- the replications of intervention findings in more tryouts to determine whether the same results may occur.

In summary, the combination of survey and educational design research approaches added a great value to this study in the sense that not only reflected the perceptions of a varied sample of Physics teachers countrywide but also allowed them to participate in the research while working in their own environment.

7.2.2 Substantive reflection

At the outset of this chapter (subsection 7.1.1) it was mentioned that one of the main problems that the Mozambican education system has been facing is linked to the limited access to education for the majority of children of school going age. In more recent years, however, the government of Mozambique has shifted the attention to the issue of efficiency of the school system. This is in recognition of the fact that, although there are still children who are out of the school system, those who are in it are not yet benefiting from the desired quality education. The reason for this is attributed to the quality of outputs at the end of different learning cycles. The steady increase of unqualified teachers in schools exacerbates this problem. This implies that there is a need for teachers (both qualified and unqualified) that should be seriously addressed through a sound teacher-training programme. Several reports (INDE, 2005; Lauchande, 2001; MinEd, 1998; Palme, 1992; Popov, 1994) indicate that secondary school teachers show a lack of skills in designing and administering valid formative tests and in addition, they have also had to face problems in formulating test items requiring the analysis and comprehension levels of cognition. The Government of Mozambique, through the MEC, expects that in the new secondary education curriculum, teachers should become familiar with the characteristics and usefulness of formative assessment and learn to develop formative assessments to be used in the classroom to inform and enhance the learning process (MEC & INDE, 2007).

This study is intended to be a modest contribution to the seeking of a solution to the problem of system inefficiency by investigating the quality of school leavers at Grade 12. Among the few studies that have been conducted about the quality of system outputs, some address issues related to teacher training and curriculum implementation (Afonso, 2007; Cupane, 2007; Fagilde, 2002; Huillet, 2002; Kouwenhoven, 2003), others look at students' alternative conceptions and beliefs (Mutimucuo, 1998; Siteo, 2006) and some others (Januário, 1997; Lauchande, 2001; Palme, 1992; Popov, 1994) address the issue of school effectiveness, with particular emphasis on the assessment of student learning. Most of these studies focus on primary education. More recently, some studies were conducted with the focus on secondary education in assessment for Mathematics (Machado, 2007), and in practical work for Biology (Cossa, 2007) and Chemistry (Chevane, 2002). Not one of these studies, however, has investigated the

way students are assessed in authentic research situations within the context of demonstration experiments and involving teachers in the design and evaluation of assessment materials. A very distinctive element of this study is its potential to generate improvement suggestions into the assessment system, drawn from the work with teachers who participated as designers and users.

7.2.3 Scientific reflection

As referred to earlier in Chapter 4 (Section 4.3), this study used a combination of qualitative and quantitative data collection methods for collecting information (Baseline Survey) and for the summative evaluation of the prototypes (Intervention Study). The combination of Baseline Survey and Intervention Study has conferred on the study a particular uniqueness in the Mozambican research context. As indicated in Chapter 3, studies conducted in and about the country in the area of assessment for secondary education are still scarce, and none of them have employed an educational design research approach. The formative evaluation activities incorporated into a cyclic process of design and evaluation of the PAM prototypes of this study, took place in naturally existing schools. Authentic research situations were used to test the expected practicality and expected effectiveness of the material. This iterative process of design and formative evaluation between users (teachers and students) and the researcher, which occurred in the ordinary classroom environment, has enabled the researcher to observe the main characteristics of the prototypes in action and to suggest timely improvements. The findings of the prototyping process in particular, and those of the demonstration experiments in general, have shed light on several issues related to classroom assessment, namely: (i) the time needed within the curriculum to accomplish the tasks; (ii) the classroom managerial aspects like grouping and assessment of students; (iii) the practicality of the POE strategy; and (iv) the role of teachers and students during the experiments. Drawing from experiences of previous studies carried out in different areas and contexts, the author of this study claims that it is acceptable to assume that, by supporting teachers in designing and trying out assessment prototypes in one type (performance assessment) and in a given context (laboratory setting), it will be possible to see teachers transfer their knowledge and skills to other assessment practices and contexts.

Furthermore, the collaborative element has also been beneficial in this study. The involvement of teachers, school directors and policy-makers (pedagogical officers and assessment specialists) in the various stages of the study had the advantage of accommodating the teachers' needs in the assessment materials and of raising the decision makers' awareness on the importance of incorporating these needs into curriculum review process and teacher training programmes.

Finally, it would be misleading to suggest that the role of the researcher during the study was consistent and went unnoticed, particularly in the intervention phase. When the researcher started his research, he assumed the role of the designer who developed the data collection instruments and the first version of the PAM prototypes. During the formative evaluation of this and the subsequent versions, he acted as co-developer with teachers and expert appraisers, a position that led to designer-researcher. In the classroom tryout of the material, this role diminished considerably giving way to a participant-observer playing more the role of task facilitator. As the designer, the researcher aimed at developing high quality materials, as facilitator at ensuring that teachers are adequately introduced to the intervention, and as researcher at being objective. These multiple roles not only brought benefits but also some practical problems. An example of a problematic situation was when students asked for clarification of how to conduct a certain experiment properly. It was compromising for the researcher to help students while maintaining his role as researcher. Being constantly aware of the potential bias of these multiple roles, the researcher had to rely on the informed opinions of critical outsiders amongst the experts, supervisors and colleagues within the Faculty of Education at UEM. In fact, in order to reduce these dilemmas, van den Akker (2002) suggests the use of multiple methods and sources of data collection, and discussions with many parties involved in the development process. This means that the perceptions and interpretations of the events by the researcher were shaped by the triangulation of multiple data collection methods (evaluation questionnaires, interviews, classroom observations) and various sources of information (teachers, students, school directors, educational officers, university students) representing multiple perspectives, which prevented bias in the interpretation of data and the analysis of findings from the development process (Krathwohl, 1998).

Overall, it can be concluded that the combination of the survey and the educational design research approaches added value to the research community in Mozambique. An informed intervention aimed at improving assessment practices involving potential users in real situations was undertaken on the basis of a needs assessment. This was crucial for the improvement of the ecological validity of the study findings (Cohen et al., 2000).

Section 7.3 presents the conclusions of the study resulting from the reflection about the findings.

7.3 Conclusions of the study

The aim of this study was to investigate and improve assessment practices in Physics used by secondary school teachers in Mozambique. Specifically, the study intended to investigate how assessment practices of teachers can be improved, with a focus on Grade 12 Physics. Methodologically, to address this aim a twofold approach was employed. An Intervention Study consisting of development and tryout of consecutive prototypes of assessment practices for Physics teachers undertaken in the context of demonstration experiments. But, prior to the intervention, a survey intended to achieve a good understanding of what was actually going on in the classroom in terms of types, quality and relevance of assessment practices used by teachers, was conducted. Based on findings obtained from various data collection instruments and drawing from reflections of the literature review, the conclusions listed below are presented and discussed.

1. Basic assessment practices undertaken by Physics teachers in Mozambican secondary schools appeared to be of poor quality and there is a need for improvement.

During the Baseline Survey, a large number of teachers could not provide answers to some of the assessment practices due to their apparent poor understanding or lack of knowledge. They showed lack of understanding between what they do in the classroom and what they were actually being asked by the researcher. One of the evidences of the teachers' lack of understanding of most of assessment practices are expressed by the teachers' claim that students can not assess their work or that of their peers (peer-assessment) because they are in school to study and could give each other

low marks (refer to Chapter 5, subsection 5.2.1). More evidence of teacher lack of understanding was shown by the inconsistencies in teacher responses from the questionnaires to the interviews (see Chapter 5, Table 5.1). Very often, teachers contradicted themselves from one data collection instrument to another and when questioned, it appeared that the lack of knowledge and understanding was the cause for this. The amount of missing data in the questionnaire also reflected teachers' unfamiliarity with basic assessment practices.

An analysis of the way teachers assessed some of the student activities also expressed how low was the validity, in terms of content, of the teacher assessment (refer to Chapter 5, subsection 5.2.2). For instance, at the question of *how often teachers assess student activity of solving problems*, some of the teachers interpreted the meaning of 'solving problems' as finding answers of some questions using calculations while, for the context of this study, it referred to generate solutions of real-world problems. Furthermore, although there were positive aspects of feedback provision (articulation, timeliness, and personalisation), generally spoken, there was a lack of input given to students in order to empower their learning.

During the intervention study, teachers also seemed to have difficulty in both reporting experimental results and evaluating student performance in demonstration experiments (refer to Chapter 6, subsection 6.3.4). Although the experiment report template contained guidelines on what to include in each section, teachers focused their attention more on student manipulative skills than other investigative skills. Student ability to communicate the results of their experiments was also not dealt with by the teachers because teachers, like their students, lacked investigative skills. It can then be concluded that teachers in schools are conducting most of the basic assessment practices with limited knowledge and skills about assessing effective student learning.

2. Developing and applying exemplary assessment materials has the potential to improve performance assessment practices associated with demonstration experiments in Physics.

Despite all the problems with teacher assessment in schools, the findings of this study have shown that it is possible to develop a potential solution for the assessment of student learning, particularly for Physics. Training teachers in designing exemplary lesson materials and supporting them with the materials, which includes an assessment

component, is essential for the improvement of both teaching and learning. Physics is, by nature, an experimental subject in which effective learning can only take place if students are required to perform real-world tasks that demonstrate meaningful application of essential skills and knowledge. According to research (Airasian, 2000; Moskal, 2003; Popham, 2002), one of the most successful assessment practices in Science education is performance assessment, because of its crucial role in assessing performance tasks. Still according to research, an effective performance assessment for experimental subject like Physics is most likely to succeed when it is undertaken in a laboratory context, where students can perform real demonstration experiments. During the classroom tryout, teachers and students were positive about demonstration experiments (refer to Chapter 6, subsection 6.3.4). It emerged that experiments allow students to develop their own explanations of the events they observe and to draw informed conclusions. The Predict-Observe-Explain (POE) strategy, associated to the experiments, not only enhance student learning of Physics but also add an element of motivation and enjoyment because *“starting any experiment by making your own prediction of the event is so fascinating that you never forget the results of the comparison (...)”*. However, two elements appear to be relevant in addressing the improvement of demonstration experiments, namely: time spent during preparation and execution of experiments and class sizes. Experiments were described as being time consuming and they can be difficult to conduct in large classes if adequate class management strategies are not adopted.

In conclusion, these study's findings indicate that performance assessment, when conducted in the context of demonstration experiments, represents a potential solution to the poor quality of teacher assessment practices in schools. It appeared to be one of the most successful means of assessing students learning of Physics and one may expect that the same approach will work for other subjects and other assessment strategies.

3. Exemplary assessment materials containing specific guidelines appear to be essential to support teachers' effective practice.

One of the most challenging problems faced by teachers in Mozambique is the availability of support materials for teaching and the limited access to libraries, especially in schools located outside Maputo. Good curriculum materials and effective

teacher training programmes may fail to produce impact in teaching and learning in schools due to the lack of supplementary lesson materials. Intervention studies in Science education conducted internationally (see Chapter 3, Section 3.5) have shown the importance of developing exemplary support lesson materials for teachers. These materials not only support teachers in aspects like subject knowledge, lesson preparation, teaching methodology, assessment and feedback, but also help students to construct their own knowledge. Characteristics of such materials are that they should be (i) based on the objectives of curriculum, (ii) developed from materials teachers are already using, (iii) made to engage students, support curriculum implementation, improve student teaming, and to report individual student progress, and (iv) made to help teachers adopt a student-centred approach.

Despite the importance of the specifications in the materials, it is important, however, to indicate that empirical evidence has shown that teachers feel uncomfortable working with lengthy materials and, therefore, they should be designed to be user-friendly (refer to Chapter 6, subsection 6.3.4). In this study, assessment materials not only contained the characteristics described above, but also had specific guidelines on how teachers can design and conduct performance assessment in the context of demonstrations experiments following the POE strategy. Specifically, the materials provided teachers with specifications on components and functions of assessment including a practice-oriented lesson plan (Appendix P, Part 2), on designing assessment practices and providing feedback (Part 3), and on conducting demonstration experiments and assessing student performance (Part 4). The participating teachers reported that, after the demonstration experiments their awareness and skills in designing and conducting experiments were enhanced. In fact, there are indications that demonstrate teacher learning from these experiments. Classroom observations conducted with teachers after the intervention, and results from the evaluation workshop, indicated that their lesson preparations (e.g., lesson plans and organisation of the materials and equipment) following specific suggestions from the exemplary PAM materials had improved significantly (refer to Chapter 6, Section 6.5). Study findings also revealed that the involvement of teachers in developing assessment materials appears to improve their confidence and their ability to think critically. Therefore, exemplary materials with detailed specifications on how

to design and use different parts of the lessons – including assessment - appear to be one possible solution to address the problem of poor teacher assessment practices.

4. The study findings from Mozambique confirm what the international literature has indicated in relation to effective classroom assessment practices.

Findings from this study confirm what international literature says and what previous studies have found in relation to classroom assessment practices in more developed contexts. An example of this is found in the role of both teachers and students in formative assessment. There is evidence from this study that formative assessment is a process requiring a close interaction between the teacher and the students (refer to Chapter 6.3.4). During the demonstration experiments, when students were asked to indicate the difference between the demonstration experiments based on PAM materials and their regular Physics laboratory lessons, they referred to the fact that the lessons in the tryout, were closely facilitated by the teacher and were accompanied by worksheets with detailed instructions on how to do things. In this regard, one student said: *“In our regular laboratory lessons the teacher would simply tell us what needs to be done and wait to see whether or not we managed to reach the desired outcome (...)”*. For the demonstration experiments to be successful, the teacher had to guide students on the POE strategy, i.e., on how to predict the behaviour of the events, how and what to actually observe during the experiments, and how to reconcile the predictions and the observations. In so doing, students had to use the information they acquired during learning to draw conclusions about their observations. This is line with what Black et al., (2003) report about formative assessment – a process in which information about learning is evoked and then used to improve the teaching and learning activities in which teachers and students are engaged.

Another example is related to the time involved in developing and trying out exemplary materials. This study showed that teachers needed more time to conduct effective demonstration experiments and to provide formative feedback to all students given the context of large classes (refer to Chapter 6, Table 6.6). The time was revealed as problematic in other intervention studies conducted elsewhere. Motswiri (2004) calls it lack of congruence of the exemplary materials where the intended practice appeared to be incongruent with the teachers’ current practice. Teclé (2006) refers that, in her study, teachers were observed encountering problems with group

work activities throughout the tryouts and the issue of time continued to be problematic. Ottevanger (2001) argued that, although teachers seemed to address the time issue in their own ways, this appeared to be a continuous problem in completing lessons.

5. Research only conducted in Maputo cannot be generalised to the rest of the country.

Research findings of this study cannot be generalised to a wider perspective, i.e., nationwide. Two aspects are the reasons for this. Firstly, although the sample of the Baseline Survey was drawn to meet a maximum variation in its representation across the country, library facilities and other teaching conditions in schools countrywide tend to decrease and worsen the further when one moves away from Maputo. This situation had implications on the way teachers perceived the various data collection instruments and interpreted the different assessment related concepts and practices within them. As a result, some teachers' responses lacked consistency as referred to in Chapter 5 (Section 5.2). Secondly, only the second version of the PAM materials designed in this study was tried out with potential users in the classroom. Whether the characteristics of the final version (Version 4) of the material will produce effective improvements on the way teachers conduct performance assessment is a point to be proven. As Yin (1994) points out, the generalisability, beyond the investigated teachers and students, of the study findings similar to the one reported in this dissertation, can only be made if several replications of intervention findings are undertaken in more tryouts and with more users (refer to Chapter 4, Section 4.4).

7.4 Recommendations

This study has shown the way teacher assessment practices in the context of Physics demonstration experiments can be improved, employing an educational design research approach. The study findings have partially confirmed results from earlier studies about the nature of design and formative evaluation of assessment materials and the potential difficulties inherent in teacher participation in formative evaluations. This final section ends by providing some recommendations formulated from three perspectives namely (i) policy and practice, (ii) further research, and (iii) further

development work. Some of these recommendations are based on the findings drawn from the study while others result from the researcher's professional experience.

7.4.1 For policy and practice

- A relevant conclusion of this study is that Physics teachers in secondary schools have limited knowledge and skills to conduct effective classroom assessments. The implication of this fact is that these teachers need support in designing and using appropriate and relevant assessment practices as well as in identifying effective assessment approaches. It is then worthwhile to recommend that the MEC and teacher training institutions promote the up scaling of Physics teacher knowledge and skills in this respect within an in-service mode to allow teachers to benefit from training while they are working.
- Promoting an effective in-service training for teachers in schools implies having exemplary support assessment materials which can help teachers not only to prepare and conduct lessons but, more importantly, to monitor student learning. But these materials, as argued by the literature (Mafumiko, 2006; Motswiri, 2004; Tecele, 2006), and supported by study conclusions, have the potential to improve teaching and learning if they are designed by the teachers themselves (refer to conclusion nr. 2). In this context, it appears to be relevant that teachers are trained on developing exemplary assessment materials for their own use in schools. This can be done in the form of regular workshops with teachers of the same school or of different schools during selected days of school vacations or over weekends. It is, however, important to note that exemplary PAM materials alone are not sufficient to support teachers in conducting demonstration experiments and, hence, improve their assessment practices. Incorporating the materials into teacher training programmes is desirable. For instance, in-service education can provide teachers with the opportunity to interact with the material, to clarify related theories, to practise and demonstrate the intended innovation. Short-term in-service courses for selected teachers and assignments for them to read and discuss can be used to promote this purpose.

- As already discussed in Chapter 6 (Table 6.6) and acknowledged in the conclusions of the study (refer to conclusion nr. 4), the two participating teachers in the classroom tryouts of the PAM prototypes expressed concerns related to the time needed to conduct the experiments following the POE strategy. The high number of students per class, which makes the supervision of all students a daunting exercise, worsens the problem. It then seems to be relevant for the Ministry of Education and Culture to reconsider the time allocated for Physics lessons within the Grade 11 and 12 Syllabus especially for the so-called 'practical lessons'. The on-going process of curriculum review is an excellent opportunity to embed the change.
- Another concern mentioned by teachers, particularly during the evaluation workshop, is the limited capacity of teachers and schools when it comes to the reproduction of student worksheets and teachers' guides. If the Ministry of Education and Culture is to accommodate the implementation of the PAM materials, school budgets and the existing system of cost sharing with Parents' Councils need to be addressed and revised. This means that the implementation in schools should not be left to the teachers alone.

7.4.2 For further research

- The main purpose of this study, as stated in Chapter 1 (Section 1.1), was to investigate and improve assessment practices used by secondary school Physics teachers in Mozambique. The teaching and learning process, however, is inherently an iterative endeavour, involving two main agents, namely teachers and students. Thus, from this perspective there is a need for other studies to be undertaken so that these assessment practices can also be investigated and improved from the students' perspective. The constructivist approach advocated by this study rests on cognitive principles according to which knowledge is not received passively but actively built up by a cognising subject. This means that student epistemologies exert influence in the way they are likely to perceive and approach knowledge (Sitoe, 2006).

- This study involved one tryout of the experimental materials (Version 2 of the prototypes) with two teachers and their students. From empirical point of view this activity cannot be sufficient for concluding that the PAM materials are practical and effective. As has been emphasised in previous chapters, the emphasis of the intervention was put on investigating the expected practicality and expected effectiveness of the material in terms of their potential to improve teacher assessment practices through experts and not via empirical testing. Further research is needed to prove the actual practicality and effectiveness of the materials.
- Any process of implementation of innovation takes time, and research is needed to explore changes in and of the process. According to a model by Hall and Hord (2001) there are some elements which play a role in the change process namely, the change facilitators who provide assistance, the individuals who implement the change and the resource systems from which support is drawn. Facilitators can probe the change and the results can be used to match resources with the users' needs. As discussed in Chapter 6, it appears that there are short-term impacts of the PAM materials on both teachers and students. It may be interesting to investigate the long-term impacts of the intervention particularly related to the use of the POE strategy in assessing student learning, and on teacher familiarity with formative assessment.
- Guskey (2000) explains that educational innovations sometimes have a counter-productive effect on those who implement them, particularly when they oppose the existing policies. This could be the case with this study from the perspective of the teachers if, for instance, some teachers recruited by the Ministry are not familiar with formative assessment materials and with the use of the learning evidence to feed the teaching process. Teachers attempting to implement the innovation may discover that certain colleagues contradict their efforts. Therefore, conflicting or supportive environmental factors are important aspects to be researched further.

- The level of access to teaching and learning materials including library facilities has proven to be difficult for teachers in Maputo and those of other parts of the country. Teachers working outside Maputo are more in short supply of support materials for their work than their colleagues from Maputo. This was reflected by the difficulties that teachers from the provinces other than Maputo had on understanding and interpreting data collection instruments and assessment related concepts. These difficulties did not emerge during the pilot process of the instruments with Maputo teachers. This fact implies perhaps that, for future studies, despite all logistical and financial constraints involved in conducting research countrywide, data collection instruments need to be piloted in other parts of the country to allow that all potential teacher difficulties are totally uncovered and timely addressed.

7.4.3 For further development work

- When designing and developing demonstration experiments, special attention needs to be paid to the *time* required to address student questions and difficulties in order to reflect on unforeseen issues, and on *class management*, given the fact that the majority of schools are characterised by overcrowded classrooms.
- The effect of innovations and changes on student learning takes time because learning outcomes are also influenced by some other factors taking place outside school environment. Demonstration experiments, like the ones reported in this study, alone cannot produce improvements in teacher assessment practices and consequently in student learning outcomes, unless the educational system is supportive enough to accommodate, for instance, the findings in teacher training programmes. Furthermore, parents, peers and the media influence the way students are actually learning. This means that in future an evaluation study using, for instance, achievement tests could help determining the extent to which demonstration experiments could ultimately have impacted on student learning.

- The POE strategy suggested by this study requires collective and individual observation, reflection, and the reconciliation of ideas. A complete and sound sequence of these events is only successful if all students are engaged. So, advice on how to deal with large classes is a challenge that needs to be carefully addressed in future studies.

Although the findings of this study indicate that the most frequently used assessment practices in Mozambican schools are paper-and-pencil tests, verbal tests, and homework, while projects, portfolios, and peer-assessments are the less used ones, a critical need for the improvement of teacher skills in designing and utilising assessment practices specially those required for science subjects such as Physics has been highlighted. Curriculum review by the MEC should further emphasise the need for assessment to become an integral part of teaching and learning, as a planned process of identifying, gathering and interpreting information about the performance of students, which would have a positive effect on the education system.

Effective teaching and learning can only take place if assessment practices are being implemented effectively as part of this teaching process. This is very true in Mozambique where teachers have many challenges to deal with both within the community and the classroom. The role of Physics education and more broadly Science education is critical to the development and economy of Mozambique and therefore it is of utmost importance to obtain effective teaching (including assessment) and learning practices across the entire country in this domain. This can only happen if the researchers, policymakers and practitioners combine and share their expertise.

REFERENCES

- Afonso, E. (2007). Developing culturally inclusive philosophy of science teacher education. In I.V. Mutimucuo & M. Cherinda (Eds.), *Proceedings of the 15th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 307-315).
- Airasian, P. (2000). *Assessment in the classroom: A concise approach* (2nd ed.). Boston: McGraw-Hill.
- Airasian, P.W. (2001). *Classroom assessment: concepts and application* (4th ed.). Boston: McGraw-Hill.
- ARG (1999). *Assessment for learning: Beyond the black box*. University of Cambridge: Assessment Reform Group.
- Bardin, L. (1977). *Análise do conteúdo*. Lisboa: Edições 70.
- Bell, B., & Cowie, B. (2001). *Formative assessment and science education*. London: Kluwer Academic Publishers.
- Bereiter, C. (2002). Design research for sustained innovation. *Cognitive Studies Bulletin of the Japanese Cognitive Science Society*, 9(3), 321-327.
- Black, P. (1998). *Friend or foe? Theory and practice of assessment and testing*. London, Philadelphia: Falmer Press.
- Black, P., & Atkin J.M. (Ed) (1996). *Changing the subject. Innovations in science, mathematics and technology education*. London and New York: Routledge.
- Black, P.; Harrison, C.; Lee, C.; Marshall, B.; Wiliam, D. (2003). *Assessment for learning. Putting it into practice*. London: Open University Press.

- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessing in Education: Principles, policy and practice*, 5(1), 7-74.
- Black, P., & William, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9-25). London, Thousand Oak, New Delhi: Sage Publications.
- Brown, G., Bull, J., & Pendlebury, M. (1997). *Assessing student learning in higher education*. London and New York: Routledge.
- Buendía Gomez, M. (1999). *Educação Moçambicana: História de um processo 1962-1984*. Maputo: Imprensa Universitária.
- Champagne, A.B., Klopfer, L.E., and Anderson, J.H. (1980). Factors influencing the learning of classical mechanics. *American Journal of Physics*, 48(12), 1074-1079.
- Chatterji, M. (2003). *Designing and using tools for educational assessment*. Boston: Allyn and Bacon.
- Chevane, V. N.V. (2002). The impact of two different ways of assessment of the chemistry laboratory classes in the basic science course at the Eduardo Mondlane University. In C. Malcolm, & C. Lubisi (Eds.), *Proceedings of the 10th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 9-14).
- Clement, J. (1982). Students' perceptions in introductory mechanics. *American Journal of Physics*, 50(1), 66-71.
- Clement, J. (1993). Using bridging analogies and anchoring intuitions to deal with students' preconceptions in Physics. *Journal of Research in Science Teaching* Vol 30, nr 10, 1241-1257.

- Cohen, L., Manion, L. & Morrison, K. (2000). *Research methods in education* (5th ed.). New York: Routledge Falmer.
- Coolican, H. (1999). *Research Methods and Statistics in Psychology* (3rd ed.). London: Hodder & Stoughton.
- Cossa, E.F.R. (2007). *A case study of practical work in a cell biology course at the Eduardo Mondlane University in Mozambique*. Doctoral dissertation. Cape Town: University of the Western Cape.
- Creswell, J.W. (2003). *Research design: qualitative, quantitative, and mixed methods approaches* (2nd ed.). Thousand Oaks, London, New Delhi: Sage Publications.
- Cupane, A.F. (2007). Towards a culture-sensitive pedagogy of Physics teacher education in Mozambique. In I.V. Mutimucuo & M. Cherinda (Eds.), *Proceedings of the 15th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 335-342).
- Dekkers, P.J.J.M., & Thijs, G.D. (1995). *A teaching sequence for the 'force' concept developed from classroom research on practical work. A strategy for providing means to solve 'dissonance'*. Presentation to the European Conference on Research in Science Education, University of Leeds, 7-11 April 2005, Leeds.
- Dekkers, P.J.J.M. (1997). *Making productive use of student conceptions in Physics education: Developing the concept of force through practical work*. Doctoral dissertation. De Boelelaan (NL): VU Huisdrukkerij, Amsterdam.
- Fagilde, S.A.M. (2002). *Gender patterns in communication in mathematics classrooms: A Mozambican case study*. Doctoral dissertation. Cape Town: University of the Western Cape.
- Fink, A. (1995a). *How to design surveys*. California: Sage publications.
- Fink, A. (1995b). *How to ask survey questions*. California: Sage publications.

- Fink, A. (1995c). *The survey handbook*. California: Sage publications.
- Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.
- Gardner, J. (Ed.) (2006). *Assessment and learning*. London, Thousand Oaks, New Delhi: Sage Publications.
- Garet, M.S., Porter, A.C., Desimone, L. Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- Garrett, R.M., and Roberts, I.F. (1982). Demonstration versus Small Group Practical Work in Science Education. A critical review of studies since 1900. *Studies in Science Education*, 9, 109-146.
- Gay, L.R., & Airasian, P. (2003). *Educational research: competencies for analysis and applications*. Columbus, Ohio: Merrill Prentice-Hall.
- Governo de Moçambique (1996). *Boletim da República: Regulamento de avaliação do ensino secundário geral*. Maputo: Imprensa Nacional.
- Gronlund, N.E. (1998). *Assessment of student achievement* (6th ed.). MA: Allyn & Bacon.
- Guskey, T.R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Hadi, S. (2002). *Effective, teacher professional development for implementation of realistic mathematics education in Indonesia*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.

- Hall, G.E., & Hord, S.M. (2001). *Implementing change: Patterns, principles, and potholes*. Massachusetts: Allyn and Bacon.
- Harlen, W. (2006). The role of assessment in developing motivation for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 61-80). London, Thousand Oaks, New Delhi: Sage Publications.
- Hestenes, D., Wells, M. & Swackhamer, G. (1992). Force concept inventory. In *The Physics Teacher Journal*, vol 30. Tempe: Arizona State University
- Hodson, D. (1993). Re-thinking old ways: Towards a more critical approach to practical work in school science. *Studies in Science Education*, 22, 85-142.
- Howie, S. J. (2002). *English language proficiency and contextual factors influencing mathematics achievement of secondary school pupils in South Africa*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- Howie, S. J. (2006). Presentation to Gauteng Department of Education. Assessment for Learning Conference, May 2006, Johannesburg.
- Huillet, D. (2002). Evolution of in-service and pre-service teachers' conceptions on limits of functions, through participation in a research community. In C. Malcolm, & C. Lubisi (Eds.), *Proceedings of the 10th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 124-129).
- INDE (2005). Proposta de estrutura para o Ensino Secundário Geral (documento em estudo). Maputo: INDE.
- INE (2007). *Statistical Yearbook, 2007*. Maputo: National Institute of Statistics.
- James, M. (2006). Assessment, teaching and theories of learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 47-60). London, Thousand Oaks, New Delhi: Sage Publications.

- James, M., & Pedder, D. (2006). Professional learning as a condition for assessment for learning. In J. Gardner (Ed.), *Assessment and learning* (pp. 27-44). London, Thousand Oaks, New Delhi: Sage Publications.
- Januario, F.M. (1997). Current and final assessment in Mozambican science education: The case study of lower primary schools. In M. Sanders (Ed.), *Proceedings of the 5th Conference of the Southern African Association for Research in Mathematics and Science Education* (pp. 230-236).
- Joyce, B., & Showers, B. (1988, 1995). *Student achievement through staff development*. (1st and 2nd ed.). New York: Longman.
- Johnstone, A.H., & Wham, A.J.B. (1982). The demands of practical work. *Education in Chemistry*, 19, 71-73.
- Kathy, H., & Burke, W. (2003). *Making formative assessment work: effective practice in the primary classroom*. London: Open University Press.
- Kemp, J. & Toperoff, D. (1998). *Guidelines for portfolios assessment in teaching English*. <http://www.anglit.net/main/portfolios/contents#contents>, retrieved August 23, 2006.
- Kouwenhoven, W. (2003). *Designing for competence: Towards a competence-based curriculum for the Faculty of Education of the Eduardo Mondlane University*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- Krathwohl, D.R. (1998). *Methods of educational and social research: An integrated approach* (2nd ed.). New York: Longman.

- Lauchande, C. (2001). Assessment in primary education in Mozambique: Some mathematics results. In I.V. Mutimucio (Ed.), *Proceedings of the 9th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 297-308).
- Lin, R.L. & Gronlund, N.E. (2000). *Measurement and assessment in teaching* (8th ed.). Columbus, Ohio: Merrill Prentice-Hall.
- Lincoln, Y.S., & Guba, E.G. (2000). Paradigmatic controversies, contradictions, and emerging influences. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 163-188). Thousand Oaks, CA: Sage Publications.
- Machado, A.J.E. (2007). A avaliação: Um dos grandes problemas no processo de ensino e aprendizagem da Matemática no ensino secundário em Moçambique. In I.V. Mutimucio & M. Cherinda (Eds.), *Proceedings of the 15th Conference of the Southern African Association for Research in Mathematics, Science and Technology Education* (pp. 220-225).
- Mafumiko, F. (2006). *Micro-scale experimentation as a catalyst for improving the chemistry curriculum in Tanzania*. Doctoral dissertation. Enschede (NL): University of Twente.
- McKenney, S. (2001). *Computer-based support for science education materials developers in Africa: exploring potentials*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- McMillan, J.H. (2001). *Essential assessment concepts for teachers and administrators*. California: Corwin Press.
- MEC (2007). *Relatório Anual de 2006. Documento número 1.01/RAR8*. Oitava Reunião Anual de Revisão do Plano Estratégico de Educação e Cultura. Maputo: MEC.

- MEC & INDE (2007). *Plano Curricular do Ensino Secundário Geral (PCESG): Documento orientador - objectivos, política, estrutura, plano de estudos e estratégias de implementação*. Maputo: Imprensa Universitária.
- Mertens, D.M. (1998). *Research methods in education and psychology: Integrating Diversity with quantitative and qualitative approaches*. Thousand Oaks, CA: Sage Publications.
- Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis: an expanded sourcebook* (2nd ed.). Thousand Oaks, London, New Delhi: Sage Publications.
- MinEd (1994). *Indicadores educacionais e efectivos escolares: Ensino primário 1983 1992*. Maputo: Direcção de Planificação, MinEd.
- MinEd (1997). *Programas de Física, 2º Ciclo*. Maputo: Ministério da Educação.
- MinEd (1998). *Education Sector Strategic Plan 1999-2003. Reviving schools and expanding opportunities*. Maputo: Ministry of Education.
- MinEd (1999). *Plano Estratégico da Educação*. Maputo: MinEd.
- MinEd (2001). *Secondary and Secondary Teacher Education Strategic Plan*. Maputo: Ministry of Education.
- MinEd (2003). *Levantamento estatístico anual*. Maputo: MinEd.
- MinEd (2004). *Education Sector Strategic Plan II, 2005-2009 (Draft)*. Maputo: Ministry of Education (Mimeo).
- Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation*, 8(14). <http://PAREonline.net/getvn.asp?v=8&n=14>, retrieved April 26, 2006.

- Motswiri, M.J. (2004). *Supporting chemistry teachers in implementing formative assessment of investigative practical work in Botswana*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- Muller, J. (2006). *Authentic assessment toolbox*. <http://jonathan.mueller.faculty.noctrl.edu/toolbox/examples.htm>, retrieved April 27, 2006.
- Mutumucuo, I.V. (1998). *Improving students' understanding of energy: A study of the conceptual development of Mozambican first-year university students*. Doctoral dissertation. De Boelelaan (NL): VU Huisdrukkerij, Amsterdam.
- National Research Council (NRC) (2001). *Knowing what students know: the science and design of educational assessment*. Committee on the Foundations of Assessment. J. Pelligrino, N. Chudowsky, and R. Glaser (Eds.). Washington, D.C.: National Academy Press.
- Nieveen, N.M. (1997). *Computer support for curriculum developers: a study on the potential of computer support in the domain of formative curriculum evaluation*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- Nieveen, N.M. (1999). Prototyping to reach product quality. In J. van den Akker, R.M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), *Design approaches and tools in education and training*. Boston: Kluwer Academic, 125-136.
- Ogunniyi, M.B. (1986). Two decades of science education in Africa. *Science Education*, 70, 111-112.
- Ottevanger, W.J.W. (2001). *Teacher support materials as a catalyst for science curriculum implementation in Namibia*. Doctoral dissertation. Enschede (NL): PrintPartners Ipskamp.
- Palme, M. (1992). *O significado da escola. Repetência e desistência na escola primária moçambicana*. Maputo: INDE.

- Pelgrum, W.J. (1989). *Educational assessment: Monitoring, evaluation and the curriculum*. De Lier: Academic Boeken Centrum.
- Piaget, J. (1972). *The science of education and the psychology of the child*. New York: Orion Press.
- Plomp, T. (2004). *Quality assurance in Netherlands education*. Workshop held at the University of Pretoria, 5 August.
- Plomp, T. (2006). *Educational Design Research: a research approach to address complex problems in educational practice*. Paper presented at The Fifth International Forum on Educational Technology, October 29-30, Central China Normal University, Wuhan, China.
- Popham, W.J. (2002). *Classroom assessment. What teachers need to know* (3rd ed.). Boston, MA: Allyn and Bacon.
- Popov, O. (1994). *Qualidade e avaliação em Ciências Naturais no Ensino Primário em Moçambique*. Maputo: INDE.
- Race, P., Brown, S., & Smith, B. (2005). *500 tips on assessment* (2nd ed.). London and New York: RoutledgeFalmer.
- Reeves, T.C. (2000). Socially responsible educational technology research. *Educational Technology*, 40(6), 19-28.
- Richey, R., Klein, J., & Nelson, W. (2004). Developmental research: Studies of instructional design and development. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2 ed.) (pp. 1099-1130). Mahwah, NJ: Lawrence Erlbaum Associates.
- Seels, B.B., & Richey, R.C. (1994). *Instructional technology: The definitions and domains of the field*. Washington, DC: Association for Education and Communications and Technology.

- Shavelson, R.J., McDonnell, L.M., & Oakes, J. (1987). *Indicators for monitoring mathematics and science education: a sourcebook*. Santa Monica (CA, USA): The RAND Corporation.
- Sitoe, A. (2006). *Epistemological beliefs and perceptions of education in Africa: An explorative study with high school students in Mozambique*. Doctoral dissertation. Groningen (NL): Centre for Development Studies.
- Stiggins, R. J. (1987). The design and development of performance assessments. *Educational Measurement: Issues and Practice*, 6, 33-42.
- Sykes, G. (1999). Teacher and student learning: Strengthening their connection. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 151-179). San Francisco: Jossey-Bass.
- Tamir, P. (1991). Practical work in school science: An analysis of current practice. In B.E. Woolnouth (Ed.), *Practical science* (pp. 13-20). Milton Keynes: Open University Press.
- Teclé, A. T. (2006). *The potential of a professional development scenario for supporting biology teachers in Eritrea*. Doctoral dissertation. Enschede (NL): University of Twente.
- Tessmer, M. (1993). *Planning and conducting formative evaluations*. London: Kogan Page.
- Thijs, G.D. (1987). Conceptions of force and movement. In J.D. Novak (Ed.), *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*. Vol. III. Ithaca, N.Y.: Cornell University, pp. 501-513.
- Thijs, A. (1999). *Supporting science curriculum reform in Botswana: the potential for peer coaching*. Doctoral dissertation. Enschede (NL): University of Twente.

- Treagust, D.F., Duit, R., & Fraser, B.J. (1996). *Teaching and learning in science and mathematics*. New York: Teachers College Press.
- UNDP (2000). *Human Development Report – Mozambique*. Maputo: UNDP.
- UNDP (2004). *National Human Development Report, 2003*. Maputo: UNDP.
- van den Akker, J. & Plomp, Tj. (1993). *Development research in curriculum: Propositions and experiences*. Paper presented at AERA Annual Meeting, April 12-16, Atlanta.
- van den Akker, J. (1999). Principles and methods of development research. In J.V.D. Akker, R. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.) (1999). *Design approaches and tools in education and training*. Dordrecht: Kluwer.
- van den Akker, J. (2003). Curriculum perspectives: An introduction. In J. V. D. Akker, W. Kuiper & U. Hameyer (Eds.), *Curriculum landscapes and trends* (pp. 1-10). Dordrecht: Kluwer Academic Publishers.
- van den Akker, J., Gravemeijer, K, McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. London: Routledge. ISBN10: 0-415-39635-2
- van den Berg, E., & Giddings, G. (1992). *An alternative view of laboratory teaching*. Bentley: Curtin University. Monograph.
- Walker, D. (1992). Methodological issues in curriculum research. In P. Jackson (Ed.), *Handbook of Research in Curriculum* (pp. 98-118). New York: Macmillan.
- Weeden, P., Winter, J. & Broadfoot P. (2002). *Assessment: What's in it for schools?* London: Routledge Falmer.
- White, R., & Gunstone, R. (1992). *Probing understanding*. London, New York and Philadelphia: The Falmer Press

Wiggins, G. P. (1993). *Assessing student performance*. San Francisco: Jossey-Bass Publishers.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, London, New Delhi: Sage Publications.

APPENDICES

Appendix A: Questionnaire for teachers and school directors

Code of school

A. ABOUT THIS QUESTIONNAIRE:

What is it?

This questionnaire is part of a doctoral study project and is designed and administered only for graduation purposes. You and your school were chosen to help the project and your school to find out more about the assessment practices currently in use and, if the need is there, how to improve them.

Why should you fill it in?

It will help you and your school to know more about the students and how they can be fairly assessed. The information may also be valuable for research. However, if there are any parts you do not wish to answer, then leave them blank.

Is it confidential?

Yes. No one in the school will know what you have written. You should fill it in without anyone seeing what you write, and without talking to anyone. When the school gets the information back, they will not know what any individual said, only the overall results.

Is it a test?

No. There are no right or wrong answers, so you should not worry about it. Please just answer as honestly as you can.

B. BACKGROUND INFORMATION:

1. First Name:

2. Last Name:

3. Gender: (write *M* or *F*)

4. Age: years

5. School:

6. Grade: 7. Class: 8. Subject:

9. Topic:

10. How many years have you been teaching Physics (including this year)? year(s).

SECTION 1
C: ASSESSMENT PRACTICES APPLIED BY TEACHERS AND THEIR QUALITY

1. Why do you do assessment in your classroom? Tick all that apply.

- a. to maintain social environment in the classroom
- b. to place students
- c. to plan and conduct instruction
- d. to provide feedback and incentives
- e. to diagnose students' strengths and weaknesses
- f. to judge and grade learning and progress
- g. to provide information to policymakers
- h. to satisfy the demands of the parents
- i. others (specify) _____.

2. How well do students understand the purpose of assessment that you apply? Tick the appropriate box.

- a. don't understand
- b. understand some of it
- c. understand most of it
- d. understand very well

3. How much of the teaching time do you spend in your class on assessment per week? Tick the appropriate box.

- a. less than 1/4h
- b. between 1/4h – 1/2h
- c. between 1/2h – 3/4h
- d. between 3/4h – 1h
- e. more than 1h

4. How often do you use each of the following assessment practices? Tick the appropriate box.

| | a. Never | b. Monthly | c. Weekly | d. Daily |
|----------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 4.1 performance assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.2 portfolio assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.3 homework | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.4 paper-and-pencil tests | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.5 projects | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.6 quizzes (verbal tests) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.7 worksheets | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 4.8 peer-assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

5. How often do you: Tick the appropriate box.

| | a. Never | b. Sometimes | c. Frequently | d. Always |
|---|--------------------------|--------------------------|--------------------------|--------------------------|
| 5.1 judge student performance using a certain criterion. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.2 use scoring sheets to explain in advance each criterion. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.3 determine how acceptable is a student's performance really is. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.4 determine a student's progress through his/her particular evolving work. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.5 allow students to evaluate their own work. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.6 allow students to supply their answers to the questions in writing. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.7 provide students with direct questions rather than incomplete statements. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.8 allow students to assess the work of their colleagues. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.9 allow students to assess their own work. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.10 use oral questions to judge the student performance. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| 5.11 give homework to students. | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |

SECTION 2

D: RELEVANCE OF ASSESSMENT PRACTICES

1. How do you inform students about a planned assessment? Tick only one box.

- a. I never inform them
- b. orally
- c. through a written note
- d. other, specify _____

2. How do you engage students in the evaluation of their performance? You can tick more than one box

- a. I don't involve them at all
- b. by handing the results out
- c. by involving them in self-assessment
- d. by sharing with them the goals to be achieved
- e. by explaining them the implications of the results
- f. by reflecting with them on the assessment data

3. What kind of advice do you give to students when you hand out the results? Tick only one box.

- a. I give no advice at all
- b. I give some comments on the students' weaknesses
- c. I give some comments on the students' strengths
- d. I give some comments on both strengths and weaknesses
- e. I do a review and reflection on assessment data
- f. other, specify _____

4. How often do you use the assessment results in class teaching and assessment?

Tick the appropriate box.

a. Never b. Sometimes c. Frequently d. Always

4.1 to assign a grade

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|--------------------------|

4.2 to identify strengths and weaknesses of the students

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|--------------------------|

4.3 to help students know and recognize the standards they are aiming for

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|--------------------------|

4.4 to encourage active involvement of students in their own learning

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|--------------------------|

E: END NOTES

Completely happy Fairly happy Partially happy Not at all happy

1. How did you feel about answering these questions? *Tick only one box*

| | | | |
|--------------------------|--------------------------|--------------------------|--------------------------|
| <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
|--------------------------|--------------------------|--------------------------|--------------------------|

2. Do you have any other comments to make?

THANK YOU VERY MUCH FOR COMPLETING THIS QUESTIONNAIRE

| B.2 TEACHING/LEARNING ENVIRONMENT | Yes | No | Condition/Comments |
|---|------------|-----------|---------------------------|
| 1. Are there displays in the classroom (such as Physics models, lens, graphs etc.)? | | | |
| 2. Are there commercial posters on the walls? | | | |
| 3. What <i>media</i> equipment is available in the classroom (such as computers, overhead projector, TV/Monitors, tape recorders, videos etc.)? | | | |
| 4. Are there teacher-made posters on the walls? | | | |
| 5. Does the teacher use worksheets? | | | |
| 6. Are there displays of students' work? | | | |
| 7. Has every student got a Physics textbook? (if a group of student shares a textbook indicate how many students are there per group) | | | |
| 8. Is Physics equipment available? (Note what equipment is available) | | | |
| 9. Is the available Physics equipment used by the teacher? | | | |
| 10. Is the available Physics equipment used by the students? | | | |

| C. DESCRIPTION OF STUDENTS |
|--|
| 1. How many students are in the class including those who are absent? |
| |
| 2. How many boys and girls are in the class? |
| |
| 3. Other comments with regard to the students. (Look at dress, general appearance of the students, jewellery, anything that could be related to SES) |
| |
| 4. Is there anything else that you observed that is important for this research that was not included in this schedule? |
| |

D. DESCRIPTION OF TEACHERS ASSESSMENT PRACTICES AND OF THEIR QUALITY

| | | | | |
|--|--|--|--|--|
| 1. Extent to which assessment practices are applied presently | | | | |
| | a | b | c | d |
| i) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | N/A | student performance is not judged | students performance is judged | students performance is judged by certain criteria |
| 1.1 Performance assessment | | | | |
| | a | b | c | d |
| ii) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | N/A | criteria are explained | some criteria are explained but not by prepared scoring sheets | some criteria are explained by prepared scoring sheets |
| 1.2 Portfolio assessment | | | | |
| | a | b | c | d |
| i) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | N/A | student progress is not determined | student progress is determined | student progress is determined by some student evolving work |
| | a | b | c | d |
| ii) | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| | N/A | there is no ability of the students to evaluate their own work | some students are able to evaluate their own work | the ability of all students to evaluate their own work is increased somewhat |
| | e | | | e |
| | <input type="checkbox"/> | | | <input type="checkbox"/> |
| | students performance is judged by multiple criteria | | | students performance is judged by multiple criteria |
| | <input type="checkbox"/> | | | <input type="checkbox"/> |
| | an criteria are explained by prepared scoring sheets | | | an criteria are explained by prepared scoring sheets |
| | <input type="checkbox"/> | | | <input type="checkbox"/> |
| | acceptability of student's performance is determined by exclusively through human judgment | | | acceptability of student's performance is determined by exclusively through human judgment |
| | <input type="checkbox"/> | | | <input type="checkbox"/> |
| | student progress is determined by particular student evolving work | | | student progress is determined by particular student evolving work |
| | <input type="checkbox"/> | | | <input type="checkbox"/> |
| | the ability of all students to evaluate their own work is greatly increased | | | the ability of all students to evaluate their own work is greatly increased |

iii) N/A the student growth is not evaluated some student growth is evaluated sometimes all student growth is evaluated over time

i) N/A a b c student do not select answers student select answers from many options student select the true or false answer

ii) N/A student do not select answers student match corresponding answers student supply answers through filling in blank spaces or constructing his/her own responses

1.3 Paper-and-pencil tests

i) N/A a b c d e homework is not given homework given but not marked students mark own homework students homework marked by the teacher

ii) N/A homework is not checked by the teacher homework checked monthly by the teacher homework checked weekly by the teacher homework checked daily by the teacher

1.4 Homework

| | | | | |
|-----------------------------------|--|--|--|--|
| iii) <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| N/A | homework is not shown to parents | homework is shown to parents | homework is checked by parents | homework is signed by the parents |
|iii) <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| N/A | homework is not checked | some homework checked for some students | all homework checked for some students | all homework checked for all students |
| | | | | |
| 1.5 projects | | | | |
| i) <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| N/A | students do not do an elaborated piece of work | students do an elaborated piece of work | students do a pre-set elaborated piece of work | students do a pre-set elaborated piece of work |
| ii) <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| N/A | students do not have assignments out of school | students do some assignments out of school environment | students do some assignments out of school environment over several days | students do some assignments out of school environment over several days |
| iii) <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> |
| N/A | students do not produce reports of their work | students produce reports of their work | students produce reports of a pre-set structure piece of work | students produce reports of a pre-set structure piece of work |

1.6 Others

2. The quality of the assessment practices as demonstrated by teachers and students

| | a. Never | b. Some of the time | c. Most of the time | d. Always | Comments (Cts) |
|--|--------------------------|--------------------------|--------------------------|--------------------------|----------------|
| 2.1. The teacher provides the students with the objectives of the assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 2.2. The teacher discusses with the students the intended learning outcomes of the assessment task | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.3. The teacher provides clear and consistent procedures for the assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.4. The resources available are adequate for the requirements of the assessment procedures | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.5. The time allocated is sufficient for the requirements of the assessment procedures | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.6. The teacher assigns homework at the end of each lesson | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.7. The teacher collects the student's assignments | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.8. The teacher takes note of each student's mark | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.9. The teacher explains the students how far they have achieved the intended outcomes | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.10. The teacher discusses the assessment results with students | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.11. The teacher discusses the assessment results with colleagues | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 2.12. The teacher gives feedback to the students | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |

3. Are assessment practices appropriate for instruction?

| | a Never | b Some of the time | c Most of the time | d Always | Cts |
|--|--------------------------|--------------------------|--------------------------|--------------------------|-------|
| 3.1 Students informed about what will be assessed and how | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.2 Relevant procedures and follow-up actions designed in consultation with students | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.3 There is a balance between giving marks and giving advice | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.4 When possible, students are engaged in the evaluation | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.5 When possible, parents are involved in the evaluation of the students | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.6 Teachers use assessment practices as integral to teaching | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.7 Teachers involve students in sharing goals | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.8 Teachers help students know and recognize standards are aiming for | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.9 Teachers provide feedback to students | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.10 Teachers involve students in self-assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.11 Teachers involve students in peer-assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.12 Teachers encourage active involvement of students in their own learning | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 3.13 Teachers and students are involved in reviewing and reflection on assessment data | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |

4. How valid and reliable are the assessment practices?

| | a Never | b Sometimes | c Frequently | d Always | Comments (Cts) |
|---|--------------------------|--------------------------|--------------------------|--------------------------|----------------|
| 4.1 Cover all the important aspects of specific Physics topic to be assessed | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | _____ |
| 4.2 Assessment methods allow one to make valid decisions on instruction and assessment | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.3 Assessment questions allow students to demonstrate the performance being assessed | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.4 Directions and wording are clear enough that students know what is expected of them | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.5 Assessment is related to what the students have been taught | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.6 The scoring procedures are clear | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.7 The scoring procedures are consistent | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |
| 4.8 The scoring procedures are unbiased | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | Cts _____ |

Appendix C: Interview schedule for teachers

This interview was designed as part of a doctoral study project and will be conducted only for graduation purposes. You and your school were chosen to help the project and your school to find out more about the assessment practices currently in use and, if the need is there, how to improve them. The interview will be conducted in an informal and conducive environment and it will not last more than thirty minutes. All the information to be gathered from the interview is confidential and your identity as interviewee will be kept unknown. Feel free to answer the questions and be as honest as possible.

1. How do you assess your students in the classroom during the year?
-How do you know your students are making progress?
2. How does this compare to the information collected by the exam at the end of the year?
3. What is the purpose of your classroom assessment?
4. Do you think that your students understand the main purposes of the assessment?
- Why?
5. How often do you assess your students per week?
6. What do you understand by "performance assessment"?
-Is this different from what you do, for instance, in paper-and-pencil tests?
7. How do you assess project work?
8. What do you understand by "portfolio assessment"?
- Have you ever determined a student's progress through his/her particular evolving work?
- How?
9. How do you assess student's homework?

10. Do you do peer-assessment in your class and can you tell me about this?

-Do you allow students to assess their own work or that of their colleagues?

11. Tell me what do you do with the assessment results of your students

12. How do you evaluate the performance of your students?

13. Do you have any other comments to make?

Appendix D: Interview schedule for school directors

This interview was designed as part of a doctoral study project and will be conducted only for graduation purposes. You and your school were chosen to help the project and your school to find out more about the assessment practices currently used by teachers in your school and, if the need is there, how to improve them. The interview will be conducted in an informal and conducive environment and it will not last more than thirty minutes. All the information to be gathered from the interview is confidential and your identity as interviewee will be kept unknown. Feel free to answer the questions and be as honest as possible.

1. How do teachers assess their students in the classroom during the year?
-How do you know the teachers are assessing their students?
2. How does this compare to the information collected by the exam at the end of the year?
3. What is the purpose of teachers' assessment?
4. Do you think that the teachers understand the main purposes of the assessment?
-Why?
5. Can you tell me about the frequency of teachers' assessment?
- How often do teachers assess their students per week and how do you come to know?
6. What do you understand by "portfolio assessment"?
-Is this different from what teachers do, for instance, in paper-and-pencil tests?
7. Tell me what do you do with the assessment results of students
8. How do you evaluate the performance of your students?
9. Do you have any other comments to make?

Appendix E: Interview schedule for pedagogical officers

This interview was designed as part of a doctoral study project and will be conducted only for graduation purposes. You were chosen to help the project and the schools to find out more about the assessment practices currently used by teachers in schools and, if the need is there, how to improve them. The interview will be conducted in an informal and conducive environment and it will not last more than thirty minutes. All the information to be gathered from the interview is confidential and your identity as interviewee will be kept unknown. Feel free to answer the questions and be as honest as possible.

1. According to the Ministry, what are the objectives of the teachers when they assess their students in the classroom during the year?
2. How does this compare to the information collected by the exam at the end of the year?
3. In your opinion, what should be the purpose of teachers' assessment?
4. Do you think that the teachers understand the main purposes of the assessment?
- *Why?*
5. Can you tell me about the frequency of the inspectorate visits to schools and teachers?
- *What are their main purposes?*
6. What are the Ministry's mechanisms of verifying how teachers are assessing their students?
- *How are they efficient?*
7. What do you understand by "portfolio assessment"?
- *Is this different from what teachers do, for instance, in paper-and-pencil tests?*
8. Tell me what do teachers do with the assessment results of students

9. How do you evaluate the performance of your students?

10. Do you have any other comments to make?

Appendix F: Guide for expert appraisal

My request is that you help me review the evaluation instruments G, H, I, J and K for *consistency* with the design guidelines in documents (a) "Demonstration experiments" and (b) "Template for the Demonstration Experiment Report". To guide the appraisal the following questions can be considered:

1. Is there consistency between evaluation instruments G to K and:
 - a) Demonstration experiments
 - b) Template for the Demonstration Experiment Report

2. Are the various items in each evaluation instrument (G to K) specific enough to convey the intentions of the developer of establishing *practicality*?

3. Please use questions 1-2 to make comments and suggestions for improvement of the material enclosed.

Appendix G: University students' questionnaire (before tryout)

Dear student,

You have been chosen to participate in the appraisal of the Physics demonstration experiment materials intended to promote Physics learning by specifically focusing on assessing students' understanding of the inertia. The questionnaire focus on how did you perceive *practicality* of the materials as used in the classroom in terms of relevance, structure, content, and presentation as well as the suitability of the POE strategy. The information that you provide will help to improve the teaching and assessment of Physics.

1. Please describe your general impression about the prototype in terms of:

a. Relevance of material:

b. Structure:

c. Relevance of the content:

d. Presentation:

1. What did you like and dislike about the experiments? Why?

2. What things would you like to have taken out of these experiments?

3. What things would you like to have added to these experiments?

4. What potential problems do you foresee about doing these experiments in class?

5. Do you feel that the POE teaching strategy is helpful for students' reasoning?

How? _____

6. Any other comments or suggestions

Thank you very much for your cooperation!

Appendix H: University student's follow-up interview (before tryout)

Dear student,

The interview focus on how did you perceive *practicality* of the performance assessment materials as used in laboratory demonstrations in terms of relevance of the content, presentation of the material, and the POE strategy.

1. What general comment can you make about the material and the experiments with inertia?
2. At what extent the prototype is useful for the preparation of the experiments?
3. What things would you like to see in the prototype and they are missing?
4. What specific comments can you make for each experiment?
5. What do you feel about the *practicality* of the POE strategy?
6. How do you consider the time allocated for teaching the syllabus to accommodate the suggested strategy?
7. Do you think the teachers will like the approach? Why or why not?
8. Do you think the students will like the approach? Why or why not?
9. Do you think that the teacher's role was clear? And that of the students? Explain.
10. Do you have any comments or suggestions for improvement?

Appendix I: Teacher's evaluation questionnaire (after tryout)

Dear teacher,

You have participated in Physics demonstration experiments intended to enhance Physics learning. The experiments followed a Predict-Observe-Explain (POE) strategy and focused on assessing students' understanding of the concepts of force and inertia. By means of this questionnaire I would like to have your opinions about the approach and the materials used, and your experiences with student performance assessment in the demonstration experiments. The information that you provide will help to improve the design and development of Physics assessment practices and the teaching of Physics in general.

1. General information

Date ___/___/___ Classes taught _____ Your age _____

Academic qualifications _____

Teaching experience (years) _____ In this subject (years) _____

2. General impression of the Physics assessment materials (Part 3 of this document)

2.1 Is the language clear and understandable for students? *Circle on the correct number.*

1. Yes 2. No

If not, explain what the problems were? _____

2.2 Was the description of the experiments clear for the students or did they have many questions? If there were many questions what were the students' questions? _____.

2.3 Were the pictures clear? *Circle on the correct number.*

1. Yes 2. No

If not, what are the needed improvements? _____

2.4 Was the teacher's guide (Part 2 of this document) useful during preparation of the experiments? *Circle on the correct number.*

1. Yes 2. No 3. Partly

Which parts were useful and why? _____

Which parts need improvements and why? _____

2.5 How much time did you spend in each activity?

| <i>Activity</i> | <i>Time spent (minutes)</i> |
|--------------------------------------|-----------------------------|
| Looking for the equipment | _____ |
| Trying out each experiment | _____ |
| Grouping the students | _____ |
| Introducing the task to the students | _____ |
| Other, specify _____ | _____ |

3. Teaching and assessment strategy

3.1 Was the POE strategy of teaching and assessment practical for students' reasoning?

Circle on the correct number.

1. Yes 2. No 3. Partly

Which parts were useful and why? _____

Which parts need improvements and why? _____

3.2 What was your role as a teacher during the experiment? *Tick one or more options*

Explainer of all students

Active participant

Guide students with difficulties

An interested spectator

Other, specify _____

3.3 Where there students who were not active during the experiments? *Circle on the correct number.*

1. Yes 2. No

If yes, why? _____

3.4 Do you feel that the main objective of these experiments was achieved? *Circle on the correct number.*

1. Yes 2. No

If not, please indicate which particular aspects were not met and why. _____

3.5 Do you feel the prototype as a whole needs any changes or additions? *Circle on the correct number.*

1. Yes 2. No

If yes, what changes? _____

3.6 Any other comments or suggestions: _____

Thank you very much for your cooperation!

Appendix J: Follow-up interview with teachers (after tryout)

The interview focus on how did you perceive *practicality* of the performance assessment materials as used in laboratory demonstrations in terms of relevance of the content, presentation of the material, and the POE strategy.

1. Personal data

Age, qualifications, teaching experience, position in school, others.

2. Expectations

- a. What did you expect about conducting the experiments about force and inertia?
 - i. Organizational problems
 - ii. Discipline problems
 - iii. Other problems
- b. While you were conducting the experiments, were they as you expected? Explain.
- c. After the experiments, have you met your expectations?

3. General impression of the material

- a. Please describe your general impressions about the material and the experiments.
- b. What specific comments can you make for each demonstration experiment?
- c. What did you like and dislike about the demonstration experiments, and why?
- d. What is your impression of the structure and clarity of the assessment materials?
- e. Do you feel that the materials are practical and usable in terms of resources/equipment needed to carry out the experiments? Explain.

- f. Was it easier to follow the POE strategy and how much of this constituted a problem in your guidance of students through the experiments?
- g. How do you consider the time you have for teaching the syllabus to accommodate the suggested practice?

4. Teaching and assessment strategy

- b. What do you feel about the *practicality* of the POE strategy?
- c. Do you think that your role as a teacher was clear? And that of the students?
- d. What do you think was expected of you during the demonstration experiments in terms of formative assessment of the students' work
- e. How do you find the assessment of the students' work through the lab report?
- f. How practical was the use of POE strategy in terms of conducting and assessing the demonstration experiments?
- g. How do you compare the POE strategy with the way you used to conduct your demonstration experiments?
- h. Do you have any comments or suggestions for improvement?

Appendix K: Students' questionnaire (after tryout)

Dear student,

During the last days you were involved in some demonstration experiments about force and inertia. I would be grateful to receive your comments about all the activities you were involved in including the writing of the demonstration experiment report. There are no wrong or right answers and, please, do not discuss your views with someone else while answering this questionnaire.

1. Date ___/___/___ Class _____ Age _____ years old

2. What did you like and dislike about the experiments?

I liked: _____

I disliked: _____

3. What did you like most about the experiments? _____

Please, state the reasons why. _____

4. What did you like least about the experiments? _____

Please, state the reasons why. _____

5. Were the demonstration experiments different from the usual Physics experiments you are used to in your class? *Please tick the appropriate box*

Yes

No

If yes, what were the differences? _____

6. How do you describe your participation in the experiments? *One option in each line*

Passive 1 2 3 4 Active

Dependent 1 2 3 4 Independent

Uninterested 1 2 3 4 Interested

Any other comment: _____

7. Do you think that all other members of your group were active? *Please tick the appropriate box*

Yes

No

If not, why? _____

8. Please tick one or more of the following if you agree with them.

- The prototype was very well structured
- The prototype was unclear and confusing
- The prototype was practical and easy to use in the lab

9. Did you face any problems during the execution of the experiments? *Please tick the appropriate box*

- Yes
- No

If yes, what the problems were? _____

10. Please, write any other comments or suggestions you may have _____.

Thank you very much for your cooperation!

Appendix L: Letter from the Ministry of Education and Culture to schools

TO WHOM IT MAY CONCERN

This is to certify that **Mr Francisco Januário**, from Eduardo Mondlane University, Faculty of Education is authorized to undertake a research project in Francisco Manyanga Secondary School in Maputo. The project, entitled *Investigating and improving assessment practices in secondary schools in Mozambique*, is undertaken in the framework of Mr Januário's doctoral studies held at the University of Pretoria, South Africa. Activities under the project include interactive sessions with Grade 11 and 12 Physics teachers meant to design and develop assessment materials for teachers to use in the classroom. These activities should be undertaken during the normal school period and professional and physical integrity of teachers should be guaranteed.

National Director of General Education

Appendix M: Letter from Eduardo Mondlane University to schools



UNIVERSIDADE EDUARDO MONDLANE
FACULDADE DE EDUCAÇÃO

A Direcção da

Nossa Ref. 287(...)/FACED/05

15 de Agosto de 2005

Assunto - Pedido de permissão para um Estudo

A Faculdade de Educação da Universidade Eduardo Mondlane tem como um dos seus objectivos a melhoria da qualidade do ensino através da elevação da competência do seu corpo docente. Neste ambito, tem levado a cabo um programa de formação de docentes em diversas áreas educacionais. Neste momento, o **dr. Francisco Januário**, docente desta Faculdade encontra-se envolvido num programa de formação na área de avaliação e controle de qualidade, mais concretamente na disciplina de Física para a 12ª classe do Ensino Secundário Geral, curso diurno.

Assim sendo, a Faculdade gostaria de solicitar permissão, por escrito, à Direcção da Escola que V. Excia dirige para efectuar este trabalho junto dos professores daquela classe e nível. O programa que o docente pretende levar a cabo com os referidos professores consiste no desenvolvimento de protótipos de ensino de Física da 12ª classe sobre um tópico (ex. Mecânica) e de um portfólio de avaliação do mesmo tópico que é uma ferramenta ou *kit* de avaliação que os professores poderão usar na sala de aulas.

Cientes da Vossa colaboração, subscrevemo-nos com alta estima.

Prof Doutor Inocente Mutimucuo

Director Adjunto Para a Investigação e Pós-Graduação

Appendix N: Letter of consent from the researcher to teachers

**PEDIDO DE PERMISSÃO DO PROFESSOR
PESQUISA E MELHORAMENTO DAS PRÁTICAS DE AVALIAÇÃO EM
FÍSICA NAS ESCOLAS SECUNDÁRIAS EM MOCAMBIQUE**

Caro professor

Agosto de 2005

Por meio desta carta é convidado a participar num projecto de pesquisa destinado a investigar as práticas de avaliação usadas pelos professores de Física da 12^a classe em Moçambique e como essas práticas podem ser melhoradas. A sua participação no projecto é voluntária e confidencial. Não lhe será pedido a revelar qualquer informação que possa conduzir à identificação da sua identidade, a não ser que manifeste o interesse de ser contactado para ser entrevistado individualmente no âmbito de certificação de informação. Mesmo neste caso, a confidencialidade será garantida e poderá tomar a decisão de abandonar a entrevista a qualquer momento se assim o desejar.

A acompanhar esta carta vai um documento explicando o seu papel no processo de pesquisa.

Os resultados deste estudo serão usados, por um lado, para ajudar na monitoração de melhoramentos qualitativos dos resultados dos alunos e no desempenho do sistema educacional em geral. Por outro lado, espera-se que com o melhoramento das práticas de avaliação por meio de desenvolvimento de protótipos (planos de aula) de ensino de Física e de *portfólios* (kits) de avaliação da mesma disciplina possam servir de ferramenta de apoio ao Ministério de Educação e Cultura na monitoração da qualidade de ensino.

Se desejar participar neste estudo, por favor, assine esta carta como declaração do seu consentimento, i.é., como indicação de que participa no projecto de livre vontade e que compreende que poderá desistir de o fazer a qualquer momento se achar conveniente. A participação nesta fase do projecto não o obriga a participar na posterior entrevista de certificação de informação. No entanto, se decidir participar nesta entrevista, tal participação é também voluntária. Em nenhuma circunstância será revelada à sua escola ou aos seus superiores qualquer informação que o possa prejudicar por ter participado neste projecto.

Assinatura do professor participante _____ Data _____

Assinatura do investigador _____ Data _____

Atenciosamente

dr Francisco Januário

xx

(Teacher's role in the research process)

O papel do professor na pesquisa

O objectivo deste estudo é investigar as práticas de avaliação usadas pelos professores de Física da 12ª classe em Moçambique e como essas práticas podem ser melhoradas. Para isso estão planificadas duas actividades de pesquisa em que o professor estará envolvido.

A primeira consiste no desenho e desenvolvimento de protótipos de ensino de Física da 12ª classe, que são exemplares de planos de aula sobre um tópico (ex. Mecânica) que o professor aborda nas suas aulas. Estes planos contêm passos e estratégias metodológicas de abordagem do tópico nomeadamente a justificação da necessidade de aprendizagem, os objectivos, os conteúdos, as actividades de aprendizagem, o papel do professor na facilitação da aprendizagem e as sugestões de avaliação formativa.

A segunda actividade compreende o desenvolvimento de um portfólio de avaliação do mesmo tópico que é uma ferramenta ou kit de avaliação que o professor poderá usar paralelamente às tradicionais formas de avaliação conhecidas.

As duas actividades decorrerão sob forma interactiva entre os professores (seleccionados) da disciplina e da classe na sua escola e o investigador. Por vezes será necessário, o envolvimento de alguns alunos seleccionados. As sessões de actividade, ou seja, o desenvolvimento destes materiais decorrerá nas horas normais da actividade do professor,

sendo que não será usado tempo extra à normal actividade laboral. Para casos em que isso possa porventura ocorrer, serão criadas condições logísticas adequadas.

No fim da pesquisa o professor terá dois instrumentos produzidos: um exemplar do plano de aula e um kit de avaliação.

Assinatura do investigador _____

Appendix O: Ethical clearance



UNIVERSITY OF PRETORIA
FACULTY OF EDUCATION
RESEARCH ETHICS COMMITTEE

**CLEARANCE CERTIFICATE
DEGREE AND PROJECT**

CLEARANCE NUMBER : CS06/03/11

PhD Assessment and Quality Assurance
Investigating assessment practices in physics in secondary
schools in Mozambique

**INVESTIGATOR(S)
DEPARTMENT
DATE CONSIDERED
DECISION OF THE COMMITTEE**

F Januano
Curriculum Studies (Centre for Evaluation and Assessment)
24 March 2006
APPROVED

This ethical clearance is valid for years from the date of consideration and may be renewed upon application

**CHAIRPERSON OF ETHICS
COMMITTEE**

Dr C Lubbe

A handwritten signature in black ink, appearing to be 'C Lubbe', written over a horizontal line.

DATE

24 March 2006

CC

Prof S Howie
Mrs Jennie Boukes

This ethical clearance certificate is issued subject to the following conditions:

1. A signed personal declaration of responsibility
2. If the research question changes significantly so as to alter the nature of the study, a new application for ethical clearance must be submitted
3. It remains the students' responsibility to ensure that all the necessary forms for informed consent are kept for future queries.

Please quote the clearance number in all enquiries.

Appendix P: Physics Assessment Materials (on a CD - Rom)

XXXXXXXXXXXXXXXXXXXXXX

Appendix Q: Certificate of language editing

BERNICE BRADE EDITING

FREELANCE WRITER, PROOF READER AND EDITOR
WEB RESEARCHER AND RESEARCH STRATEGIST
ENGLISH SPECIALIST

Tel. and Fax +27 11 465 4038
Cell 072 287 9859
Email edit@iafrica.com
28 August 2007

P O Box 940
LONEMILL 2062
South Africa

To whom it may concern

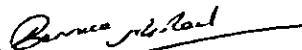
This letter serves to confirm that in August 2007 I proofread and did the language editing for the Doctoral Thesis of

FRANCISCO JANUARIO

This document is being submitted in fulfilment of the requirements for the degree

**Doctor Philosophiae (Assessment and Quality Assurance)
in the Department of Curriculum Studies
Faculty of Education
at
the University of Pretoria.**

I have proofread and edited the entire work, except for the Annexures. This editing principally involves proofreading, language, style and grammar editing; and also checking the text for clarity of meaning, sequence of thought and expression and tenses. I have also noted any inconsistencies in thought, style or logic, and any ambiguities or repetitions of words and phrases, and have corrected those errors which creep into all writing. I have written the corrections on the hard copy and have returned the document to the author who is responsible for inserting these.



Bernice McNeil



Proprietor: Bernice McNeil BA Hons, NSTD
<http://www.ancefire.com/journal/edR/aboutbba.html>