

**THE EFFECTIVE USE OF MULTIPLE-CHOICE QUESTIONS IN ASSESSING  
SCIENTIFIC CALCULATIONS.**

by

HESTER ALETTA TERBLANCHE

submitted in accordance with the requirements for

the degree of

**MASTER OF EDUCATION WITH SPECIALISATION IN SCIENCE EDUCATION**

at the

University of South Africa

Supervisor: Professor Johannes Dreyer

February 2017

Student number: **6697097**

## DECLARATION OF ORIGINALITY

I declare that the dissertation entitled

**THE EFFECTIVE USE OF MULTIPLE-CHOICE QUESTIONS IN ASSESSING  
SCIENTIFIC CALCULATIONS,**

is my own work and that all the sources that I have used or quoted have been indicated and acknowledged by means of complete referencing.

I further declare that I have not previously submitted this work, or part of it, for examination at Unisa for another qualification or at any other higher education institution.

---

SIGNATURE

(Mrs H.A. Terblanche)

---

DATE

## ABSTRACT

This study investigated the effective use of online Multiple-Choice Questions (MCQs) with immediate formative feedback, and the granting of partial credit for correct second or third chance answers when assessing and assisting students' conceptual learning at higher cognitive levels. The research sample comprised first year engineering science students at the Tshwane University of Technology (TUT), Pretoria campus. The differences between using online MCQ-assessment for problem-solving calculations and using constructed written questions (CRQs)<sup>1</sup> in the assessment of problem-solving calculations were explored. Furthermore, the differences between the assessment of problem-solving calculations using online MCQs without immediate formative feedback, and with immediate formative feedback and the granting of partial credit were analysed. The findings revealed that students' marks were lower when answering problem-solving calculations using online MCQs without immediate formative feedback than when answering the same questions using CRQs. This clearly indicates that using online MCQs without immediate formative feedback is not effective in assessing scientific problem-solving calculations. Alternatively, online MCQs proved effective in assessing problem-solving calculations when immediate formative feedback and partial credit were employed. The statistical analysis showed that students performed significantly better when immediate formative feedback was given and partial credit was granted for correct second or third attempts. This was due to online MCQs utilising immediate formative feedback, which made it possible to grant partial credit when students chose the correct answers after feedback. This showed that online MCQs with immediate formative feedback and partial credit being granted can be an effective assessment tool for scientific problem-solving calculations. It increases performance and supports learning from assessment. Students can thus correct their calculations whilst in the process of doing them.

---

<sup>1</sup> Constructed written questions also known as constructed response questions (CRQs) in the assessment of problem-solving calculations are questions that are answered in the traditional written step-by-step way.

### **Key terms:**

Multiple-Choice Questions; MCQs with immediate formative feedback; Guessing in MCQs; Problem-solving calculations; Assessing calculations using MCQs; MCQs in scientific calculations; MCQs at University; Using MCQs to study; MCQs with multiple opportunities; Online MCQs.

## **DEDICATION**

I lovingly dedicate this dissertation to my daughter, Marelize, who was always there to encourage and carry me beyond her own challenges, and through difficult times. Marelize, I admire your kind and serving spirit. May the Lord bless you and keep you!

## **ACKNOWLEDGEMENTS**

I thank the Lord who gave me the instruction to write this dissertation. He indeed keeps His promise to never let me go and guide me through to the end. This was for sure the most important academic journey that I have gone on.

I owe my sincere gratitude to:

- My family and friends for their encouragement, daily prayers and constant love. They were always understanding, taking no for an answer when there was no time for coffee.
- My supervisor, Prof Johann Dreyer for his guidance and critique, especially at the end.
- Melissa, the language editor, for supporting me. Your testimony gave me the last inspiration that I required.

## TABLE OF CONTENTS

CHAPTER 1 : ORIENTATION TO THIS STUDY .....	1
1.1 Introduction.....	1
1.2 Background to and motivation for the research .....	3
1.3 Problem statement and purpose statement.....	5
1.4 Research questions and sub-questions .....	6
1.5 The aim and objectives of this study .....	6
1.6 Research design .....	7
1.6.1 Research paradigm.....	7
1.6.2 Research approach.....	7
1.6.3 Research type .....	7
1.6.4 Research methods .....	8
1.7 Trustworthiness .....	11
1.7.1 Validity .....	12
1.7.2 Reliability.....	12
1.7.3 Objectivity .....	13
1.7.4 Ethical considerations .....	13
1.8 Chapter division.....	14
1.9 Summary .....	15

CHAPTER 2 : LITERATURE REVIEW .....	16
2.1 Introduction.....	16
2.1.1 Definition of terms.....	16
2.2 Contextual and conceptual framework.....	17
2.2.1 Assessment .....	18
2.2.2 Methods of formative assessment.....	23
2.2.3 Multiple-choice questions.....	26
2.2.4 Online MCQs .....	37
2.2.5 Online MCQs with feedback .....	38
2.2.6 Online MCQs with immediate formative feedback.....	40
2.2.7 Online MCQs with immediate corrective formative feedback.....	43
2.2.8 MCQs with immediate formative feedback and the granting of partial credit	44
2.2.9 Student preference for MCQs.....	46
2.3 Theoretical framework .....	49
2.3.1 Bloom’s taxonomy .....	49
2.3.2 The Assessment Component Taxonomy.....	51
2.3.3 Item Response Theory.....	53
2.3.4 Feedback interventions .....	54
2.4 Summary .....	54

CHAPTER 3 : RESEARCH METHODOLOGY .....	56
3.1 Introduction.....	56
3.1.1 The research sub-questions .....	56
3.2 Rationale for empirical research.....	57
3.3 Research design .....	57
3.3.1 Research paradigm.....	58
3.3.2 Research approach.....	58
3.3.3 Research type .....	58
3.4 Research methods.....	59
3.4.1 Sampling.....	59
3.4.2 Data collection .....	59
3.4.3 Data analysis .....	63
3.4.4 Measures for trustworthiness.....	66
3.4.5 Ethical considerations .....	69
3.5 Summary .....	70

CHAPTER 4 : DATA ANALYSIS AND RESULTS .....	71
4.1 Introduction.....	71
4.1.1 Research design and methodology .....	71
4.1.2 Research question and sub-questions .....	71
4.2 Research process .....	72
4.2.1 Reliability and validity .....	75
4.3 Data analysis.....	75
4.3.1 Data analysis procedures .....	75
4.3.2 Distribution analysis .....	76
4.3.3 The testing of the hypotheses.....	77
4.4 Data interpretation.....	81
4.5 Concluding remarks .....	81
4.6 Summary .....	82

CHAPTER 5 : SUMMARY, CONCLUSIONS AND RECOMMENDATIONS.....	83
5.1 Introduction.....	83
5.2 Summary of the research findings.....	83
5.2.1 Key findings from the literature study .....	83
5.2.2 Key findings from the empirical study.....	85
5.3 Research conclusions.....	87
5.4 Recommendations.....	88
5.5 Avenues for further research .....	90
5.6 Limitations of this study.....	90
5.7 Concluding remarks .....	91
REFERENCES.....	92
APPENDIX A.....	106
APPENDIX B.....	113
APPENDIX C.....	118
APPENDIX D.....	126
APPENDIX E.....	130

## LIST OF TABLES

Table 2.1: Mathematics Component Taxonomy and cognitive level of difficulty .....	51
Table 2.2: Mathematics Assessment Component Taxonomy and cognitive skills .....	52
Table 2.3: Comparison of the successes of MCQs and CRQs .....	53
Table 3.1: Presentation of the differences of the pre- and post-tests.....	66
Table 4.1: Description of the marks for the three assessments .....	76
Table 4.2: Statistical symbols for hypothesis testing.....	78

## LIST OF FIGURES

Figure 1.1: Graphical presentation of the research process.....	10
Figure 2.1: Example of an IFAT® form.....	45
Figure 2.2: Bloom's Taxonomy .....	50
Figure 3.1: The comparison between the average marks of all three tests .....	61
Figure 3.2: Example of computer program screen for assessment M .....	62
Figure 3.3: Example of computer program screen for assessment N.....	62
Figure 3.4: Example of how marks are awarded in CRQs .....	63
Figure 3.5: Example of how marks are awarded in MCQs without IFF.....	64
Figure 3.6: Example of how marks are awarded in MCQs with IFF and the granting of partial credit .....	65
Figure 4.1: Graphical presentation of the research process.....	74
Figure 4.2: Distribution of the three assessments.....	76

## LIST OF ACRONYMS

AUC	Answer-Until-Correct
CRQs	Constructed Response Questions
FI	Feedback Interventions
IFAT	Immediate Formative Assessment Techniques
IFF	Immediate Formative Feedback
IRT	Item Response Theory
MCQs	Multiple-Choice Questions
TUT	Tshwane University of Technology

# CHAPTER 1: ORIENTATION TO THIS STUDY

## 1.1 INTRODUCTION

The aim of this study was to contribute to the current pool of knowledge regarding the use of online Multiple-choice Questions (MCQs) to assess problem-solving calculations with immediate formative feedback (IFF) and with the granting of partial credit. Problem-solving calculations are an integral part of most introductory engineering courses, and students' ability to solve them is traditionally assessed by means of written responses to questions where they have to give constructed answers. This is why these assessments are known as Constructed Response Questions or CRQs. Thus, these problem-solving calculations are usually solved in a constructive way (written down step-by-step), which is regarded as an extremely important part of the teaching and learning of problem-solving calculations in science (Huntley, Engelbrecht & Harding, 2010:141-171). This is especially the case as the procedure to obtain the answer is considered to be as important as the answer itself. Valuing the calculation steps, as well as the answer makes it possible to grant partial credit for some correct steps, although other steps, and the answer, may be wrong. For the assessment of calculations to be fair, it must therefore allow for partial credit for a partially correct answer (i.e. some correct steps in the calculation). This is probably the biggest reason why calculations are mostly not assessed using MCQs if only an answer is given/chosen and it is not possible to grant partial credit because the calculating steps are not presented.

The importance of good assessment practice (also in assessing calculations) is emphasised by Boud (1995:35), who argues

*...that the effects of bad practice are far more potent than bad practice in any other aspect of teaching. Students can, with difficulty, escape from the effects of poor teaching, they cannot escape from the effects of poor assessment.*

Formative assessment encourages student learning by contributing helpful feedback and feed-forward to improve their performance. Formative feedback can thus be very valuable in teaching and learning (Black & Wiliam, 2004:20-50). Gipps and Simpson (2005:175) state that formative assessment without feedback may not add much value to student learning. Therefore, it is important to provide feedback on assessments (Gogri, Shaikh & Venugopal, 2013:2-5) as it is, in fact, the feedback that makes it formative in nature. To

use formative assessment with feedback in CRQs in the assessment of problem-solving calculations is therefore a common occurrence these days. There are various ways of giving feedback that can help ensure learning gains, in higher education too (Engelbrecht & Harding, 2004:217-231), and one of these is to do it online.

Currently, online assessment is becoming more common and more important, particularly as part of e-learning and it should be as effective, trustworthy, reasonable and beneficial to students (Kearns, 2012:34) as other forms of assessment. It is important for students to attain the same results as those who do CRQs in the assessment of problem-solving calculations.

Some American colleges and universities assess students' knowledge with MCQs after completing a chapter of a textbook or a section of work (Mathews, 2006:6). To confirm the importance of assessing students' knowledge at the end of each chapter, Biju and Seibu (2010:403-408) found that approximately 40 to 50 MCQs per chapter were posted on the Blackboard<sup>2</sup> forum as online MCQ-assessments. The students needed to complete these at their own pace after a chapter was done. They were encouraged to do it as part of their learning exercises, although they did not get any marks for it. After the completion of the assessments, through receiving immediate post assessment feedback, they could see the answers and therefore discover what mistakes they had made. They further found in their study that there was a direct alignment between students who made use of the MCQs and better performance in the final examination (Biju & Seibu, 2010:403-408). Online MCQs are thus becoming a popular way of doing online assessment and revision (Tsze, 2013:19-32).

Currently, however, online MCQs are mostly used to assess theory. In his study, though, Tsze (2013:19-32) found that, unlike using CRQs, MCQ-assessments do not ordinarily consider partial understanding of the answers for credit (Tsze, 2013:19-32). Although immediate post assessment feedback is provided, it has no formative function. If online MCQ-assessments are to be more useful and fair, immediate formative feedback (IFF) during the assessment process should be provided as it would assist students to learn from assessment. Not only do students then receive immediate feedback (e.g. the answer

---

<sup>2</sup> The Blackboard system is a virtual communication forum between lecturers and students through announcements used by universities. It is also used to manage the course for students to access.

is wrong), they receive formative corrective feedback (e.g. this is what you probably did wrong, so correct it and try again). One could express it as online MCQs with IFF during the assessment and the granting of partial credit for choosing the correct answer after considering the feedback.

To facilitate this, the online MCQs should be structured in such a way that students need to be given an opportunity to use corrective feedback so that they can get to the correct answer, in other words, they need to be given the chance to choose a different answer. If this is possible, students are penalised for initially choosing a wrong answer (making a mistake), but are awarded partial credit for then correcting the mistake and getting to the correct answer.

One last challenge exists in making online MCQ-assessments with IFF and the granting of partial credit viable, which is determining which distractors should be used when setting online MCQs. These distractors should anticipate the mistakes that students could make so that formative feedback can be built into it to indicate to them what they did wrong and how to correct it. To help to determine which distractors to use, lecturers need to note the common mistakes that students make when doing the same or similar CRQs in the assessment of problem-solving calculations. The ideal scenario would be to set up a database over a period of time, capturing the mistakes that students commonly make, as well as the most common mistakes that lecturers know students may make from their own experience. It is important to realise that it is impossible to make provision for all possible mistakes, but one can use the most common ones as possible distractors. The focus of this study was therefore to investigate the possibility of developing online MCQs designed to provide immediate formative feedback during the assessment process so that students can use this feedback to recalculate and find the correct answer. They could then be awarded partial credit when they correct their mistakes, which indicates that this process could be a possible alternative to the use of CRQs.

## **1.2 BACKGROUND TO AND MOTIVATION FOR THE RESEARCH**

In this study, the researcher has been a lecturer at the Tshwane University of Technology (TUT) in Engineering Science for the past nine years. She lectures science in the extended programme of the course where first year students complete the same curriculum in one year as diploma students do in six months. At TUT, there is a gradual

move towards e-learning and subsequently, towards online assessment. In the Engineering Science extended programme, the problems at the end of each chapter of the textbook are currently assessed online using MCQs, which are placed on the Blackboard platform. These traditional MCQs are also known as fixed choice selected response items because they require students to choose a response from the fixed options presented to them. These types of questions can be marked quickly and provide immediate post assessment feedback (showing if the answer is correct or incorrect) to the students. TUT's management software system, the Blackboard system, MyTutor via Respondus, is a user-friendly platform that can be used to assess the theoretical aspects of the curriculum by means of online MCQs. However, there is a need in engineering courses for the online assessment of problem-solving calculations. At present, this is not an option as the MCQs in use do not make provision for IFF during the assessment process that could allow students multiple opportunities to recalculate and find the correct answer when utilising the feedback, and which could make the granting of partial credit possible.

Engelbrecht and Harding (2003:57-65) indicate that the use of CRQs is more appropriate and effective than the use of MCQs for assessing the logical and consolidation components of learning. Huntley, Engelbrecht and Harding (2011:3-16) find in the assessment component taxonomy, which was mentioned as a theory in the theoretical background, that MCQs may possibly be used to assess these higher cognitive levels. Hadsell (2009:135-141) emphasises that MCQs can have a positive effect if immediate feedback is given. A possible disadvantage of MCQs could be that MCQs are often poorly designed and can only assess the lower two levels of Bloom's Taxonomy (Lister, 2000:158-162; Gomes & Mendes, 2009:2547-2554) which is also utilised in the theoretical background of this study.

The researcher experienced that if traditional MCQs without formative feedback are used, the results may not necessarily be a true reflection of the students' understanding, since partial credit cannot be granted as is the case with CRQs. Consequently, only a few lecturers actually use MCQs. According to previous success and pass rate statistics in the courses at TUT, the average marks of students showed that they could get higher marks in CRQs compared to when the same content was assessed by means of MCQs without formative feedback and the granting of partial credit. The researcher regards this

as unfair assessment practice. Accordingly, there is a need to design quality MCQs with regard to the assessment of logical and consolidated assessment components (Engelbrecht & Harding, 2003:57-65; Huntley et al., 2011:3-16) and to build in IFF and the granting of partial credit when assessing calculations.

This background and these personal observations provided the motivation for this research to investigate the possibility of designing online MCQs with IFF and the granting of partial credit during the assessment process. This is posited as important as it provides students multiple opportunities to find the correct answer after utilising the feedback.

### **1.3 PROBLEM STATEMENT AND PURPOSE STATEMENT**

E-learning is a popular trend in teaching and learning in higher education (Kearns, 2012:43). Although the utilisation of online MCQs is a recognised assessment tool in e-learning, it may be considered unsuitable for use in assessing problem-solving calculations in science as it is uncommon to grant partial credit when using it (Engelbrecht & Harding, 2004:217-231). The post assessment feedback that students receive on MCQs can therefore only indicate if an answer is correct or incorrect, and the assessment resultantly has little formative value (Siddiqui, Bhayser, Bayser & Bose, 2016:114-121). When IFF is incorporated into the process of answering MCQs, this can, however, have a formative function. Guessing will always be part of the use of MCQs, but it may be used less if IFF during the assessment process and the granting of partial credit are included in the design. Students get assistance with the solving of the problems, and they can earn partial marks for the correct answer even if it is their second or third choice after correcting their mistakes.

The purpose of this study was to determine whether online MCQs with IFF and the granting of partial credit during the assessment process, utilised in assessing problem-solving calculations, will result in equal or better student performance in comparison with online MCQs without IFF or in comparison with CRQs.

## 1.4 RESEARCH QUESTIONS AND SUB-QUESTIONS

The research question formulated to address the above-stated purpose of this study is:

***Can online multiple-choice questions, designed to provide immediate formative feedback with the granting of partial credit, be used as an effective alternative for constructed written questions in the assessment of problem-solving calculations for Engineering Science students?***

### **The research sub-questions**

The following sub-questions were formulated to answer the main research question:

- When setting online MCQs, how can an assessor determine which distractors would anticipate the mistakes that students could make so that formative feedback can be built in to indicate to them what they did wrong and how to correct it?
- Can online MCQs without IFF replace CRQs in fairly assessing students' ability to do problem-solving calculations in science?
- Are online MCQs with IFF, and the granting of partial credit, more reliable and fair than MCQs without IFF in assessing students' ability to do problem-solving calculations in science?
- Can online MCQs with IFF and the granting of partial credit be used to assess problem-solving calculations as a fair and valid alternative to CRQs?

## 1.5 THE AIM AND OBJECTIVES OF THIS STUDY

The aim of this study was to contribute to the current pool of knowledge regarding the use of online MCQs to assess problem-solving calculations with IFF and the granting of partial credit.

### **Specific objectives**

The following objectives were formulated to assist with the achievement of this study's aim, namely, to:

- Determine the general mistakes that students make when completing constructed written questions in the assessment of problem-solving calculations. The common mistakes made by students need to be identified and used as distractors in the

design of the online MCQs, and in designing relevant formative feedback for each MCQ.

- Determine the effectiveness of using online MCQs without IFF compared to the use of CRQs in the assessment of problem-solving calculations.
- Investigate the level of reliability of online MCQs with IFF and the granting of partial credit compared to the reliability of online MCQs without IFF.
- Determine the effectiveness of using online MCQs with IFF and the granting of partial credit compared to the use of CRQs.

## **1.6 RESEARCH DESIGN**

Bless, Higon-Smith and Kagee (2007:71) describe a research design as the description of acceptable processes to be achieved in order to test comprehensive hypotheses under specified circumstances. A research design is the strategy of how to define the nature of the relationship between the different variables (Bless & Higon-Smith, 1995:46). The researcher decided to use a quantitative design because it was determined that an experiment was needed to find the answers to the sub-questions, and therefore the research question. A research method involves data collection and analysis methods and the interpretation of data. Purposive convenience sampling was used to select an experimental group and the statistical data was obtained from the experiment (Chadwick, et al., 1984). This will be discussed in more detail in Chapter 3.

### **1.6.1 Research paradigm**

This research was approached from a positivist point of view as it was based on measurements from an experiment, which were used to test the hypotheses.

### **1.6.2 Research approach**

A quantitative research approach was followed in this study. It was determined that an experiment would be needed to obtain the data to answer the research question regarding the comparison of the effectiveness of different types of assessment.

### **1.6.3 Research type**

Pre-experimental research was used to investigate the effectiveness of MCQs with IFF and the granting of partial credit in comparison to other types of assessment. This was

the case as pre- and post-tests were needed to determine the effectiveness of each type of assessment.

#### **1.6.4 Research methods**

Research methods explain the ways in which the research was undertaken and includes the sampling process and the way the data are collected, analysed and interpreted for the purpose of this study.

##### *1.6.4.1 Sampling*

The Engineering and Built Environment extended course at TUT was used as the population of interest in this research. The course comprises 160 first-year students, none of whom met the academic requirements to enrol for the equivalent diploma course. The students were randomly assigned to one of four groups, which consisted of 40 students each. Purposeful convenience sampling was used to select one of the four groups as the researcher had to have access to the group of students who would be involved in the experiment. This specific group of 40 students furthermore agreed to take part in the voluntary experiment for which they did not receive any credit. This then represented the experimental group.

##### *1.6.4.2 Data collection*

A baseline assessment was performed to identify the common mistakes made by the students when doing problem-solving calculations. This was done previously to use the common errors to design the distractors for the online MCQs. This meant that IFF could be built in to help students to correct their answers when they chose the wrong distractor. This baseline assessment was solely done to identify common mistakes.

As part of the experiment, a CRQ-assessment (pre-test) was carried out in a familiar venue during normal lecturing times to make the circumstances as natural as possible for students. Four calculation questions on the topic of Work and Energy, as required in the syllabus, were given to the 40 Engineering Science first-year students to complete in a 90-minute period. According to the ethical clearance requirements, Brenda Ngoma fulfilled the role of an independent assistant during the assessments (see Appendix A for the consent form).

She gave every student a random number between 700M and 790M, which served as a confidential password to access the online MCQ-assessments.

**The data instrument** used in this study was one experiment consisting of the following three assessments. The students from the experimental group answered the questions for the pre-test in writing on a paper with a carbon copy paper attached (**Assessment 1**). The original paper was handed in for marking by the examiner with partial credit given, and the students used the duplicate as reference when completing tests M and N (first and second post-tests). The students completed test M (post-test 1), consisting of the same set of questions as in the pre-test but now in an online MCQ-format without IFF (**Assessment 2**). They then completed test N (post-test 2), which was done online and was designed by adding IFF and the granting of partial credit to test M (**Assessment 3**). This was required to allow the students who chose the incorrect alternative to make further choices after applying the corrective feedback that they received after making the wrong choice.

A special program was written for the purpose of carrying out the experiment from which the data were collected. The **independent variable** - the problem-solving calculations (CRQs) – was manipulated by the researcher, while the **dependent variables** – online MCQs with IFF and the granting of partial credit, and online MCQs without IFF– was measured at the end of the experiment (Maree & Pietersen, 2011:147). The use of online MCQs with IFF and the granting of partial credit were chosen to determine if this was an effective alternative to CRQ problem-solving calculations.

Figure 1.1 is a presentation of how the experiment and its associated pre- and post-tests were set up. The CRQs are presented as the pre-test. The post-test 1 is then the online MCQs with the same questions. The common mistakes from the baseline assessment were then given as distractors. Post-test 2 then assessed online MCQs where immediate feedback was given.

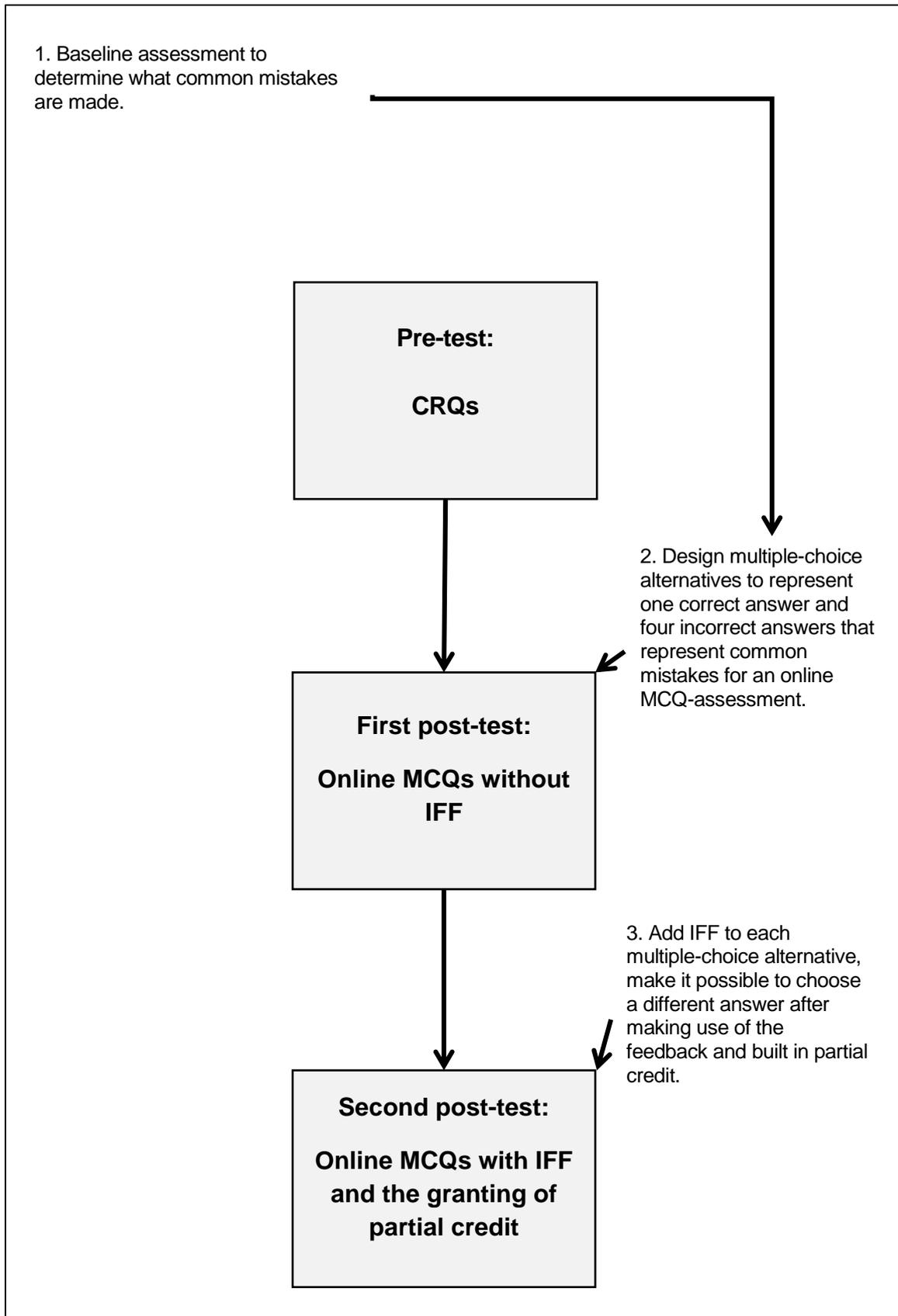


Figure 1.1: Graphical presentation of the research process

#### 1.6.4.3 Data analysis

The following three hypotheses were postulated to be tested:

##### **First hypothesis**

**H<sub>0</sub>:** There is no significant difference between the average mark for CRQs and online MCQs without IFF.

**H<sub>1</sub>:** The average of CRQs is significantly higher than the average of online MCQs without IFF.

##### **Second hypothesis**

**H<sub>0</sub>:** There is no significant difference between the average for online MCQs with IFF and the granting of partial credit and online MCQs without IFF.

**H<sub>1</sub>:** The average of online MCQs with IFF and the granting of partial credit are significantly higher than the average of online MCQs without IFF.

##### **Third Hypothesis**

**H<sub>0</sub>:** There is no significant difference between the average for CRQs and online MCQs with IFF and the granting of partial credit.

**H<sub>1</sub>:** The average of CRQs is significantly different to that of online MCQs with IFF and the granting of partial credit.

The dependent two-sample t-test was used to decide if the null hypothesis should be rejected or not. This test was used because the average marks obtained by the experimental group for the same set of questions, which were assessed in two different ways, were compared. Finally, a conclusion must be reached that tests the acceptability of the state of the hypotheses (Maree & Pietersen, 2011:203) (see Section 4.4).

## **1.7 TRUSTWORTHINESS**

Lincoln and Guba (1985) refer to the term *trustworthiness* as the way in which a researcher is able to proclaim his/her findings to be meaningful and that the research is of a high standard.

### **1.7.1 Validity**

Validity means that a quantity or instrument measures what it is supposed to measure (Maree and Pietersen, 2011:147). The way to measure validity depends on the ability of a number of professionals to obtain the same answer for a specific problem (Haladyna, 2011:183-202).

#### ***Internal validity***

Internal validity refers to the factors that might have an influence on the relationship between the dependent and independent variables. All three assessments were done on the same questions, and it was impossible for anyone to get the questions from somebody else. This implies that no external communication could influence the students' marks. Therefore, no student could be positively or negatively influenced. The students were not forced to take part in this research. An independent assistant was present to observe that all assessment requirements of the university were met. This study could therefore be considered to be internally valid.

#### ***External validity***

The external validity in this research was ensured by the following:

The researcher could not favour any individual student. All students of the specific group were part of the sample group, irrespective of their characteristics or level of knowledge. In this study, the difficulty levels of the chosen questions were selected carefully.

Unfortunately, it was not possible to make provision for every calculation mistake that a student could possibly make, but provision could be made for the the most common mistakes that students could make because of a lack of subject knowledge or misunderstanding. This research tested a specific part of the science syllabus.

### **1.7.2 Reliability**

Reliability measures how the results will fluctuate when an identical test is repeated (Williams, 2006:283; Palmquist, 2016:1). It therefore has to do with the stability of, or to what extent the instrument is repeatable (Maree & Pietersen, 2011:215).

The software program, written for the purpose of this study, can be used for other groups and for other problem-solving calculations, even those in mathematics. This meets the requirement for repeatability of the instrument.

### **1.7.3 Objectivity**

The marking of online MCQs cannot be influenced by subjectivity. The more questions there are in a MCQ-assessment, the more reliable the results would be. The current study could have been more accurate if more questions with different levels of difficulty could have been used. The variables are, however, objectively measured and the researcher remained distanced from the students to draw unbiased conclusions.

### **1.7.4 Ethical considerations**

Professional bodies, whose members are involved in research, have created codes and ethics that researchers must respect (Babbie, 2007:27-71). The anonymous information that was used for the research purposes are not of any academic, study or disciplinary value. The data will be stored for five years as hard copies in a locked cupboard and soft copies will be stored on a computer with password protection.

See Appendix A for the informed consent that all 40 of the participating students signed.

There was no pressure on any student to take part in this study. In effect, they could change their minds during any stage of this study without any consequences. Ethical clearance for the research was granted by both TUT and UNISA (see Appendix B).

## **1.8 CHAPTER DIVISION**

This dissertation is comprised of five chapters, which are briefly described below:

### **Chapter 1**

The overview and orientation in this chapter provides a brief summary of this study, including a brief discussion of this study's rationale, the literature study, and an overview of the research design and procedures.

### **Chapter 2**

The theoretical framework guiding this study is presented. In this chapter, the key constructs are further discussed and unpacked.

### **Chapter 3**

The research design is motivated and explained in this chapter.

### **Chapter 4**

The data analysis and interpretation thereof are presented and the results from the statistical analysis and the findings are discussed.

### **Chapter 5**

This study is summarised and conclusions and recommendations are presented.

## 1.9 SUMMARY

In this chapter, an overview was given of the research on online formative assessment with feedback and the granting of partial credit for online MCQs to assess problem-solving calculations in science. Currently, students' knowledge of science is traditionally assessed by means of constructed written questions (CRQs) in the assessment of problem-solving calculations.

By involving intelligent e-learning systems, and allowing students in higher education to make use of online assessments, e-learning can be made accessible with increasing convenience and efficiency. The MCQ-assessment method is popular for online assessment. The current challenge is that MCQs are commonly used only to assess theoretical knowledge, whereas there is a need to assess problem-solving calculations as well. A gap in knowledge was thus identified and it was decided to carry out research to determine how MCQs can be designed to effectively assess higher cognitive levels, specifically in assessing problem-solving questions.

The research design was discussed, research methods were explained and a chapter division was presented.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 INTRODUCTION

In this chapter, a review of the relevant literature on the topic under study is undertaken. The chapter starts with the literature relevant to the contextual and conceptual framework of this study, and ends with a theoretical framework derived from the overview of the literature.

### 2.1.1 Definition of terms

The definitions of the terms used in this study are given below:

**Assessment component taxonomy** is a method that was developed to identify alternative assessment methods like online MCQs.

**Blackboard system** is a computer-generated forum where students and lecturers communicate. This is especially used by universities as a platform for educational events, and as management for course activities (Heirdsfield, 2011:1-16).

**Constructed Written Questions, also called Constructed Response Questions (CRQs)** are questions where students are required to create their own answers instead of choosing the correct one from a list of pre-setup alternatives (Kuechler & Simkin, 2003:389).

**Feedback interventions (FIs)** are among the most widely functional psychological interventions where students are given information regarding their task performance.

**Formative assessment** intends to promote students' learning by giving valuable supportive feedback on their performance and suggestions to better their performance. It refers to assessment *for* learning (Stiggins, 2002:758-767; Black & Wiliam, 2003:623-637).

Black and Wiliam (1998:7-74) and Toplis (2015:267) describe, in broader terms, that any feedback that can be used to modify teaching or learning processes should be seen as formative assessment.

**Immediate feedback** is when feedback is given directly as the questions are answered (e.g. when an answer is chosen when doing online MCQs) rather than doing it at a later stage. It helps to focus attention on common student mistakes and how these can be corrected (Higgins & Tatham, 2003:1-11).

**Immediate formative assessment techniques (IFAT)** is a technique where an IFAT® form is used to scratch the correct MCQ-option. IFAT® form is a paper that is covered with a thin opaque film. The correct answer is indicated by a star or a blank for an incorrect answer. If it is incorrect, another opportunity is given. Partial credit will then be given if the correct answer is chosen (Maier, Wolf & Rander, 2016:85-98).

**Item Response Theory (IRT)** is a theory where a mathematical function is used to give the probability that a student with a specific ability will answer a question correctly.

**Moodle** stands for ‘Modular Object-Oriented Dynamic Learning Environment’ and is an online-system used to manage a course (Brandl, 2005:16-23).

**Multiple-Choice Questions (MCQs)** consists of a problem with a few possible answers (alternatives). The problem is called the stem, and the incorrect possibilities are called distractors. There is a correct choice, which represents the correct answer to the question (Brame, 2014).

**Problem-solving calculations** represent a mathematical method of solving word problems.

**Respondus** is a tool used to create exams that can be printed or published to the Blackboard platform (Poutre, Hedlund & Nau, 2015).

## **2.2 CONTEXTUAL AND CONCEPTUAL FRAMEWORK**

A conceptual framework is an analytical tool or an organising image of the phenomenon to be investigated. “It determines which questions are to be answered by the research and how empirical procedures are to be used as tools in finding answers to these questions” (De Vos, Strydom, 2011:35). Assessment and MCQs as a form of formative assessment are discussed.

### 2.2.1 Assessment

Assessment is considered to be the most important driving factor of student learning because students usually focus on the material to be assessed (Htwe, Ismail & Low 2014:502). Although students need to be graded through assessment, they must also be encouraged to learn from it (Gretton & Challis, 1999).

Assessment can be of a summative or formative nature (Black & Wiliam, 2003:623-637). Summative assessment is described as the assessment *of* learning while formative assessment refers to assessment *for* learning (Stiggins, 2002:758-767; Black & Wiliam, 2003:623-637). The intention behind the assessment determines the assessment method. Assessment *of* learning is more appropriate if the intention is to grade students for a test. However, if the intention is to improve teaching, then assessment *for* learning is more suitable (Wiliam, 2006:283-289). This study explored the formative nature of assessment. Two of the types of formative assessment are CRQs and MCQs (Gogri et al., 2013:89-93). MCQs are considered in more detail due to the nature of this study.

In order for an assessment activity to enhance learning, it needs to provide information that lecturers and their students can use. This information is not only used to assess themselves and one another, but to provide feedback to improve both teaching and learning activities. Assessment is only considered formative once the feedback is used to adjust teaching to satisfy learning needs (Wiliam, 2014).

In their study, Engelbrecht and Harding (2004:217-231) found that students have a greater interest in the summative component of assessment. They argue that a student benefits effectively from formative assessment only if he/she makes mistakes. Formative assessment does not provide additional information to a student with full marks. The feedback given to well-performing students leads mostly to peace of mind and increased confidence. Therefore, the true value of formative assessment is only visible to a student when he/she loses marks in a test and gets feedback on how to improve this result (Engelbrecht & Harding, 2004:217-231). This is of significant value for this study because feedback could make the difference for a student who has lost marks. Feedback will then help to improve their marks via IFF and the granting of partial credit.

### 2.2.1.1 *Formative assessment*

Scriven (1967:41) was the first to describe an evaluation process that has “a role in the on-going improvement of the curriculum”, as being formative. Shute (2008) describes formative assessment as information that is presented to the student for the purpose of improving learning by changing his/her thinking pattern or behaviour.

Black and Wiliam (2004:20-50) did pioneer work with regard to formative assessment. They view assessment that is designed and used to predominately improve student learning as assessment *for* learning. They state that formative assessment differs from assessment where accountability is the main focus. The assessment becomes formative when the information is applied to improve teaching and learning (Black, Harrison, Lee, Marshall & William, 2004:8-21). Black and Wiliam later rephrased their initial definition of formative assessment.

They proposed that:

*An assessment functions formatively to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (William, 2014:10).*

When lecturers know what their students have achieved after each chapter or section of the work, it allows them to determine if the students are at the level that they are supposed to be. This enables them to equip students with knowledge that they might still need (Buchanan, 2000:193-200; Coyne, 2015:24-26). Social constructivists describe the concept of a zone of proximal development as the difference between the current and the potential level of understanding that a student can accomplish with the lecturer’s guidance (Vygotsky, 1978).

It takes a positive culture of assessment to be able to improve learning for both lecturers and their students. This culture requires development over a period of time with the commitment of leadership for lecturers and students. These parties must work together in developing goals, setting assessment criteria, and providing manual feedback that is useful for both parties (Wiliam, 2014). In this study, it is proposed that lecturers must

work together to set up a proper database for common mistakes that students may possibly make. These can then be used as distractors and accordingly used for feedback. Formative assessment appears to be an increasing trend in education research and practice, while researchers such as Black, Wiliam and Heritage (2004:8-21) continue to advance its theoretical framework. Black and Wiliam (1998:7-74) support the positive impact of formative assessment on a student's learning at a diversity of cognitive levels.

In this research, the impact of immediate formative assessment may be of significant value to a student's learning at different cognitive levels.

#### *2.2.1.2 Formative assessment with feedback*

Toplis (2015:267) considers any feedback that can be used to modify the teaching or learning processes as formative in nature. There are, however, certain criteria that formative assessment needs to adhere to for it to be useful (Brosvic, Epstein & Cook, 2006:205-218). Students need to have enough time left to incorporate the feedback into their learning process. William (2011:3-14) states that merely telling students that they need to improve is also not useful. Valuable feedback focuses on student learning and what they need to improve on and how they should do this.

Boud (2000:151-167) and Boud and Soler (2016:400-413) argue that the effectiveness of feedback can only be measured when students are allowed to apply the given feedback to repeat the same assignment. Feedback, therefore, has been accepted to be an effective tool for student learning (Mory, 2003:745-783; Hattie & Timperley, 2007:81-112; Shute, 2008:153-189). Hattie and Timperley (2007:81-112) also address the positive influence of feedback on learning and constructed a model of feedback to enhance learning. Four types of feedback, together with factors illustrating their efficiency, are identified in this model. The first type indicates if the answer is correct or incorrect. Secondly the feedback process explains the processes of the task. Self-regulation feedback that can help students with plans to improve their own work is stated to be the third type and self –feedback is considered to be the fourth type of feedback.

Their research revealed that students may even benefit from self-feedback by increased commitment, determination and self-confidence (Hattie & Timperley, 2007:81-112). Task

specific feedback is, however, the most popular type of feedback used in the classroom (Black & Wiliam, 1998:7-74; Hattie & Timperley, 2007:81-112).

Hyland and Booth (2000:234) state that formative feedback has the potential to change each question into an instrument for further development. Hattie (1987:197-212) concurs and adds that feedback is the most dominant single-influence on student learning. Black and Wiliam (1998:7-74) confirm that formative assessment has a significantly positive influence on students' learning, motivation and success. In this study, the statement that formative feedback may change each question into an instrument for further learning is of significance because the proposed assessment method could provide feedback to improve students' knowledge.

Hattie (1987:197-212) indicates three conditions that should be met for formative assessment to support the learning process. It firstly needs to determine what the current position of the student is. It should, secondly, convey to the learner what is needed in order to be successful. Thirdly, learners require guidelines on how they can improve.

In this study, the current position of the student is determined after the first attempt to answer a MCQ. The immediate feedback given will guide the student towards the correct answer. The students can then learn from their mistakes to improve their knowledge. Wiliam and Thompson (2007:53-82) identify five ways in which formative assessment can be successful. It can describe and share learning purposes and principles for success. It can initiate efficient discussions and several learning tasks in the classroom to help students to understand the work properly. Formative assessment can provide feedback to improve students' learning. It motivates students to assist each other, and stimulates them to take responsibility for their own learning. Lastly, the data from the assessment can be used to determine student performance and identify what part of the curriculum needs attention if the mistakes are common to a group (Maier et al., 2016:85-98).

Nicol and Macfarlane-Dick (2005) identify the principles of good feedback as follows: it can explain expected standards, aims and criteria as good performance. The progress of thinking and self-learning is assisted. Feedback gives high quality evidence of the student's learning. It inspires discussion about the assessment amongst lecturers and peers. Helpful motivational principles and self-confidence are encouraged. Feedback

offers chances to close the gap between present and anticipated presentations. It gives information to lecturers that can help to form their teaching. Moreover, not only is the extent of feedback important, but also the timing of when this feedback is presented (Toplis, 2015:158).

Therefore, the provision of feedback is significantly important in supporting learning (Gipps & Simpson, 2005:3-31) as it is also proposed in this study that immediate feedback should be used to assist students to find the correct solution.

#### *2.2.1.3 Formative assessment with immediate feedback*

An important variable of feedback is timing (Hattie & Timperley, 2007:81-112). These authors explain that the timing of feedback is related to the content. In their opinion, immediate feedback is most effective when a common error is immediately corrected. Shute (2008:153-189) maintains that high achievers benefit the most from delayed feedback and lower achievers from immediate feedback. In contrast, Brosvic et al. (2006:205-218) note that the sooner feedback is given, the more students can benefit from it. Toplis (2015:267) agrees that immediate feedback is more effective in learning than delayed feedback.

Gogri et al. (2013:89-93) state that feedback reflects the problem areas of students' knowledge and the need for improvement of their learning. Immediate feedback gives information to show students what they did wrong and how to reach their learning goals. Gogri et al. (2013:89-93) explicate that students understand and remember the feedback the best when it is immediately given. Almost 79% of the students who participated in their study agreed that feedback must be given immediately because each phase of the learning process builds up from the previous phase. If no feedback is given, the next phase may be built on misconceptions. Taking this statement into consideration, the current study proposes that the use of IFF and the giving of partial credit can help prevent misconceptions.

#### *2.2.1.4 Formative assessment with immediate online feedback*

Formative assessment can be presented in different ways. Online assessment is becoming a leading medium of assessment, probably because the flexible nature of online assessments could be beneficial to students (Morton, Xu & Joplin, 2012:806).

Online assessments are flexible since students can undertake the assessment in their own time, easily repeat it, and receive feedback on their performance immediately (Buchanan, 2000:193-200).

The movement in favour of the need to give immediate feedback raises the question of how universities will present feedback in a resource-efficient manner (Gogri et al., 2013:89-93). Lopez-Pastor, Pintor, Muros and Webb (2012:163-180) confirm that to move from the traditional way of assessment toward online assessment will initially be time-consuming until a database of questions is set up. Thereafter it will be time-efficient.

Gibbs and Simpson (2005:3-31) confirm that the frequency of feedback and its immediacy, made possible by online MCQs, contributes to the overall quality of learning and therefore has a direct impact on student performance. A number of researchers agree that immediate formative online feedback has a positive impact on students' inspiration and commitment (De Lange, Suwardy & Mavondo, 2003:1-14). Shute (2008:153-189) goes even further to say that it is best to give feedback not only immediately after the assessment procedure, but during it. This finding is of the utmost significance in this current study as it is proposed that immediate feedback given during the answering of a specific online multiple-choice assessment question does not only tell a student that an answer is wrong, but indicates what mistake was made and what can be done to correct the mistake. Most important of all - it provides an opportunity for the student to immediately make use of this information to find the correct answer, thereby also allowing the assessor to allocate partial credit for the corrected answer.

### **2.2.2 Methods of formative assessment**

Formative assessment can take on many different forms. MCQs and CRQs are classified as two of the methods of formative assessment (Gogri et al., 2013:89-93).

The methods of MCQs can be classified as follows:

- True/False-type assessments, which are relatively easy to formulate. Unfortunately, students have a 50% chance to guess right and these can therefore not be used to test knowledge properly.

- The student needs to state how many of the possible five answers are correct. This suggests that there can be more than one correct answer. In this way, a student can be penalised although he/she has 80% of the knowledge.
- Match column A with column B. The probability of guessing is less, but for the examiner there is more than one possible correct answer to be checked.
- Elimination is also an option to use where guessing plays a lesser role.
- The type of MCQ that is most often used is to choose the correct answer out of five possible alternatives (Benvenuti & Cohen 2008:21; Gogri et al., 2013:89-93).

The opposite of MCQs is described as constructed written questions (CRQs). CRQs mean that students are required to create their own answers instead of choosing the correct one from a list of pre-set-up alternatives (Kuechler & Simkin, 2003:389). They define oral examinations, answers in short exercises, essays and self-tests to be different methods of CRQs.

Question papers can consist of MCQs and CRQs. It cannot be fair to use only the results of the MCQ section or only the results of the CRQ section of the paper as it may give the wrong indication of the student's knowledge (Webb, 1989:216-218). Jonanovic and Aburaghiem (2006:233-252) proved that MCQs can test knowledge remembrance, but a balanced paper should include MCQs and CRQs to create a multiple strategy assessment. Siddiqui et al. (2016:114-121) agrees that MCQs can be combined with CRQs when setting a paper to provide a more reliable assessment. Lukhele, Thissen and Wainer (1994:234-250) studied assessments that included both MCQs and CRQs and found that a single essay gives the same amount of information as four to eight MCQs. Testing 16 MCQs in 75 minutes was as reliable as a three-hour CRQ-assessment.

Engelbrecht and Harding (2003:57-65) find that MCQs can effectively be used to assess students' achievement of mathematical insight and knowledge:

- MCQs and CRQs can both be used successfully to assess undergraduate mathematics.
- They decided to make use of MCQs and CRQs in exam papers where partial credit would be given for correct aspects of an answer although the final answer might be wrong. This is perceived to be one of the serious shortcomings in online assessment (which this study has challenged).

- Although short answer questions are the only CRQs that can be used online, all different formats of MCQs are useful in online courses.
- Another finding from their study was that students performed better in paper-based CRQs<sup>3</sup> than in online CRQs<sup>4</sup> due to:
  - The absence of partial credit in online CRQs.
  - The fact that the fear factor associated with the assessment of paper CRQs may cause anxiety in written exam conditions.
  - Students are not forced to write down the logical steps when solving the problem using online CRQs. This leads to an unnecessary loss of marks.

Finally Engelbrecht and Harding (2003:57-65), concluded that students perform better in paper CRQs than in online CRQs, and better in online MCQs than in online CRQs. The outcome of a study done by Engelbrecht and Harding (2003) where the use of MCQs with IFF and the granting of partial credit was explored and found that MCQs were an advantage, is of great significance to this study.

Hift (2014:249) suggests that careful consideration is necessary before CRQs can be entirely replaced by MCQs. He insists that although MCQs are being used, we clearly do not understand the cognitive design of them that well. MCQs as an assessment format performs excellently for knowledge-based and theoretical questions in contrast with CRQs. If MCQs are well-constructed and rich in context, they can be valid and capable of assessing higher order cognitive skills (Huntley et al., 2010:141-171). Kastner and Stangl (2011:263-273) add that CRQs can be better assessed compared to MCQs with multiple correct responses, which is where more than one alternative can be correct and no negative marking is used. Siddiqui et al. (2016:114-121) indicate that the challenges in setting up proper distractors can be solved by using multiple correct responses.

Engelbrecht and Harding (2003:57-65) found that when MCQs are used, students memorise only bits of information. They are therefore not encouraged to understand the topic overall. Huntley, Engelbrecht and Harding (2011:3-16) propose the

---

<sup>3</sup> Paper CRQs refer to when an assessment is done in the traditional way, on paper, where the lecturer marks it with partial credit given.

<sup>4</sup> Online CRQs are when an assessment is done on the computer without any possibility of choosing the correct option, like using MCQs without partial credit.

assessment component taxonomy, which is part of the theoretical background of this study (Section 2.3.2). Beyond expectation, one of the findings was that MCQs were a slightly better option than CRQs in the problem-solving component as an assessment method. For the problem-solving component, a mathematical method to solve word problems is essential. The way in which Huntley et al. (2011:3-16) approached the problem-solving questions was to divide these into sub-questions concerning different cognitive skills. Marks were then allocated step-by-step to guide the students to the correct answer. This study proposes to guide the student step-by-step towards the answer, but the challenge is to use MCQs with IFF and with the granting of partial credit, and not to divide the problem-solving calculation into sub-questions. Huntley et al. (2011:3-16) reveal that special skills are required, especially when MCQs are designed for logical and consolidation assessment components. They recommend that assessments must therefore include both MCQs and CRQs to balance a paper. This indicates that MCQs can be used to test cognitive skills at higher cognitive levels too (Williams, 2006:283-289; Haladyna, 2011).

### **2.2.3 Multiple-choice questions**

A multiple-choice item consists of a problem with alternative answers. The problem is referred to as the stem, while the alternatives exist of the correct answer and the distractors (Brame, 2014). MCQs have been successfully used as a form of assessment for more than 100 years (Nazeem, Vinayak, Bhayser & Sukhwant, 2016:114-121). MCQs are increasingly used as a supplement or even a replacement for other assessments (Nicol & Dick, 2007:199-218) and can be very useful for students who need to self-study. Often, students have to study out of class and take responsibility for their own learning due to limited contact time in class. However, it still remains a challenge for staff to motivate students to work on their own consistently (Einig, 2013:847-848). Using regular online MCQ-assessments may contribute to alleviating this challenge.

It is quite popular for assessors to use MCQs in at least one portion of a paper, especially with large groups (Rollnick, Brenner & Moletsane, 2010:51-65). It might be more accurate to use MCQs in testing higher cognitive skills if challenges like guessing and issues with the marking scheme are addressed (Nazeem et al., 2016:114-121). Researchers agree that MCQs are useful for the memorisation of questions and the

recollection of factual knowledge. According to these authors, MCQs to assess higher cognitive levels are unfortunately not yet proven to be reliable and fair (Chiheb, Faizi & Afia, 2011:69-72; Ventouras, Triantis, Tsiakas & Stergiopoulos, 2011:616-624; William, Lee, Harrison & Black, 2010:49-65). These are challenges that are addressed in this study in order to try and make a positive contribution to this field.

In choosing one of five possible correct answers in an MCQ-assessment, students may be limited in their thinking. This could be the reason why students do not learn to organise their thoughts or create their own answers (Brame, 2014). William et al. (2010:49-65) argue that if MCQs are constructed poorly, insufficient attention may be given to the writing process, which could affect the validity and reliability of the assessment.

Einig (2013:847-848), however, indicated in her study that the use of MCQs could be a significant form of formative assessment. In relation to assessment at higher levels of Bloom's taxonomy (see Section 2.3.1) as the theory in the framework, Alford, Herbert and Frangenheim (2006:176-224) suggest that expertly constructed MCQs could potentially empower valid and reliable understanding. The use of MCQs could assist lecturers with their marking load as well. Limited feedback and telling the students if their answer is right or wrong are options. Unfortunately, the higher cognitive skills suggested by Bloom cannot be reliably assessed yet (Einig, 2013:847-848).

The University of South Wales (2014) has the opposite view, and states that MCQ-assessments can be designed to test higher order cognitive skills, such as problem solving, creative thinking, and synthesis. This is a significant view in terms of the objectives of this study, namely, to design and implement reliable and fair online MCQ-assessments for problem-solving calculations.

#### *2.2.3.1 Advantages of MCQs*

Morton et al. (2012:806) reveal that online MCQ-assessments are currently a popular method of assessment with several benefits such as automated marking, reliable scores, reduced marking time, and consistency between different assessors.

Al-Sadi and Al Halabi (2010:176-182) indicate that MCQ-assessments are an easy and effective way of assessing large groups. William, Lee, Harrison and Black (2010:49-65), as well as Malau-Aduli, Assenheimer, Choi-Lundberg and Zimitat (2013:510-522) agree with this statement. Large groups of students are becoming more of a reality, which requires a balance between the manageability, the validity, and reliability of assessments. Because of lecturing staffs' high workload, they can rely on MCQ-assessments with online feedback given to students to lessen their marking load. In a study done by Snowball (2014:823-838), he explained that group sizes expand more than most Universities can absorb. From studying 50 first year economic students, Snowball found that the majority of students agreed to replace one lecture per week with online resources, which improved learning. Twenty percent of the same group, however, preferred more lecturing and less time online. The findings further indicated that in-class lectures should never be neglected, although online MCQs were a valuable addition to improve the performance of students (Snowball, 2014:823-838).

William (2011:3-14) extends the original theory of Bloom by stating that the individualisation of learning is a crucial concept of assessment that has the potential to improve learning. To construct MCQs to assess higher levels of cognitive skills is a very demanding task and needs an investment in time to set up proper questions. When it comes to the marking of assessments, the return of this investment is, however, seen.

Pittenger and Lounsbery (2011) see student-generated MCQs as another advantage. To use MCQs as part of peer-assessment, students can see common errors made by their peers. This engages students with the content of the course and therefore enhances their cognitive skills.

Marx (1998) states that the use of MCQ-assessments can be to the benefit of students with limitations in the use of English. Students may choose the wrong answer due to a lack of English reading skills when using CRQ-assessments. It could be easier to recognise a correct option than to write it from scratch (William et al., 2010:49-65). Fellenz (2004:703-719 and Bush (2006:398-404) concur, explaining that students may not be penalised because of English not necessarily being their first language.

The benefits can be expressed as efficiency and objectivity in assessment.

## **Efficiency**

The efficiency of online MCQ-assessments lies in their productivity and cost-effectiveness (Hift, 2014:249). Kuechler and Simkin (2003:389) identified the most important advantage of MCQs, which also relate to this study. This saves time and can be done by huge numbers of students. When MCQs are done online, cheating can be controlled by randomising the questions. Einig (2013:847-848) finds that almost all students can use MCQs in their learning. Statistical investigation shows a significant relationship between the use of MCQs and learner performance (Einig, 2013:847-848).

Al-Sadi et al. (2010:176-182) state that giving immediate feedback to online MCQs is effective, while Siddiqui et al. (2016:114-121) add that such feedback creates an opportunity for self-assessment. Kutluay (2005) explains the efficiency of online MCQs to be administrated easily because students can see their marks immediately after answering the questions. Online MCQs take less time to complete, therefore more of the syllabus can be assessed in the same time as CRQs with less questions. MCQs can therefore test a wider spectrum of the curriculum (Siddiqui et al., 2016:114-121). Pittenger and Lounsbery (2011) indicate that student-generated MCQs can effectively be used as assessments and therefore MCQs can be an integral part of teaching and learning.

CRQs normally take much more time to complete and the marking of the scripts can also be very time-consuming. CRQ-assessments require four to 40 times as long to manage as MCQs, and with the same reliability (Wainer & Thissen, 1993:103-118). Lukele, Wainer and Thissen (1994:234-250) report that the cost of marking CRQs was 300 times that of marking MCQs. Benvenuti (2010:1) showed that if MCQs are carefully designed and well-structured, misconceptions can be eliminated, higher levels of cognitive skills can be assessed and there can be discernment between strong and weak performing students. This is a significant finding in terms of this study as it is a long-standing critique that MCQ-assessments can only be used for assessing lower cognitive levels such as the recall of knowledge.

## **Objectivity**

Schuwirth and Van Der Vleuten (2004:974-979) state that MCQs are objective and reduce the potential partiality of the examiner. Higgins and Tatham (2003:1-11) agree

that there is no subjectivity from the assessors when marking MCQs, which makes MCQs fair assessments (Nazeem et al., 2016:114-121).

Siddiqui et al. (2016:114-121) state that MCQs are easy to grade but that assessors must ensure valid and reliable results by designing questions that are fair and of high quality. In mass lecturing, the greatest advantage of MCQ-assessments is that they are marked by a computer (Einig, 2013:847-848). MCQs can therefore be set up in a database that can be used multiple times (Al-Sadi et al., 2010:176-182). A database of MCQs can be very useful to support student learning as students can access a database of questions for revision purposes and redo them until they understand the work. Hift (2014:249) expresses that although the set-up of MCQs at first is very timeous, time will be saved in the long run when a proper database is set up by different participating lecturers. A platform like Blackboard can be used to build and expand a database that can be utilised repeatedly by lecturers and students.

Hughes and Quinn (2013) mention another objective advantage of MCQ-assessments, namely, that assessment can take place without the need to decode handwriting of students.

**As seen in this section**, there are many advantages of using MCQs. The literature indicates that MCQs are more effective when using them online and when lecturing large groups. The use of MCQs can reduce the partiality of assessors because it can be marked by computer. Although there are many advantages, there are also disadvantages that should be taken into take into consideration.

### *2.2.3.2 Disadvantages of MCQs*

Wenning and Vieyra (2015) emphasise that it is not easy to set up proper MCQs. Distractors must be chosen with great care and standard questions must be avoided. This links to low item difficulty and other disadvantages like guessing, structure, and language issues regarding MCQs.

### **Guessing and negative marking**

Guessing what the correct answer is when completing MCQ-assessments has been a concern since their inception. The theory of the three parameter logistic Item Response Theory Model (IRT) is used in this study (see Section 2.3.3). This is a

model, as part of the theoretical framework that challenges the reaction of a student to a question topic (Lord, Melvin, Novick & Birnbaum, 1968). A mathematical function is used to give the probability that a student with a specific ability will answer a question correctly. Hambleton, Swaminathan, and Rogers (1991) proved that the guessing of correct answers may be unfair to well-performing students who could score low marks, and poorer-performing students, who could score high marks because of the randomness of guessing. In an effort to overcome the guessing temptation, the two-tier test was developed (Maier et al., 2016:85-98). The first tier is like traditional MCQs and the second tier gives a set of reasons for the given answer in the first tier. This gives the students an opportunity to explain their answers and learn from their mistakes (Loh, Subramaniam & Tan, 2014:229-250).

In their study, Gay and Thomas (1993:130-134) found that a quarter of 199 seventh- and eighth-grade students could not explain the reasons for their choices for the MCQ-answer chosen. This indicates that the students simply guessed the answers. It was found in IRT that students make estimated guesses and do not just guess randomly (Barnard, 2013:172). The respondents in Dempster's (2007:47-60) study explained to him a range of plans that they had when choosing an alternative given for MCQs if they did not know the answer. Some students would start to eliminate possible alternatives that they did not recognise, or simply choose the answer with the same term as used in the question. Others looked at the pattern of choices from previous questions. If the pattern was, for example, A, A, B, B, then the 'obvious' guess for them would be to choose C for the next answer.

Engelbrecht and Harding (2003:57-65) also expressed their concern that there is a 50% possibility for students to guess the correct answer, referring to True/False questions. When the pass rate is 50%, as it is at most universities, it may be possible that students can pass irrespective of their knowledge or lack thereof. This would differ if MCQs were used where several alternatives are available for students to choose. In standard MCQs, where one out of five options could be correct, students could receive a 20% ( $\frac{1}{5} \times 100\%$ ) average if the answer is guessed correctly. To pass with 50%, a student effectively needs to know only 37.5% ( $\frac{50-20}{100-80} \times 100\% = \frac{30}{80} \times 100\%$ ) of the work, which is definitely not a standard pass mark. This could be even worse if there were more possible distractors. Rollnick, Brenner and Moletsane (2010:51-65) agree

with Engelbrecht and Harding (2003:57-65) that the possibility of guessing dilutes the effectiveness of MCQ-assessments. William et al. (2010:49-65) are also of the opinion that the guessing factor may reduce the validity and reliability of the assessment.

Higgins and Tatham (2003:1-11) have a different viewpoint of the guessing factor. They explain that for students to guess in MCQs is not worse than writing information for the sake of getting some marks in CRQs. However, this problem does not have a total solution yet, although possible solutions have been suggested such as to discourage guessing in MCQs by using negative marking, making use of more distractors, or raising the overall pass mark for the assessment to prevent a student passing by luck (Bush, 2006:398-404; Ventouras et al., 2011:616-624). Lesage, Valcke and Sabbe (2013:188-193) state that the selection of a correct alternative can sometimes point to guessing, but this temptation can be reduced by providing negative marking.

When negative marking is used and the wrong option is chosen, a quarter of the mark can be subtracted if there are four possible answers (Karandikar, 2010:1042-1045). Harper (2003:3-8) also suggests that negative marking can be a valid alternative.

Another option is to set different marks for more challenging questions or to give more alternatives for each question. Although Ali, Carr and Ruit (2016:1-14) agree, they also recognise the fact that it is challenging to create many suitable distractors. The use of negative marking is, however, not allowed at some institutions (Engelbrecht & Harding, 2003:57-65).

An option not suggested in the literature are online MCQ-assessments where there is direct and immediate formative feedback on each question as soon as a distractor is chosen. The students will then be allowed to choose a different answer based on the corrective feedback that they receive. If partial credit is given for finding the correct answer – although not with their first try – they are encouraged to try and find the correct answer without guessing. Marks are then used as encouragement rather than as punishment, and students can learn from the mistakes that they make. The current study explored this possibility.

## **Item difficulty**

According to Al-Sadi et al. (2010:176-182) the most important disadvantage of MCQs is that they cannot assess high level cognitive insights or skills. The opposite opinion is, however, expressed by Huntley et al. (2011:3-16). They maintain that it depends on the construction and the level of the test whether some higher cognitive levels can be tested, as stated in the assessment component theory from the theoretical framework. The nature of the questions in mathematics can be measured to ensure that the quality of student learning is tested at several levels. At the University of South Wales (2014), it is acknowledged that MCQs are currently inadequate at assessing creative thinking and problem-solving.

McLachlan and Whiten (2000:788-797) find that the outcome of an MCQ-assessment can only be representative if the level of difficulty is exactly the same. Rather than testing only the difficulty, their taxonomy gives tools to measure if tasks are completed successfully at the correct level.

To determine item difficulty, Kutluay (2005) explains that the ratio of students who answer the question correctly is the parameter that demonstrate the difficulty of the questions. These parameters could be from 0.00 to 1.00. If it is more than 0.50, this implies that most of the students answered correctly. The higher the value of the parameter, the easier the assessment is. Therefore, assessors must not use questions that are too easy or too difficult as this will compromise the reliability of the assessments.

Aubrecht, Gordon and Aubrect (1983:613) explain that the average mark of an answer needs to be about 65%. In the case where there are five alternative answers to choose from, an average of 20% or even 100% shows that it is useless for assessment purposes. Bishop (1990:83) supports Aubrect et al. (1983:613), and states that the success rate of a question – the fraction of the total students whose answers were correct – must approximately be 60% to 70%, which indicates that a question is good. If the success rate is less than 50%, the question was too difficult and requires more time and explanation. Bishop (1990:83) further investigated the frequency distribution of the distractors for each question; because this could expose sections of the work, it needs more attention. Questions with a success rate of 20% or 30% are seen as

unreasonable.

These findings may be of significant value in choosing valuable and valid distractors for the instrument used in this study.

### **Structure and language issues**

When one MCQ-question depends on a previous MCQ, both can be wrong if the first is misinterpreted. If MCQs are asked unconnectedly, the structure flow of the entire content could easily get mixed up, according to Stiggins (2002:758-767).

Moreover, a language issue identified by Al-Sadi et al. (2010:176-182) was that students' skills in writing and analysing cannot properly be stimulated when MCQs are used. The selection of the correct answer from a set of alternatives does not test their writing skills. In an online article from the University of South Wales (2014), it was mentioned that creative thinking is not stimulated when MCQs are used because MCQs are not open-ended as CRQs are.

### **The set-up of MCQs**

Nicol and Dick (2007:199-218) recognise that most MCQs provide limited feedback when the assessment is constructed because the alternatives are not necessarily considered with care when they must be chosen. Distractors had to be chosen from the most common mistakes that students could possibly make to assist the student. This is an easy way for assessors to remove random distractors that would not be of any help to the students (Nicol & Dick, 2007:199-218). When the distractors are chosen incorrectly, it can lead to bad assessment and will definitely not meet the required standard (Rollnick et al., 2010:51-65). This may not inspire effective learning at all and would therefore be classified as a disadvantage. The University of South Wales (2014) specifies that assessors need special skills and must be experts to be able to construct proper MCQs. The higher the cognitive skill to be assessed, the more challenging it becomes to design MCQs in the absence of a proper database.

The disadvantages stated above can be used constructively in this study. The guessing problem could be addressed if IFF and the granting of partial credit were to be included. When a proper database is set up, the level of difficulty can be

standardised, as well as the choice of distractors from common mistakes that students normally make.

### *2.2.3.3 Guidelines for developing MCQs*

Rollnick et al. (2010:51-65) have discovered that time and design skills are necessary to set up MCQs of a high standard. Pre-testing is also needed to ensure design accountability. The questions need to meet the expected cognitive standard and this can sometimes be challenging for academics who do not have the skill of item design. If academics are alert to their incompetence, they usually use existing databases (which may be far from ideal), while those who are unaware of their lack of ability may set items that do not necessarily meet the required standards, which is to the disadvantage of their students (Rollnick et al., 2010:51-65).

Designing proper MCQs is complex and challenging, for instance, distractors like 'none of the above' or 'all of the above' are inappropriate and should not be used. It is even more difficult for assessors whose first language is not English. The good design of items is considered to be even more important where MCQs are the only method of assessment used (although this is never good practice) (Rollnick et al., 2010:51-65). In the next section, guidelines to set up the stem and suitable distractors are discussed in detail.

#### **Guidelines to set up the stem**

Brame (2014) explains that an MCQ item consists of the problem with a correct answer, and a list of possible answers, called the distractors. Only one of the alternatives is the correct answer. The incorrect alternatives provided are called distractors. Brame (2014) provides some guidelines to set up proper MCQs. It is important that the stem defines a definite problem with emphasis on the outcome. Assessors must be careful not to give irrelevant information, which could cause a decrease in validity and reliability. None of the alternative options can be the same. Using negative statements in the stem is not good practice, especially if it is followed by a negative in one of the alternatives. The stem must be a question rather than a partial sentence to complete. If a student needs to complete a sentence, they will probably try to complete it with all of the possible answers, which could be time consuming.

## **Construct effective distractors**

Brame (2014) indicates that every alternative for an MCQ must be a reasonable possibility. Distractors must be so plausible that students who do not know their work properly will find it an acceptable choice to make. If any do not satisfy this criterion, they must rather not be included. Distractors must be set clearly and briefly with no more words than necessary.

To prevent unintentional hints being given to the students, all distractors must be related to the content, while clues such as differences in grammar, length or structuring can also lead students to the correct option. If 'all of the above' or 'none of the above', are used, students may choose the correct answer despite having limited knowledge. Distractors should be presented in a logical way to avoid preference for certain options, and terms such as 'all', 'only' or 'never' should be avoided (Brame, 2014).

William et al. (2010:49-65) add that assessors must be aware of the use of similar wording in the stem and the correct answer that could give clues regarding the correct answer. William et al. (2010:49-65) further identify that grammar and punctuation are as important as the choice of distractors, and therefore suggest some guidelines. If the stem is a question, the distractor must start with a capital letter. If the stem needs to be completed, the distractor must start with a lower-case letter. When a distractor is numerical, the use of a decimal comma instead of a full stop is advised to avoid a misunderstanding.

Berry and Chew (2008:305-312), and William et al. (2010:49-65) report that the answers of students in previous assessments can be used as realistic distractors in future MCQs as they were plausible enough to be chosen. This is an important principle for design and in this study, the incorrect answers of the students in the baseline assessment were used as distractors for the post-test. Siddiqui et al. (2016:114-121) concur that distractors must be believable and that they can comprise common errors made by previous students. This is a valuable guideline as assessors can use their own experience of common mistakes from previous assessments to use as distractors. It may then, unfortunately, happen that a student will most likely find a correct option between the well-chosen distractors. Although the answer is not correct, the student may feel satisfied because the answer of his/her calculated answer was

found to be one of the alternatives given as an option (Siddiqui et al., 2016:114-121). Benvenuti and Cohen (2010:1) state that feedback can then be of great value in learning.

#### **2.2.4 Online MCQs**

A need was observed by the researcher for MCQs to be used in assessing online problem-solving calculation questions. Guo, Palmer-Brown, Lee and Cai (2014:369-383) state that e-learning has a very important place in higher education. By using e-learning, higher education can become more accessible, effective and of a higher quality.

In considering the use of CRQs, it is shown that online technology will contribute to better learning. In a study by Friel and Johnstone (1978:717-719), 83 students were assessed after doing CRQ-online questions (when an assessment is done on a computer without any possibility of choosing the correct option, like using MCQs without partial credit) and CRQ paper-based questions (when an assessment is done in the traditional way on paper where the lecturer marks it and give partial credit). It was found that the students did better in the paper CRQs than online, even without partial credit. If partial credit had been considered, the difference would have been even greater. Friel and Johnstone (1978:717-719) found that the difference was too significant to ignore. Although the study was carried out in 1978, the outcome is still valid today (Engelbrect & Harding, 2003:57-65). It is better to teach students to do proper calculations and not to rely on partial credit. Therefore, there is a place for online CRQs where only the answer is considered (Engelbrect & Harding, 2003:57-65). This is how the baseline assessment of this research was developed.

Nicol and Dick (2007:199-218) explain that although the use of online MCQs is a useful way to assess knowledge, it lacks the validation of the answer. The advantages of online MCQs, according to Nicol and Dick (2007:199-218), are time efficiency, objectivity and quality that is guaranteed.

Escudier (2011:440-447) adds that the advantages for assessors can be that less time is required for marking. The marking is objective, quick response to the data is observed, and online assessment is convenient. Velan, Kumar, Dziegliwski and Wakefield (2002:282-284) express that although assessors do not need to do

marking, they still have admission to the performance of the students and can still comment on it. Students are therefore able to receive immediate feedback. Escudier (2011:440-447) and Hewson (2012:488-498) ranked the performance of students when using online MCQs to be similar to that of CRQs.

Kuechler and Simkin (2003:389) specify that one advantage of computer-based assessments is the capacity to capture the assessment results in an online format simultaneously with the completion of the assessment. All data are then available in an online format that is neither time-consuming nor loads of paper work to file. Blackboard software is an example of web software that can be used. Some programmes even give statistical analysis of the results, which can be very useful. Guo et al. (2014:369-383) developed a web based e-learning system that offers immediate feedback and online assignments that save time and provide stability in learning. The feedback given can be used endlessly and large numbers of students can benefit from this system.

### **2.2.5 Online MCQs with feedback**

Guo et al. (2014:369-383) state that good feedback should explain standards, goals and criteria in order to give proper feedback and should enable the progress of self-learning. Feedback should give proper evidence to the students to help with their learning and inspire discussion between the assessor and students. Feedback should motivate students' self-confidence through a positive response, giving opportunities to improve knowledge, and giving information to lecturers to invest in student learning. Maier et al. (2016:85-98) have identified the following categories of feedback:

- Simple or verification feedback – This is only to tell the student if the answer was right or wrong.
- Elaborated feedback – This is specific or expanded feedback.

There are three sub-categories of elaborate feedback:

1. Task-specific feedback is given when students know the correct answer, and implies positive feedback (like well done!).
2. Instruction-based feedback gives an explanation to help students in their understanding and learning.
3. Extra-instructional feedback explores extra information to add to the

feedback.

Maier et al. (2016:85-98) assert that if the feedback shows that the answer was correct, it increases the students' confidence but distracts them from further feedback given. However, if the feedback message is that the answer is still wrong, students are more motivated to absorb this detail. According to the authors, it is preferred to give rather simple than more detailed feedback. The Theory of Feedback Interventions (FIs), as used in the theoretical framework of this study, reveals that feedback can either disappoint or compliment a student (Kluger & DeNisi, 1998:67-72). To use FIs in combination with goal-setting, the attention will be directed to focus on the task and not the self. This will motivate students to increase their knowledge.

Alternatively, Higgins and Tatham (2003:1-11) find that the better the distractors are, the greater the ability of the assessor to provide valuable feedback. Students can benefit more from proper feedback if sufficient detail is given. Only saying that the answer is wrong does not identify the mistake and is not helpful to students. It was further claimed that they could forecast all common errors that students could possibly make and use these as possible distractors. It is not valuable to simply give the correct answer (Higgins & Tatham, 2003:1-11). Guo et al. (2014:369-383) disagree in saying that it cannot always be possible to foresee all possible errors that students could make. Therefore, they choose to give more general feedback for a group of questions. In a study performed on university students by Jordan (2012), he proved that the better and the more specific the feedback when using MCQs, the more useful the information and the better the results are. This sentiment is shared by the researcher and this principle was used in this study when feedback was designed for MCQ-items.

Nicol and Dick (2007:199-218) found in their research on feedback effects that there can be significant inconsistencies in the expected results, which could be due to the assessor, content, timing, and to the value of feedback after an option was chosen. This variance in effectiveness could also exist due to the learning content, the task difficulty, the enthusiasm of students, self-credit, or learning goals. This can be ascribed to the FIs as these were used in the theoretical background (Section 2.3.4) of this study. Some categories of feedback may be more significant than others, especially for online MCQs (Epstein et al., 2002:187-201; Higgins & Tatham 2003:1-11; Kuechler & Simkin 2003:389).

Payne and Brinkman (2007:71-75) studied the efficiency of corrective explanatory and video feedback that students use online to support their learning. These categories of feedback give an exact indication of which questions were answered correctly or not. The students need to use the feedback given and repeat the questions until they understand the concepts and answer the questions correctly. Every group of answers needs to have well designed feedback on the students' level of thought without any misunderstandings. Students must be able to use the diagnostic feedback to test their own cognitive level. When they then repeat the test, new feedback will be given according to their knowledge status. This will enforce students to spend more time on self-assessment. By using concept-based feedback, guessing will be less tempting. They are then forced to read the feedback carefully or else they will not even know if their answer was wrong. The focus of a study done by Payne and Brinkman (2007: 71-75) was on the formative assessment used in MCQs. They made use of an experienced examiner to investigate the students' answers to give proper diagnostic feedback.

According to Maier et al. (2016:85-98), feedback, especially when used in online MCQ-assessment, is of the utmost importance but it is often neglected. Feedback encourages the learning process by giving information about incorrect answers. Feedback can be used effectively because of the interactive skills used in online assessments.

Previous research shows that there is a difference in the effect of online and personal feedback. Recent research shows that various types of feedback can be used, while computer-based formative assessment is currently used the most (Maier et al., 2016:85-98)

### **2.2.6 Online MCQs with immediate formative feedback**

Immediate means as quickly as possible during the actual assessment, while formative feedback entails information that is made available to enable a student to change an answer during the process. In online MCQs, immediate formative feedback (IFF) can be given directly after the completion of each item, enabling students to immediately see that a mistake was made, and giving them the opportunity to choose another alternative based on the formative feedback provided.

Delayed formative feedback is feedback that is given to a student after a specified interval (Dempsey & Wagner, 1988:20-25). Delayed feedback can be given directly after the completion of an assessment. This might happen in less than one hour, one day or even one week after completion or before the next session. Delayed feedback is given after all of the questions have been answered (Clariana, Ross and Morrison, 1992:5-17).

The timing as to when feedback must be provided is a significant variable. According to Hattie and Timperley (2007:81-112), and Shute (2008:153-189), feedback must rather be given during the assessment procedure and not later. They also propose that the best time to give feedback depends on the content of the feedback. The timing of feedback is influenced by the kind of learning and assessment taking place. On the one hand, immediate feedback seems to be helpful for procedural learning or where the students need some help (as when doing calculations) because of the level of difficulty. On the other hand, delayed feedback may be more useful for tasks that students find less challenging to do (William, 2011:3-14). Most of the students who formed the sample for their survey preferred to get feedback immediately (Zendejas, Cook & Farley, 2010:432-438). It also resulted in progress in student performance and knowledge.

Roediger and Butler (2011:20-27) note that when immediate feedback is used in assessment, more efficient understanding will result in the long term. Statistical analysis of online MCQs shows that the most used method is to give marks for the correct option and nothing for a wrong or missing answer. Bush (2006:398-404) has identified two critical problems with this, namely, that students need to choose one correct answer out of five possible alternatives without showing their thought process. Also, there is a delay in learning irrespective of whether the answer is right or wrong.

Feedback is normally given after the first response, although for online questions, a second try can easily be allowed (Dempsey & Driscoll, 1989) or may even be repeated until the answer chosen is correct (Pressey, 1926:373-376; 1950:417-447; Buchanan, 2000:193-200). Allowing students to retry several times has significant implications. Answer until correct (AUC) can help students to learn from their mistakes, and they will know if the final answer is the correct one. This was significant for this study as immediate feedback was given to the students to assist them in their thinking.

William and Thompson (2007:53-82) state that MCQs can be used for students to repeat the content of their curriculum until they know it well. When students have challenges, extra feedback will be given and discussions in groups will help to solve the problem(s) (Ramsden, 1992). Immediate feedback can be given on online MCQs to assist students with problem areas and to help them to do their revision. Students feel comfortable using online MCQs because these indicate where the problem areas are that they need to revise. Students can also use the discussion board to assist one another as an important part of peer assessment. This type of learning is not done very often yet. Online MCQs give students the opportunity to study in their own time. Immediate feedback can show students where to improve if they take responsibility and redo questions until they have the correct answer.

Hadsell (2009:135-141) explains that the effectiveness of MCQs is in the immediacy of feedback. Thus, MCQs need to be designed with care to give students immediate, quality feedback. Epstein et al. (2001:889-894) find that immediate feedback is one of the most important recent improvements in online MCQs. Some online MCQs can be set up to give immediate feedback to students while they are busy doing their assessment, while others are set up to give extended feedback afterwards (Escudier, 2011:440-447). Response systems like clickers, where immediate feedback can be given, are available and are adequate to use with large groups of students (Schnell, Lukoff & Mazur, 2013:233-261). Heaslip, Donovan and Cullen (2014:11-24) outline that if a student feels lost regarding a certain part of the work done, knowledge can be recovered by using online MCQs with immediate feedback frequently. Brosvic et al. (2005:205-218) state that immediate feedback is an important part of formative assessment to guarantee that students do not repeat the same mistakes. Higgins and Tatham (2003:1-11) state that to give immediate rather than delayed feedback, common errors should be emphasised. Students can then learn from their own mistakes. The impact of feedback is the highest if it is given immediately after the assessment. This was significant in this study where immediate formative feedback played a prominent part in the development of the proposed MCQs.

In the previous section, the immediacy of feedback was comprehensively discussed, but usually, this refers to post-assessment feedback. In the proposed MCQs central

to this research, the feedback needs to be available during the assessment process and is usually referred to as immediate corrective feedback. In the next section, the significance of immediate corrective feedback will be discussed.

### **2.2.7 Online MCQs with immediate corrective formative feedback**

Much research has been done to use online MCQs without IFF to assess theoretical knowledge (Webb, 1989:216-218; Benvenuti, 2010:21;; Siddiqui et al., 2016:114-121). The literature shows that online MCQs are currently used to give IFF but without granting partial credit, e.g. Key train (LeFebvre & Buddin, 2014), Respondus (Poutre et al., 2015), and Moodle (Brandl, 2005:16-23). Key training is where students can answer a question by choosing one out of a few possible answers. If the answer is wrong, immediate feedback, but no partial credit, is allocated to direct the student to the correct answer. This means that although the student gets feedback, the score is either 0 for a wrong answer or 1 for a correct answer. The questions set up in Key train are furthermore mostly theoretical in nature (LeFebvre & Buddin, 2014). Respondus and Moodle are mediums using the Blackboard platform where lecturers can set up class tests or assignments using multiple distractors. The students can get their results with feedback immediately after completion, but no partial credit is given (Poutre et al., 2015; Brandl, 2005:16-23). In all of these examples, simple feedback is used. The immediate feedback given is to tell students if their answer is right or wrong. Sometimes task specific feedback is used to tell students that the calculation was well done when they did it right or immediately show them the correct answer (Maier et al., 2016:85-98).

Clariana, Ross and Morrison (1991:5-17), and Dempsey and Litchfield (1993:303-327) describe a few different feedback approaches. Knowledge of response feedback indicates if the answer is right or wrong. Knowledge of correct response feedback is when feedback will indicate if the answer was right or wrong and give the correct answer too. Answer Until Correct (AUC) is when the question can be done more than once and the student will be allowed to make another choice until the correct answer is found. Elaborated feedback indicates to the student if the answer was right or wrong, and additional explanations are provided. Students need to learn from their own mistakes, and when corrective immediate feedback is given, they benefit more (Engelbrecht & Harding, 2003:57-65). This is where the current study can be of

significant value.

### **2.2.8 MCQs with immediate formative feedback and the granting of partial credit**

Tamir (1990:563-573) found that almost 30% of students chose the correct MCQ answer for the wrong reason. Engelbrecht and Harding (2003:57-65) agree and add that the marking system could be to blame. They explain that the reason for a student making the wrong choice is an absence of knowledge or the presence of confusion. Students may, alternatively, have the understanding but get the wrong answer because of a minor calculation error. One of the disadvantages of MCQs is that most often, minor and major errors are treated as the same and no provision is made to grant partial credit for correct aspects of the answer or for getting the answer correct after making use of IFF.

Much thought was given to ways to grant partial credit for MCQs that would decrease guessing (Harper, 2003:3-8; Smith, 2013:108). Engelbrecht and Harding (2003:57-65), however, express their concern that students can then pass because of the partial credit granted for each question without ever getting full marks for a question.

Immediate formative assessment techniques (IFAT) can currently be used to grant partial credit to students. This technique was introduced by Epstein, Epstein and Brosvic in 2001:889-894, according to Smith (2013:108). It was born out of their frustration that there was no difference in score for a student who guessed and a student who chose by eliminating their choices. Because MCQs can easily be an unsuccessful method of assessment, IFAT was developed (Merrel, Cirillo, Schwartz & Webb, 2015:50-55). By using the IFAT® forms, a student can immediately see if the chosen answer is correct or not. This is the approach of simple response feedback, as mentioned in the previous section. When the answer is incorrect, another opportunity is given. Partial credit will then be granted. The IFAT® form is a paper that is covered with a thin opaque film. The students then need to scratch the film away from the chosen option. The correct answer is indicated by a star or a blank for the wrong answer (Merrel et al., 2015:50-55) and the score depends on how many times a student had to choose a different answer to get to the correct one. See Figure 2.1 below for an example.

IMMEDIATE FEEDBACK ASSESSMENT TECHNIQUE (IF AT®)					
Name <u>Team #3</u>		Test # <u>2</u>			
Subject _____		Total _____			
SCRATCH OFF COVERING TO EXPOSE ANSWER					
	A	B	C	D	Score
1.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	4
2.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2
3.	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	4
4.	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1
5.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
6.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
7.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
8.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
9.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
10.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	

© 2005 M.L. & S.B. Epstein      ● Form# D008

Figure 2.1: Example of an IFAT® form

This technique was developed to consider learning after the first choice made is wrong. Statistically, a student will have a 25% chance of guessing the correct option the second time if there are four possible options left – that is, one out of the four remaining options. In a study of second responses, 44.9% of the second response was significantly higher as compared to random guessing. This emphasised that students learned from their wrong answer. This shows that the IFAT® forms contribute to learning (Merrel et al., 2015:50-55).

High performing students benefit the most from this procedure (DiBattista & Gosse, 2006:311-328). In this technique, the immediacy of the feedback lies in the awareness of knowing that the answer is right or wrong, knowledge of simple response feedback. Partial credit is granted if the answer was incorrect (Epstein et al., 2002:187-201), but second or third choices are not based on any formative feedback as none is provided.

The technique also benefits assessors as results are available immediately after the completion of the assessment (Brosvic et al., 2006:205-218). The students even need to write down the marks scored for each question depending on how many repetitions were completed, lessening the lecture time spent on marking. Studies show that students benefit more after the completion of an assessment using the IFAT® forms as compared to traditional MCQs (Brosvic et al., 2006:205-218). The actual reason for

the invention of IFAT was to give students a second opportunity to choose the correct answer instead of guessing. The opposite is, in fact, true as due to partial credit granted, students can continue trying until they get the correct answer and it promotes guessing. It is nonetheless a popular choice for MCQ-assessments among educators. Slepko (2013:782-791) and Merrel et al. (2015:50-55) encourage science educators who currently experience challenges using MCQs to use IFAT. Using IFAT was not an option in this study as immediate corrective formative feedback was not given and the MCQ-assessments were not online. The students could not make use of it before choosing a new option in trying to get to the right answer. This implies that a set of MCQs must be designed with built-in IFF, coupled with the granting of partial credit so that successive student choices are informed by the formative feedback and are not random.

### **2.2.9 Student preference for MCQs**

Assessment can be done using online CRQs or MCQs; research has been undertaken to gain the opinions of students regarding their preference for one or the other.

#### *2.2.9.1 Students' preferences for online CRQs*

Foong (2015) finds that some students prefer CRQs due to:

- The expression of own thoughts compared to the selection of provided options. This allows them to think instead of simply choosing an option.
- It requires a skill to structure and explain their own answers.
- It gives a better opportunity to prepare for the exam.
- More effort is required to think, frame and communicate.
- The flow of concepts is easier to follow.
- Since MCQs often focus on only one part of a question, CRQs provide better conceptual understanding.
- More explanation can be given for an answer.
- Prior knowledge needs to be used, which enforces revision.
- Ideas need to be expressed thoroughly and scientifically.
- A full range of answers are allowed where MCQs give limited options.

- CRQs allow limited guessing. Deep and relevant thought must be given before answering to frame the answer in their own words.
- Keywords in MCQs can give clues pointing to the correct answer. CRQs allow thinking about own keywords.
- Skills to formulate thinking are required for the examiner to understand the students' way of thinking.
- Answers from other students can be used for revision purposes.

#### 2.2.9.2 *Students' preferences for online MCQs*

Foong (2015:87-89) further reveals that some students prefer MCQs because:

- It is more time consuming to write out the answer.
- Completing an assessment using CRQs, more stationary is needed and must be brought to class.
- Some students do not prefer to write answers in sentences. They prefer using a clicker.
- When answering in writing, the student may be under the impression that their answer is right, which is not always the case. They will only know when receiving the results.
- Student may sometimes write silly answers for the sake of answering if they do not know the answer.
- MCQs can help students to understand why the answer is right but the other distractors are wrong (depending on if they are given multiple opportunities to find the correct answer).

Bridgman (1992:253-271) found that 81% of students preferred to do MCQs rather than CRQs. Hamilton (1994) in a similar study, discovered that 70% of mathematics students prefer MCQs to CRQs. Some stated that it was not even necessary for them to read all the alternatives to know what the answer ought to be. They also admitted the possibility of guessing the correct answer when using MCQs. This can sometimes be an option if one really does not know the correct answer.

Students place high value on the following features of online MCQs ( Sambell, Sambell & Sexton, 1999:179-191; Canfield, 2001:152-158; Marriott & Lau, 2008:73-90):

- Immediate formative feedback given;
- Option to work at your own pace;
- Identification of areas to improve; and
- Useful to redo for revision.

### *2.2.9.3 Students' preferences for online assessment*

The reasons that students give for preferring online assessment include (Engelbrecht & Harding, 2004:217-231):

- When answering online, you can complete the assessment in your own time and under less exam stress that can cause you to forget what you learnt.
- Time can be more organised in online assessments due to the immediate feedback of online MCQs and the convenience in receiving results immediately.
- When you can see your results immediately, you will see what parts of the work need more attention.
- Online assessments can be done in the students' own time. This results in the students planning their own study time and doing the online assessment where ever it fits into their schedule. It does not require physically going to campus.
- Online assessments expose students to the latest technology. Therefore, their computer skills may also improve.

### *2.2.9.4 Students' preferences for assessment on paper*

The reason students give for preferring written paper assessments are (Engelbrecht & Harding 2004:217-231):

- Partial credit is granted when doing paper CRQs, which can turn the scale towards passing or failing. This is unfortunately not yet common when using online assessment.
- Students are used to the constructed way of writing an exam on paper.

### *2.2.9.5 Comments offered by students with no preference*

Some students think that both methods are equally acceptable when doing assessments. They might prefer online assessment more in some circumstances but find the written part practical when, for instance, a computer is not available.

There are benefits and drawbacks to both methods. It is possible to guess the correct answer in MCQs, but partial credit is granted in CRQs where the answer is wrong due to some aspect being misunderstood or a mistake being made in some part of a calculation.

Einig (2013:847-848) did a study where she asked students to rate the efficiency of MCQs vs CRQs on a five-point Likert scale, where five is very useful and one is not useful at all. The average outcome for the two different cohorts were both more than four out of five. This implies a positive preference for MCQs. She further asked the students to clarify the reason for their answers. The students answered that MCQs gave them the chance to practise in their own time, get immediate formative feedback, and it helped them to recognise the gaps in their knowledge.

In conclusion, the explored literature has the following implications for this study:

- Online assessments are becoming more common and are often preferred by lecturers and students;
- Unfortunately, online MCQ-assessments are, at present, not yet used much to assess higher level conceptual understanding. Calculations are not yet assessed validly and fairly using MCQs as limited delayed feedback is given, partial credit is not granted, and not much learning can take place if students do not know what they did wrong;
- There probably could be a way of assessing calculations using online MCQs as long as immediate formative feedback is given during the assessment. Students must also be given additional opportunities to find the correct answer by using the feedback and if partial credit is granted for finding the correct answer.

## **2.3 THEORETICAL FRAMEWORK**

The purpose of a theoretical framework is to indicate how the current research fits into the familiar context of the reader (Maxwell, 2013). It provides a representation of what is known in general. This current study was built on four theories.

### **2.3.1 Bloom's taxonomy**

Bloom stated in his taxonomy (1969) that evaluation (assessment) should not only focus on the recollection of facts. The higher forms of thinking require students to analyse and

evaluate information, which is of great importance. Bloom explained assessment to be the evaluation of student learning (William, 2006:283-289).

Assessment at all levels is therefore necessary, according to Bloom (1969), as demonstrated in Figure 2.2.

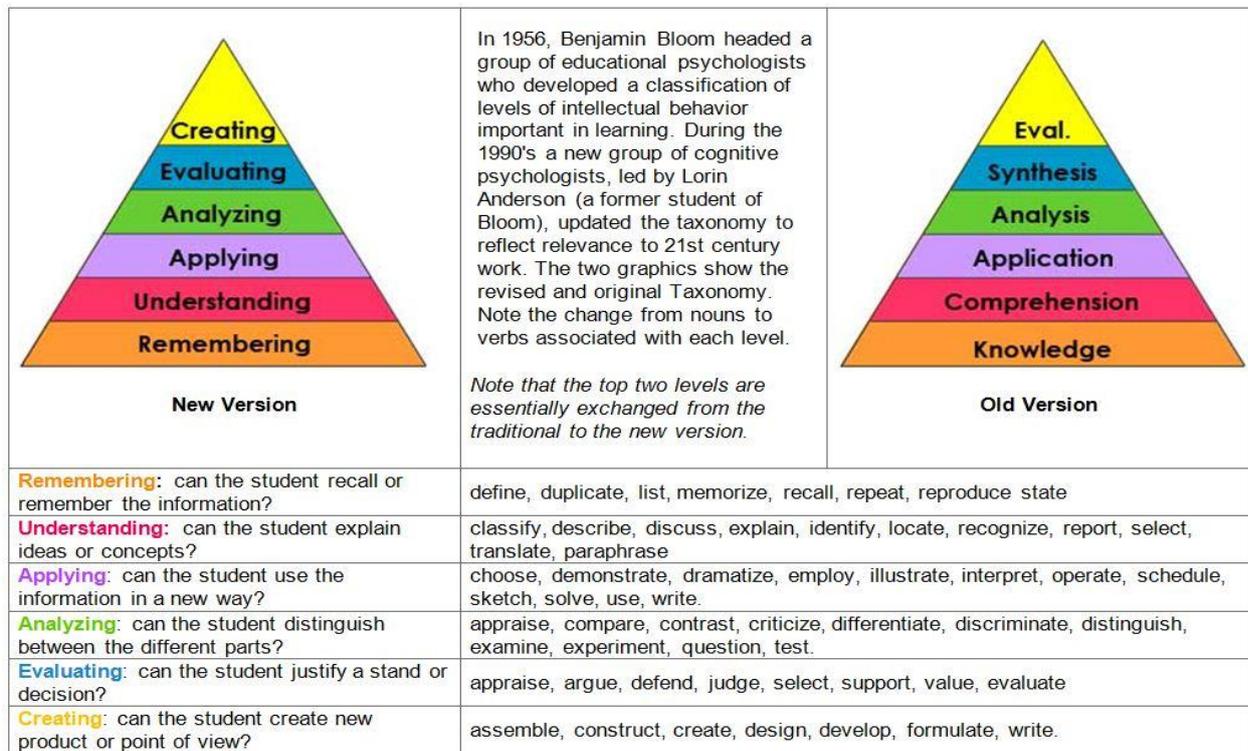


Figure 2.2: Bloom's Taxonomy

Gibbs and Simpson (2005:3-31) emphasises the value of thorough assessment in higher education. He states that assessment based on knowledge alone is being developed into an assessment of skills, competencies and abilities rather than knowledge only. Assessment can be seen as a very persistent system that control students, even more than what most lecturers would like to acknowledge. Currently, MCQs are mostly used to test theoretical knowledge while the urge in this study was to propose that the assessment should rather be based on the testing of knowledge only, and not on skills, competencies and abilities.

This research proposes the use MCQs with IFF, and the granting of partial credit to assess knowledge and skills at higher cognitive levels.

### 2.3.2 The Assessment Component Taxonomy

Though assessments were traditionally given at the end of a block to test the knowledge of students, it is increasingly being used to motivate and support learning (Geysler, 2004). Therefore, there is a difference between learning from assessment and learning for assessment. Huntley et al. (2011:3-16) used this background to propose a taxonomy named the Assessment Component Taxonomy.

The Assessment Component Taxonomy was developed to identify the components in mathematics where alternative assessment methods like MCQs could be used. Smith et al. (1996:65-77) modified Bloom's Taxonomy to structure assessment tasks. He called his taxonomy the Maths Taxonomy. By using Bloom's Taxonomy (Bloom, 1969), and the Maths Taxonomy (Smith et al., 1996:65-77), the mathematics assessment components are categorised according to the cognitive levels of difficulty. The taxonomy consists of seven mathematical components; see Table 2.1 (Huntley et al., 2011:3-16). The researcher is of the opinion that this is also relevant for science because calculations are used in both subjects.

Table 2.1: Mathematics Component Taxonomy and cognitive level of difficulty

Mathematics assessment components	Cognitive level of difficulty
1. Technical 2. Disciplinary	Lower order / Group A
3. Conceptual 4. Logical	Middle order / Group B
5. Modelling 6. Problem solving 7. Consolidation	Higher order / Group C

In Table 2.2, Huntley et al. (2011:3-16) align the assessment components with the cognitive skills required.

Table 2.2: Mathematics Assessment Component Taxonomy and cognitive skills

Mathematics assessment components	Cognitive skills
1. Technical	<ul style="list-style-type: none"> <li>○ Manipulation</li> <li>○ Calculation</li> </ul>
2. Disciplinary	<ul style="list-style-type: none"> <li>○ Recall (memory)</li> <li>○ Knowledge (facts)</li> </ul>
3. Conceptual	Comprehension: <ul style="list-style-type: none"> <li>○ algebraic</li> <li>○ verbal</li> <li>○ numerical</li> <li>○ visual (graphical)</li> </ul>
4. Logical	<ul style="list-style-type: none"> <li>○ Ordering</li> <li>○ Proofs</li> </ul>
5. Modelling	Translating words into mathematical symbols
6. Problem solving	Identifying and applying a mathematical method to arrive at a solution
7. Consolidation	<ul style="list-style-type: none"> <li>○ Analysis</li> <li>○ Synthesis</li> <li>○ Evaluation</li> </ul>

From the results of a study carried out by Huntley et al. (2011:3-16), as illustrated in Table 2.2, it was found that MCQs performed better as an assessment method in the technical and modelling components than CRQs. It is still challenging to set up MCQs in the logical and consolidation components because CRQs performed better in those areas. There is no noticeable difference in using either MCQs or CRQs in the disciplinary assessment component. In the conceptual and problem-solving assessment components, MCQs performed only slightly better than CRQs. The problem-solving component is focused on finding a mathematical method to solve word problems. The way in which Huntley et al. (2011:3-16) approach the problem-solving questions is to divide them into sub-questions concentrating on different cognitive skills. Marks are then allocated step-by-step. The comparison of CRQs and MCQs and their successes is presented in Table 2.3 below.

Table 2.3: Comparison of the successes of MCQs and CRQs

Mathematics assessment Component	Comparison of success
Technical	MCQs can be used successfully
Disciplinary	No difference
Conceptual	MCQs can be used successfully
Logical	CRQs more successful
Modelling	MCQs can be used successfully
Problem solving	MCQs can be used successfully
Consolidation	CRQs more successful

The Assessment Component Taxonomy was chosen as the theoretical framework for this research as it indicates where in the taxonomy MCQs may be used successfully.

### 2.3.3 Item Response Theory

The pioneering work of the Item Response Theory (IRT) was done between 1950 and 1960 by Frederic M. Ford, George Rasch and Paul Lazarsfeld. Benjamin Wright and David Andrich later did more research on this topic (Lord, Melvin, Novick & Birnbaum, 1968).

IRT is a model that tests the reaction of a student to a question. On the one hand, a student with good skills will most probably answer an easy MCQ due to a correct calculation and therefore will choose the correct alternative. On the other hand, if a student has poor skills and the question is difficult, he/she will probably answer the MCQ wrong due to making a common mistake within the calculation process (Hambleton et al., 1991). The IRT is the mathematical function that gives the probability that a student with a specific ability will answer the question correctly. IRT does not assume that the level of difficulty for all questions are the same (Moncrief & Foster, 2012). It sets the focus of the theory on the item that is compared to another item. This is an improvement on the Classical Test Theory, which sets the focus on the test-level.

The three parameter IRT model includes a parameter that indicates the chances that a student could guess the answer correctly. The guessing parameter is a person parameter because the guessing of an item is done on an individual basis. The distributions of these parameters indicate the degree of guessing (Hambleton & Swaminathan, 2013).

It was found in IRT that students make estimated guesses and do not just guess randomly. In this study guessing was shown to be a concern when using MCQs as an alternative assessment method.

### **2.3.4 Feedback interventions**

Kluger and DeNisi (1996:67-72), who did important work on Feedback Interventions (FIs), define it as one of the most widely functional psychological interventions. This feedback is psychological because it can determine if attention is shifting away from the task towards other goals, like the self. FIs give students feedback regarding their task performance (Ilgen, Fisher, & Taylor, 1979::349; Kluger & DeNisi, 1996:67-72), which could either disappoint or compliment the student (Kluger & DeNisi, 1996:67-72). Using FIs in combination with goal-setting, attention will be directed to the task at hand and not the self.

Feedback from a computer will increase task motivation, and therefore performance too. The feedback source, as well as the usefulness thereof will make a difference in performance. Those are not the only factors; Alder (2007:157-174) indicates that the timing and frequency of the feedback are also important. If participants are motivated and treated fairly, it will reflect positively in the quality of their work, but not necessarily negatively on the quantity thereof. In contrast, unfair treatment will lead individuals to sacrifice quality with no increase in quantity of production (Alder, 2007:157-174).

In this study, the proposed feedback, given online by a computer, increased task motivation even further due to the use of IFF with the granting of partial credit. This was hypothesised to make a difference in performance because the participants would be motivated and treated fairly. The researcher posited that it would reflect positively in the quality of work by taking the focus away from the self and back to the task.

## **2.4 SUMMARY**

The literature review explored the scholarly work relevant to this study.

This study focused on the use of MCQs to assess problem-solving calculations and the following aspects derived from the literature study were relevant in this endeavour.

When assessment is used to satisfy the learning needs of students, it is considered to be formative. Formative assessment can only be useful if feedback is given to the participants. The extent of the feedback is important, but the timing even more so as the sooner the feedback is given, the more the students will benefit from it. Immediate formative feedback is helpful if it is given as soon as it is allowed. If the assessment was to be given online, more students could make use of it in their own time and could see the outcome of their assessments immediately.

MCQs and CRQs are the two main methods of formative assessment. Online MCQs are currently being used more often and are useful for self-study. This saves marking time and enforces consistency between different assessors. Cheating can also be prevented by randomising the questions.

The distractors for MCQs, however, must be chosen with great care. If the distractors are not functional, the validity and reliability of the questions will be affected negatively. Students can then also choose the correct answer for the wrong reasons. Moving from CRQ- assessment methods to MCQ-assessments will initially increase the workload, but after a database has been set up, the workload will be substantially smaller.

Online MCQs without IFF cannot be used to assess problem-solving calculations effectively. This is because students get no immediate feedback that they can utilise and for which they can be awarded partial credit if they correct their mistakes and find the correct solution. IFAT was designed to give students a second chance to choose the correct answer, but immediate corrective feedback is still not granted for giving the incorrect answer. At present, there are therefore no MCQs in any format entirely suitable for assessing problem-solving calculations effectively.

In a quest for a solution to this dilemma, this research investigated the possibility of designing MCQs in which formative corrective feedback is given during the assessment process, where students get multiple opportunities to find the correct answer after utilising the feedback, and where partial credit is granted that may overcome the present shortcomings of MCQ-assessments. This would make them suitable for assessing problem-solving calculations effectively.

The next chapter deals with the research methodology used to conduct this research.

# CHAPTER 3: RESEARCH METHODOLOGY

## 3.1 INTRODUCTION

As indicated in the previous chapter, the literature shows that online MCQ-assessment is a popular way to assess knowledge and it is most often used without giving immediate formative feedback.

This chapter provides an overview of the research design and research methods used to investigate if an appropriate form of online MCQs with IFF and the granting of partial credit can be developed and used to assess problem-solving calculations effectively. In this vein, Creswell (2012:3) explains, “Quantitative research includes a substantial amount of literature at the beginning of the study to provide direction to research questions or hypotheses.” This is believed to have been done effectively in the previous chapter as meaningful research question, sub-questions and hypotheses were derived from the literature study. This will be discussed forthwith.

### 3.1.1 The research sub-questions

The following sub-questions were formulated to answer the main research question:

- When setting online MCQs, how can an assessor determine which distractors would anticipate the mistakes that students could make so that formative feedback can be built in to indicate to them what they did wrong and how to correct it?
- Can online MCQs without IFF replace CRQs in fairly assessing students’ ability to do problem-solving calculations in science?
- Are online MCQs with IFF, and the granting of partial credit, more reliable and fair than MCQs without IFF in assessing students’ ability to do problem-solving calculations in science?
- Can online MCQs with IFF and the granting of partial credit be used to assess problem-solving calculations as a fair and valid alternative to CRQs?

### 3.2 RATIONALE FOR EMPIRICAL RESEARCH

Much research has been done on the use of online MCQs without IFF to assess the theoretical knowledge of students. The literature shows that IFF feedback can be given when doing online MCQ-assessments but without granting partial credit. Key training (LeFebvre & Buddin, 2014), Respondus (Poutre et al., 2015), Moodle (Brandl, 2005:16-23) and two and three tier <sup>5</sup> (Caledon & Subramaniam, 2010:939-961) are examples of MCQs with immediate feedback but without the granting of partial credit. Online MCQs can currently also be used to grant partial credit to students but without immediate corrective formative feedback, such as IFAT (Merrel et al., 2015:50-55). This study was therefore undertaken to investigate the effective use of online MCQs for problem-solving calculations where IFF is given during the assessment process, where multiple opportunities are given to find the correct answer after making use of corrective feedback and where it is possible to grant partial credit.

### 3.3 RESEARCH DESIGN

The aim of this study was to contribute to the current pool of knowledge regarding the use of online MCQs to assess problem-solving calculations with IFF and with the granting of partial credit.

De Vos, Fouché and Delport (2011:142) describes a research design as the method used to collect detailed data. Furthermore, it is a plan to guide the study to gain maximum control over the factors that can interfere with the validity of the research results (Mzee, 2016:1-19). According to Burns and Grove (2001:223),

*The research design is the researcher's overall plan for obtaining answers to the research questions guiding the study. Designing a study helps researchers to plan and implement the study in a way that will help them obtain the expected results.*

The research design of this study is the plan and technique that includes the assumptions made from the assessment. Formative assessment with IFF and the granting of partial

---

<sup>5</sup>Two-tier tests are a more cultured form of MCQ. The first tier is in an MCQ format but gives the students two choices. The second tier is in the MCQ format again, where the reason for the choice made in the first tier is asked.

credit was the specific focus of investigation (Creswell, 2012:3). This research was designed to determine the type of evidence that was required to answer the research questions posed in this study (De Vos, Fouché & Schurink, 2011:324). The researcher chose to collect data from an experiment to prove or disprove the hypotheses made (Creswell, 2009:16). The focus therefore was on finding answers to the sub-questions and therefore the research question.

### **3.3.1 Research paradigm**

The researcher felt comfortable using the positivist paradigm as the research approach because the research to be undertaken was quantitative in nature, would involve an experiment, and the research findings would be noticeable and measurable. The role of the researcher was focused on data collection and interpretation, where after the hypotheses were tested.

### **3.3.2 Research approach**

A quantitative approach was used as the research was undertaken to explain the relationship between variables and to try and validate them (Leedy & Ormond, 2005). The variables of interest – the application of CRQs, MCQs without IFF, and MCQs with IFF and the granting of partial credit – were objectively measured and the researcher remained distanced from the students to draw unbiased conclusions (De Vos, Strydom & Delport, 2011:292). The statistical two-sample t-test procedure was used to draw conclusions from the data and the hypotheses tested (Maree & Pietersen, 2011:210). All of these aspects obviously fell within the gambit of quantitative research.

### **3.3.3 Research type**

One group of pre-test, post-test, and pre-experimental research (De Vos, Fouché & Delport, 2011:145) was undertaken because the research was not focused on how the different groups would react to different types of assessment or the absence thereof, but on how one group would experience the application of different types of assessment. The comparisons were not made between the different groups but between the different types of assessment and their outcomes. In this research, one group was subjected to three types of assessment in the form of a pre-test and two post-tests in order to compare the

outcomes with one another. This research therefore also required the postulation of three hypotheses, as described under Section 3.4.2.

### **3.4 RESEARCH METHODS**

Research methods entail the ways of collecting, analysing and interpreting data that researchers use to find answers to their research question and sub-questions.

#### **3.4.1 Sampling**

The students of the Engineering and Built Environment extended course at TUT comprised the population of interest. The extended course consists of 160 first-year students who exhibit the same demographic characteristics and who do not meet the academic requirements to enrol for the equivalent diploma course. The students were randomly assigned to one of four groups, which consisted of 40 students each, and one class of engineering science students was selected due to convenience (Chadwick et al., 1984) as it was a group of students that the researcher was lecturing. One group of 40 students represented the experimental group because they attended the same lectures and were taught the same content by the same lecturer over the same period of time. The experimental group was furthermore selected through purposive convenience sampling since it was the only group in which all of the students agreed to take part in the voluntary tests, for which they did not receive any credit.

#### **3.4.2 Data collection**

The data were collected by undertaking a one group pre-experiment (pre- and post-tests) (De Vos, Fouché & Delport, 2011:145) as explained in Section 3.3.3, to test the following three hypotheses:

##### **First hypothesis**

**H<sub>0</sub>:** There is no significant difference between the average mark for CRQs and online MCQs without IFF.

**H<sub>1</sub>:** The average of CRQs is significantly higher than the average of online MCQs without IFF.

## Second hypothesis

**H<sub>0</sub>:** There is no significant difference between the average for online MCQs with IFF and the granting of partial credit, and online MCQs without IFF.

**H<sub>1</sub>:** The average of online MCQs with IFF and the granting of partial credit is significantly higher than the average of online MCQs without IFF.

## Third hypothesis

**H<sub>0</sub>:** There is no significant difference between the average for CRQs and online MCQs with IFF and the granting of partial credit.

**H<sub>1</sub>:** The average of CRQs is significantly different to that of online MCQs with IFF and the granting of partial credit.

A dependent two-sample t-test was used to decide if the null hypothesis should be rejected or not. This test was used because the average marks obtained by the experimental group for the same set of questions, which were assessed in two different ways, were compared.

A **baseline assessment** was undertaken on an independent group of students (from a prior year) to identify the common mistakes that students make when doing a number of problem-solving calculations. The assessment required the students to answer the stated questions and to write down the calculations step-by-step to derive their answers. The same questions used in the baseline assessment were asked in the pre- and post-tests to the current students. This assessment was not done by the participants of the current study because the content of the assessment could not be known. The common mistakes identified were used to design the distractors for the MCQs and to design the online IFF used in the second post-test.

The CRQs were written in a venue during normal lecturing times to make the circumstances as natural as possible. According to the ethical clearance requirements, an independent assistant was needed to assist in this study. Brenda Ngoma fulfilled the role of the independent assistant (see Appendix A for the consent form). She gave every student a random number between 700M and 790M, which also served as the confidential password to access the MCQs. It was compulsory that assessment M had to

be done before assessment N. Every random number given could only be used once. Assessment M represented the MCQs without IFF, and assessment N represented MCQs with IFF and the granting of partial credit. Four calculation questions on the topic *work and energy*, as required in the syllabus, were given to 40 first-year Engineering Science students to complete in a 90-minute period (Appendix C).

The students answered the questions for the **pre-test** in writing on a paper with a carbon copy paper attached. The original paper was handed in for marking and the students used the duplicate as a reference when completing assessment M and N (**the first and second post-tests**). They used the problem-solving calculations of their copy from the CRQs to choose one of the five alternatives. At that moment they did not know if their answers were correct. Upon the completion of the online MCQs without IFF (assessment M), the students did the online MCQs with IFF and were granted partial credit (assessment N). Assessment N was designed by adding IFF and the granting of partial credit to assessment M. This was required to allow the students who chose the incorrect alternative to make further choices after applying the corrective feedback they received when making the wrong choices. The results of the different tests were then compared to analyse the three hypotheses (see Figure 3.1).

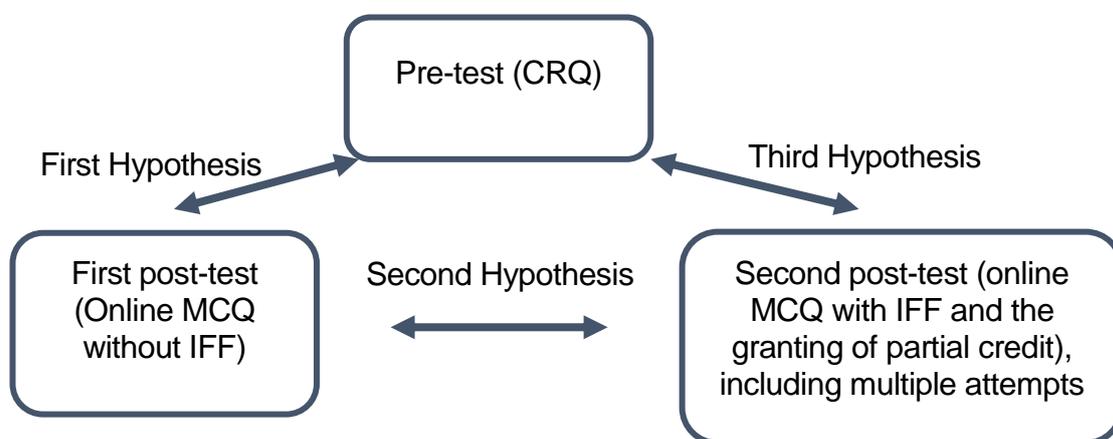


Figure 3.1: The comparison between the average marks of all three tests

A computer program was designed by a programmer, Henry de Nysschen, especially for the purpose of this study. The program can be found at the following link: <http://at.lain.co.za>. Provision was made for the students to do assessment M before they were allowed to do assessment N. M and N were chosen to ensure that the students would be forced to do the one assessment before the next assessment. Figure 3.2 and

Figure 3.3 show examples of assessment M and assessment N as they were given to the students.

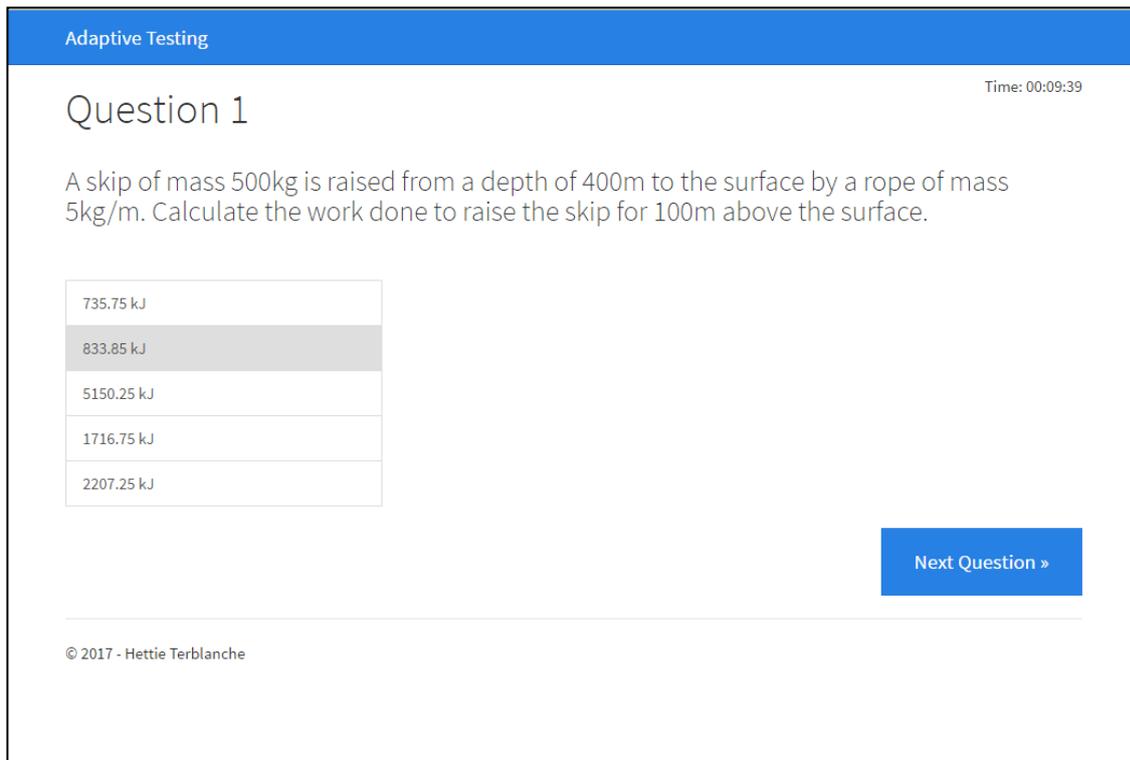


Figure 3.2: Example of computer program screen for assessment M

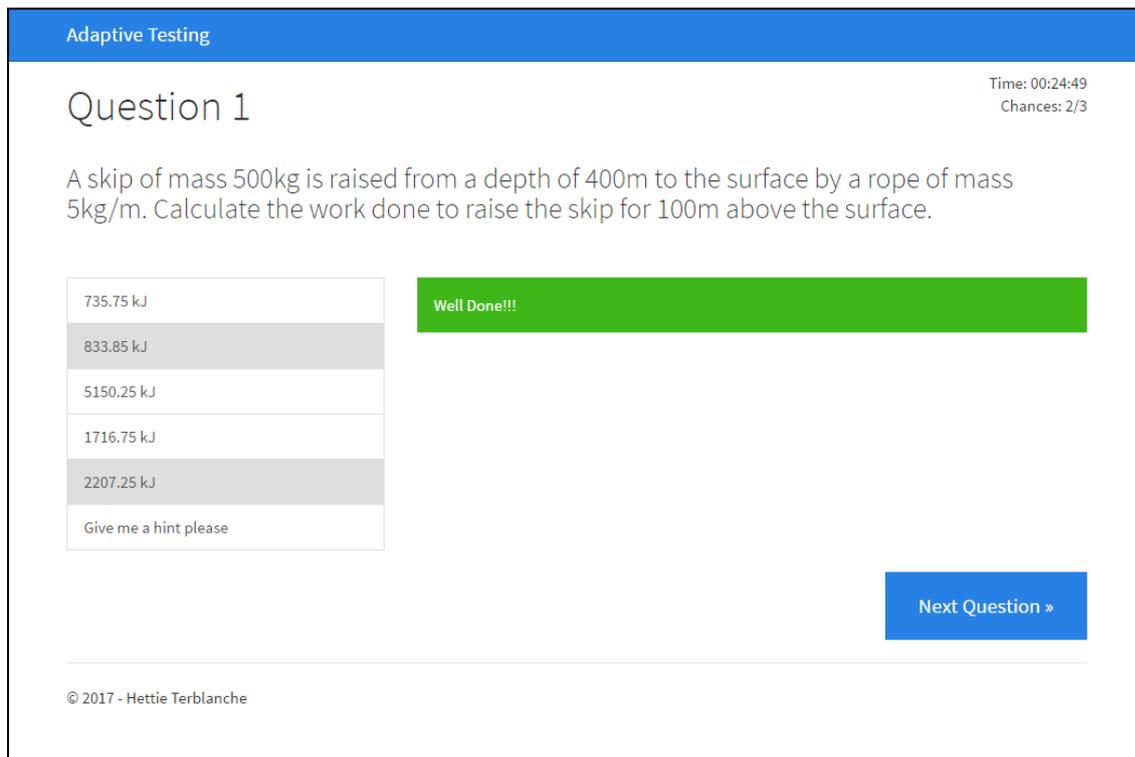


Figure 3.3: Example of computer program screen for assessment N

See screenshots of the program in Appendix D.

### 3.4.3 Data analysis

The pre- and post-test were assessed in the following manner:

In the **pre-test**, the CRQs were marked by the examiner with the granting of partial credit. A student who completed all of the calculation steps and the answer correctly received full marks. A student with an incorrect answer received partial marks for the correct aspects of the step-by-step calculations. Figure 3.4 shows that student A has full marks – five out of five, and student B made one minor mistake and received three out of five.

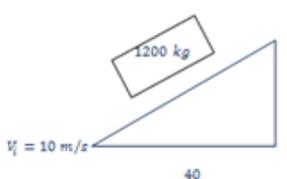
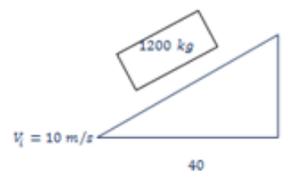
Student A	Student B
<p><b>Question 4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p>  <p style="text-align: right; margin-right: 50px;"><b>5/5</b></p> $E_{pot} + E_{kt} + E_{add} = E_{pf} + E_{kf} + E_{loss}$ $\frac{1}{2}(1200)(10)^2 + 0 + E_{add} = \frac{1}{2}(1200)(45)^2 + (1200 \times 9.81 \times \frac{300}{40}) + (300 \times 1.2\text{ton} \times 50\text{N/ton})$ $60000 + E_{add} = 1215000 + 88290 + 18000$ $E_{add} = 1261.29 \text{ kJ}$	<p><b>Question 4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p>  <p style="text-align: right; margin-right: 50px;"><b>3/5</b></p> $E_{pot} + E_{kt} + E_{add} = E_{pf} + E_{kf} + E_{loss}$ $\frac{1}{2}(1200)(10)^2 + 0 + E_{add} = \frac{1}{2}(1200)(45)^2 + (1200 \times 9.81 \times \frac{300}{40}) + (300 \times 1.2\text{ton} \times 50\text{N/ton})$ $60000 + E_{add} = 1215000 + 88290 + 18000$ $E_{add} = 1261.29 \text{ kJ}$

Figure 3.4: Example of how marks are awarded in CRQs

The **first post-test** was marked by means of the specifically designed computer program. A student who chose the correct alternative in the MCQs without IFF assessment received full marks in contrast to a student who chose a wrong distractor, and who received no marks, these are shown in Figure 3.5.

Student A	Student B
<p><b>Question4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p> <p>A) 290.1 kJ B) 4704.6 kJ C) 1243.35 kJ D) 1315.29 kJ <input checked="" type="radio"/> E) 1261.29 kJ</p> <p style="text-align: center;">5/1 5</p>	<p><b>Question4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p> <p>A) 290.1 kJ B) 4704.6 kJ C) 1243.35 kJ <input type="radio"/> D) 1315.29 kJ E) 1261.29 kJ</p> <p style="text-align: center;">5/1 0</p>

Figure 3.5: Example of how marks are awarded in MCQs without IFF

The **second post-test** allowed for multiple opportunities in the MCQs with IFF and the granting of partial credit. The assessments were marked as follows:

- a) If the answer was correct, the computer program would give the student five marks for the correct answer.
- b) If the student chose a wrong alternative, a specific IFF would be given to tell him/her what he/she probably did wrong and how it could be corrected. Four marks would be allocated if the correct answer was then chosen on the second attempt.
- c) Step b would be repeated if the student chose another wrong alternative. Two marks would be allocated if the correct answer was chosen on an attempt after the second round of feedback.

The fifth distractor was, “Can I have a hint please?”

- i) If this option was chosen, a general hint was given and the student would still get a maximum of three out of the five marks for then choosing the correct alternative.
- ii) Step b above was repeated if the student chose a wrong alternative but the student could then only get a maximum of one out of the five marks.

Figure 3.6 illustrates how a student who chose a wrong alternative could benefit by receiving partial credit when applying the IFF received in the next attempt. A student who chose the correct alternative on his/her first attempt would still receive full marks but would not benefit additionally from the IFF, as referred to in Section 2.2.1.

Student A	Student B
<p style="text-align: center;"><i>First attempt</i></p> <p><b>Question4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p> <p>A) 290.1 kJ B) 4704.6 kJ C) 1243.35 kJ D) 1315.29 kJ <input checked="" type="radio"/> E) 1261.29 kJ F) Give me a hint please</p> <div style="text-align: center; color: red; font-size: 2em; border: 1px solid red; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin: 20px auto;"> <span style="writing-mode: vertical-rl; transform: rotate(180deg);">0/1 5</span> </div>	<p style="text-align: center;"><i>First attempt</i></p> <p><b>Question4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p> <p>A) 290.1 kJ B) 4704.6 kJ C) 1243.35 kJ <input checked="" type="radio"/> D) 1315.29 kJ E) 1261.29 kJ F) Give me a hint please</p> <p style="text-align: center;"><i>IFF received</i></p> <p style="color: red;"><b>Recalculate <math>E_k</math></b></p> <p style="text-align: center;"><i>Second attempt</i></p> <p><b>Question4</b> The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the use ful energy supplied by the engine.</p> <p>A) 290.1 kJ B) 4704.6 kJ C) 1243.35 kJ D) 1315.29 kJ <input checked="" type="radio"/> E) 1261.29 kJ F) Give me a hint please</p> <div style="text-align: center; color: red; font-size: 2em; border: 1px solid red; border-radius: 50%; width: 40px; height: 40px; display: flex; align-items: center; justify-content: center; margin: 20px auto;"> <span style="writing-mode: vertical-rl; transform: rotate(180deg);">0/1 3</span> </div>

Figure 3.6: Example of how marks are awarded in MCQs with IFF and the granting of partial credit

In the pre-test where the CRQs were used as the assessment method, the calculations were done step-by-step in writing. Partial credit was given and therefore the possible marks allocated could be 0, 1, 2, 3, 4 or 5.

In the first post-test when the MCQs without partial credit were used, alternative answers could be chosen but a student could only get 0 or 5 marks.

In the second post-test where the MCQs with IFF and the granting of partial credit were used, alternative answers could be chosen but if the answer was wrong, IFF was given and a student could get another chance to choose the correct answer based on the immediate corrective feedback. Therefore, a student could receive partial credit and could get either 0, 1, 2, 3, 4 or 5 marks (see Table 3.1).

Table 3.1: Presentation of the differences of the pre- and post-tests

Assessment	Written step-by-step calculation	Select alternative answers	Multiple attempts	IFF	Partial credit	Possible marks (out of 5)
Pre-test (CRQs)	X	-	-	-	X	0, 1, 2, 3, 4 or 5
First post-test (MCQs without IFF)	-	X	-	-	-	0 or 5
Second post-test (MCQs with IFF and the granting of partial credit).	-	X	X	X	X	0, 1, 2, 3, 4 or 5

### 3.4.4 Measures for trustworthiness

Results are considered to be trustworthy if the study is shown to be reliable and valid.

#### 3.4.4.1 Reliability

“Reliability has to do with the consistency or repeatability of a measure or an instrument. High reliability is obtained when the measure or instrument will give the same results if the research is repeated on the same sample” (Maree & Pietersen, 2011:147). Consistent scoring in MCQs is almost guaranteed (McCoubrie, 1990:709-712).

The software program, written for the purpose of this study, can be used for other groups and for other problem-solving calculations. It can even be used for mathematics, which meets the requirement for the repeatability of an instrument.

#### 3.4.4.2 Validity

Validity is a quantity or instrument that measures what it is supposed to measure (Maree & Pietersen, 2011:147). The way to measure validity depends on whether a number of professionals will get the same answer for a specific problem (Haladyna, 2011). The validity of this study was established in several ways. Vos

##### **Internal validity**

“Internal validity threats are experimental procedures, treatments or experiences of the participants that threaten the researcher’s ability to draw correct inferences from the data about the population in an experiment” (Creswell, 2009:162).

Internal validity refers to the factors that might have an influence on the relationship between the dependent and independent variables. The three assessments that comprised the experiment were completed during a continuous period of time so that no time passed between the different assessments. The first assessment was on CRQs (the independent variable), while the second assessment was on MCQ without IFF (the dependent variable). The third assessment was on MCQs with IFF and the granting of partial credit (the dependant variable). All three assessments were done on the same questions and nobody could get the questions from anywhere or anyone else. This implies that no external communication could influence the students’ marks. Therefore, no student could be positively or negatively influenced. The students were not forced to take part in this research; and an independent assistant was present to observe that all assessment requirements of the university were met. This study could therefore be considered to be internally valid.

##### **External validity**

The validity of a test is determined by how well it models the knowledge, skills and abilities that students must obtain. The external validity in this research was ensured by the following:

The researcher did not favour any individual student. All of the students of a specific group were part of the sample group irrespective of their characteristics or level of knowledge. In this study, the difficulty levels of the chosen questions were selected carefully. During

the time when the experiment was carried out there were no power cuts. Because it was an online experiment, this could have caused serious trouble, fortunately it did not happen.

Unfortunately, it was not possible to make provision for every calculation mistake that a student could possibly make, but provision could be made for the the most common mistakes that students could make because of a lack of subject knowledge or a misunderstanding. A specific part of the science syllabus was tested in this study, namely, the section on *work and energy*, which was properly taught to the students. The questions were chosen to test the range of the work done on *work and energy*. The purpose was then to determine if there was a proper understanding of the work done. If the students did not understand the work and needed more guidance, IFF was used to help them and they could be granted partial credit.

#### 3.4.4.3 Objectivity

Marking the online MCQs could not be influenced by subjectivity. The more questions there are in an MCQ-assessment, the more reliable the results will be. Therefore, this study could have been more accurate if more questions with different levels of difficulty were used. The variables were, however, objectively measured and the researcher remained distanced from the students to draw unbiased conclusions.

### **Research variables**

The independent variable in this research was the problem-solving calculations that needed to be done as CRQs.

The dependent variables were:

- MCQs without IFF; and
- MCQs with IFF and the granting of partial credit.

Distractors were chosen carefully for the MCQs and tested thoroughly. A baseline study was done the previous year by giving students the same questions to do. The most obvious mistakes made by the students, were used as distractors for this study. Care was taken that no irrelevant information was given in either the stem or the distractors. The

students were not guided to the correct choices, but were given indications of what mistakes they made so that they could recalculate and find the correct answer.

### **3.4.5 Ethical considerations**

De Vos et al. (2000) highlights the importance of considering ethical guidelines and ethical problems that might appear in the research. Subsequently, he points out several consequences that should be avoided:

- Informed consent (Appendix A) requires a clear indication of the aim of this study, the procedures, advantages and disadvantages, the risks to which the participant may be exposed, and the reliability of the researchers, which were all discussed with the participant and his/her guardian.
- The purpose of this study must be emphasised to minimise any harm to the participants (De Vos & Strydom, 2011:115). In this study, no personal questions were asked, which prevented any emotional harm being inflicted on the students. All information about the research was made clear to the students (De Vos & Strydom, 2011:115).
- All 40 students that participated in this research signed a consent form to allow the researcher to use the information anonymously as part of this study. The information collected during the project, which was anonymous, was only used for research purposes and was not released for any academic assessment, study progress and/or disciplinary purposes.
- The data will be stored for five years as hard copies in a locked cupboard and soft copies on a computer with password protection.
- The independent assistant, Brenda Ngoma, also filled in a consent form (see Appendix A).

McMillan and Schumacher (2001:150) suggest that the following cautionary ethical guidelines should be implemented:

- “To respect the participant’s privacy.
- To make every effort to minimise an inaccurate interpretation of the data.
- To make known the results of the research to the participant and his/her guardian(s)”.

See Appendix A for the informed consent forms that all 40 participated students signed. There was no pressure on any student to take part in this study. They could withdraw at any time without any negative consequences. Ethical clearance for the research was granted by both TUT and UNISA (see Appendix B).

### **3.5 SUMMARY**

The researcher identified a need for research regarding the use of online MCQs to effectively assess problem-solving calculation questions in science. Quantitative research was undertaken using pre-experimental, one-group pre- and post-tests with a purposeful convenience sample. This was done to determine if the MCQs with IFF and the granting of partial credit could be used as an effective alternative for CRQs. Ethical clearance for the research was granted by both TUT and UNISA.

In the next chapter, the data from the experiment will be analysed and interpreted.

# CHAPTER 4: DATA ANALYSIS AND RESULTS

## 4.1 INTRODUCTION

In the previous chapter, the research design and methodology were discussed in detail. In this chapter, the data will be analysed and interpreted.

### 4.1.1 Research design and methodology

A quantitative approach was used to provide data via an experiment that was carried out. This data were analysed to verify if it could make a difference in the use of MCQs. The researcher set up a set of hypotheses from an experiment through a positivist paradigm approach. The main aim of this pre-experimental research was to identify the links between the use of MCQs with IFF and the granting of partial credit, and CRQs. The researcher studied a single group and was available during the assessment of the experiment for questions. Purposive convenience sampling was used to select an experimental group from the population who could participate in the experiment.

### 4.1.2 Research question and sub-questions

The main research question of this study is:

***Can online multiple-choice questions, designed to provide immediate formative feedback with the granting of partial credit, be used as an effective alternative for constructed written questions in the assessment of problem-solving calculations for Engineering Science students?***

#### **The research sub-questions**

In order to answer the research question, the following sub-questions are formulated:

- When setting online MCQs, how can an assessor determine which distractors would anticipate the mistakes that students could make so that formative feedback can be built in to indicate to them what they did wrong and how to correct it?
- Can online MCQs without IFF replace CRQs in fairly assessing students' ability to do problem-solving calculations in science?

- Are online MCQs with IFF, and the granting of partial credit, more reliable and fair than MCQs without IFF in assessing students' ability to do problem-solving calculations in science?
- Can online MCQs with IFF and the granting of partial credit be used to assess problem-solving calculations as a fair and valid alternative to CRQs?

## 4.2 RESEARCH PROCESS

The data collection process was completed as proposed in the methodology section of this study (see section 3.4.2). One of the anticipated challenges was the possibility of load shedding that could cause some problems because some of the assessments that were part of the experiment were done online – fortunately, that did not happen. The internet connection also functioned well, and the students did not experience any problems completing the online assessments.

After the process had been explained to the students, the question papers were handed out, as well as the two blank papers interfaced with carbon paper. After each student received their confidential password, they completed the written CRQs. The assessment consisted of four calculation questions on the topic *work and energy*, as required in the syllabus. It was given to 40 extended Engineering Science students to complete in a 90-minute period. One of the copies was handed in and was marked granting partial credit where parts of the calculations were correct, although the final answer might have been wrong. The other copy was used by the students to complete an online MCQ-assessment with no immediate feedback. The students were required to use the answers from their copy of the CRQ to complete the MCQs. The MCQs were marked and the results were compared with the results from the CRQs. The students then had to do the third assessment where IFF was included. This allowed them to obtain feedback and could change their answers based on the corrective feedback, and could gain partial credit for finding the correct answer on a second or third attempt. At this time in the post-test, students would not yet know the results of the previous two assessments.

A specially designed computer program was used to allocate partial credit. Depending on the number of choices the students needed in order to find the correct answer, he/she could receive five, four, three, two, one, or zero out of five for each question.

The average marks for the CRQs were compared with the average marks for the online MCQs without IFF, and then also with the average marks of the online MCQs with IFF. This allowed multiple-choices to find the correct answer and the granting of partial credit depending on the number of attempts needed to do so. Lastly, the average marks of the two MCQ-assessments were compared.

Figure 4.1 is a presentation of how the experiment and its associated pre- and post-tests were set up. The CRQs are written as pre-test. Post-test 1 was then online MCQs with the same questions. The common mistakes from the baseline assessment were given as distractors. Post-test 2 was then assessed via online MCQs where immediate feedback was given.

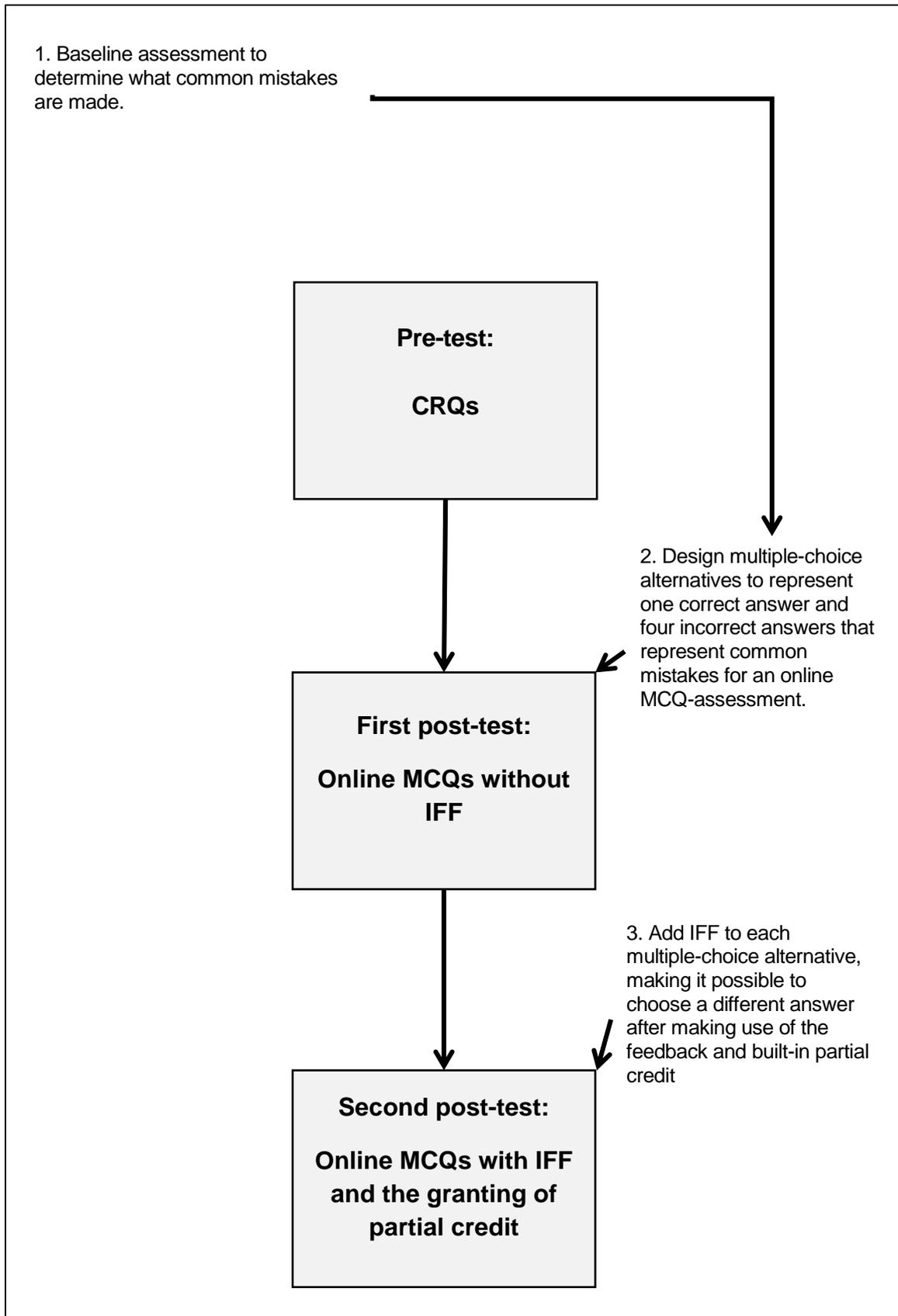


Figure 4.1: Graphical presentation of the research process

### **4.2.1 Reliability and validity**

Objectivity when marking MCQs make these much more reliable. This method of assessment is valid because no maturation or regression can take place.

#### *4.2.1.1 Reliability*

It is much more reliable to mark MCQs because the answer cannot be influenced by subjectivity, which entails more stability. The assessment was given to all four groups; and it was found that the outcomes for the students who had completed all three assessments were similar for all four groups. The data obtained from the one group were thus reliable. The single group was selected non-randomly through purposive sampling in the pre-experimental procedure.

#### *4.2.1.2 Validity*

The three assessments were written during the same period so that no time passed during the different tests of the experiment and no communication could influence the students' marks. No maturation or regression could therefore take place. In this research, only four questions could be used to collect the data in one period of 90 minutes. The validity could have been stronger if more questions were asked and the level of difficulty differed.

## **4.3 DATA ANALYSIS**

The analysis of the pre-and post-tests were done to make comparisons and to come to conclusions regarding the set hypotheses.

### **4.3.1 Data analysis procedures**

The data were analysed by calculating the average standard deviation and variance for each assessment. The standardised *AVERAGE*, *STDEVA* and *VAR* statistical functions in Microsoft Excel were used. Each assessment counted out of a total of 20 marks, which was converted to a percentage. The statistical description of the data is given in Table 4.1 below.

Table 4.1: Description of the marks for the three assessments

Assessment	Number of students	Average	Standard deviation	Variance
Pre-test: CRQs	40	58.925%	15.952	254.48
Post-test 1: Online MCQs without IFF	40	36.875%	24.011	576.522
Post-test 2: Online MCQs with IFF and the granting of partial credit	40	61.125%	17.704	313.446

### 4.3.2 Distribution analysis

The relative frequency distributions of the average marks for CRQs, online MCQs without IFF, and online MCQs with IFF and the granting of partial credit were created and are shown in Figure 4.2.

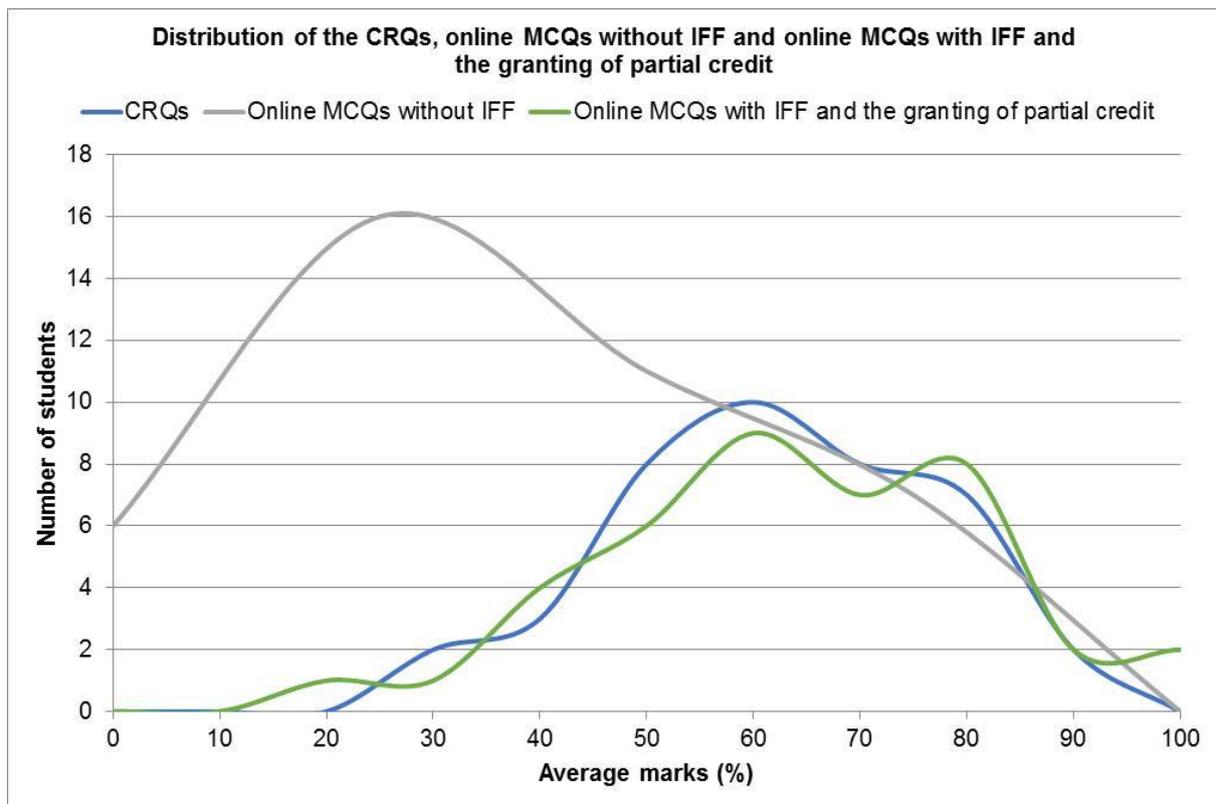


Figure 4.2: Distribution of the three assessments

The distribution of CRQs is indicated by the blue line in Figure 4.2. The distribution has a bell-like shape with a maximum frequency of 60%, which is relatively close to the average mark of 58.9%.

The grey line in Figure 4.2 indicates the distribution of online MCQs without IFF. The students could only score 0%, 25%, 50%, 75% or 100% for the assessment since no partial credit was given. Most of the students scored 25% for this assessment with the combined average for all the students being 36.9%. The distribution of online MCQs without IFF does not match the shape and slope of the distribution of CRQs. This leads to the conclusion that online MCQs without IFF are not an effective alternative to CRQs. This will be further analysed in Section 4.3.3.

The distribution of online MCQs with IFF and the granting of partial credit (indicated by the green line in Figure 4.2) have a maximum frequency of 60%, with 80% following closely thereafter. The shape and slope of the distribution seems close to the distribution of CRQs, which leads to the conclusion that online MCQs with IFF and the granting of partial credit could be an effective alternative to CRQs. This will be analysed further in Section 4.3.3.

### **4.3.3 The testing of the hypotheses**

Statistical inferences were made from the data obtained from the experiments by means of hypothesis testing. The pre- and post-tests were conducted independently from each other since the pre-test was written before the post-tests and the first post-test before the second. The students did not know their results from the pre-test when completing the post-tests, or the results of the first post-test when doing the second.

The dependent two-sample t-test was used to compare the average marks obtained by the experimental group in the two different types of assessment. The two-sample t-test depended on the marks to be normally distributed for each assessment. However, the normality of the distribution will not have an effect on the results of large groups (greater than 30) (see Appendix E for statistical tables).

The statistical symbols used for the hypothesis testing in this section are given in Table 4.2 below.

Table 4.2: Statistical symbols for hypothesis testing

Symbol	Description
$n_i$	Number of students who took assessment $i$
$\bar{Y}_i$	Sample average of assessment $i$
$\mu_i$	Population average of assessment $i$
$S_i^2$	Sample variance of assessment $i$
$\sigma_i^2$	Population variance of assessment $i$

Three hypothesis tests were conducted to answer the main research question.

### First Hypothesis

The average marks for CRQs ( $i = 1$  from Table 4.2) and online MCQs without IFF ( $i = 2$  from Table 4.2), when assessing problem-solving calculations were compared. A t-distribution was fitted to the data and the one-sided, two-sample t-test was used to determine if the average mark from the CRQs was significantly higher than that of online MCQs without IFF. The following hypothesis statements were postulated:

**H<sub>0</sub>:** There is no significant difference between the average mark for CRQs and online MCQs without IFF.

**H<sub>1</sub>:** The average of CRQs is significantly higher than the average of online MCQs without IFF.

A test statistic of 4.84 was calculated using  $T = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$ . The distribution could be

estimated with a standard normal distribution since the degree of freedom was 78.

A p-value of 0.000 001 298 was calculated to determine the level of significance for this test. The null hypothesis was *rejected* for all levels of significance  $> 0.000 001 298$  – also at a 5% level of significance. There was evidence that the average for CRQs was significantly higher than the average of online MCQs without IFF with 95% confidence.

## Second Hypothesis

The average marks for online MCQs with IFF and the granting of partial credit ( $i = 3$  from Table 4.2), and online MCQs without IFF ( $i = 2$  from Table 4.2) when assessing problem-solving calculations were compared. A t-distribution was fitted to the data and the one-sided, two-sample t-test was used to determine if the average mark of online MCQs with IFF and the granting of partial credit was significantly higher than that of online MCQs without IFF. The following hypothesis statements were postulated:

**H<sub>0</sub>:** There is no significant difference between the average for online MCQs with IFF and the granting of partial credit and online MCQs without IFF.

**H<sub>1</sub>:** The average of online MCQs with IFF and the granting of partial credit is significantly higher than the average of online MCQs without IFF.

A test statistic of 5.14 was calculated using  $T = \frac{(\bar{Y}_3 - \bar{Y}_2) - (\mu_3 - \mu_2)}{\sqrt{\frac{S_3^2}{n_3} + \frac{S_2^2}{n_2}}}$ . The distribution could be

estimated with a standard normal distribution since the degree of freedom was 78.

A p-value of 0.000 000 287 was calculated to determine the level of significance for this test. The null hypothesis was *rejected* for all levels of significance  $> 0.000 000 287$  – also at a 5% level of significance. There was evidence that the average for online MCQs with IFF and the granting of partial credit was significantly higher than the average of online MCQs without IFF with 95% confidence.

## Third Hypothesis

The average marks for CRQs ( $i = 1$  from Table 4.2) and online MCQs with IFF and the granting of partial credit ( $i = 3$  from Table 4.2) when assessing problem-solving calculations were compared. A t-distribution was fitted to the data and a two-sided, two-sample t-test was used to determine if the average mark of CRQs was significantly different to that of online MCQs with IFF and the granting of partial credit. The following hypothesis statements were postulated:

**H<sub>0</sub>:** There is no significant difference between the average for CRQs and online MCQs with IFF and the granting of partial credit.

**H<sub>1</sub>:** The average of CRQs is significantly different to that of online MCQs with IFF and the granting of partial credit.

A test statistic of -0.58 was calculated using  $T = \frac{(\bar{Y}_1 - \bar{Y}_3) - (\mu_1 - \mu_3)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_3^2}{n_3}}}$ . The distribution could be

estimated with a standard normal distribution since the degree of freedom was 78.

A p-value of 0.5594 was calculated to determine the level of significance for this test. The null hypothesis was *not rejected* for any level of significance < 0.5594 – also at a 5% level of significance. There was, therefore, not a significant difference between the average marks for CRQs and online MCQs with IFF and the granting of partial credit at a 95% confidence level.

The difference between the average for CRQs and online MCQs with IFF and the granting of partial credit was -2.2% (58.9% - 61.1%). This difference seems insignificant since the null hypothesis above was not rejected. This finding was checked to confirm that the difference was, in fact, significantly small. A 95% confidence interval was constructed around the difference to determine what size the difference could be for the null hypothesis not to be rejected at a 5% level, as well as the probability of committing a type II error (to not reject the null hypothesis where the alternative is true) with this decision.

The 95% confidence interval for the difference was calculated as follows:

$$(\bar{Y}_1 - \bar{Y}_3) - 1.96 \sqrt{\frac{S_1^2}{n_1} + \frac{S_3^2}{n_3}} \leq (\mu_1 - \mu_3) \leq (\bar{Y}_1 - \bar{Y}_3) + 1.96 \sqrt{\frac{S_1^2}{n_1} + \frac{S_3^2}{n_3}}$$

$$-9.59 \leq (\mu_1 - \mu_3) \leq 5.19$$

The probability of wrongly concluding (on a 95% confidence level) that the difference of the averages between CRQs and online MCQs with IFF and the granting of partial credit was not significant if the difference was greater than -9.89 or 5.19, was 0.0783. The online MCQs with IFF and the granting of partial credit could therefore be used as an effective alternative to CRQs.

#### 4.4 DATA INTERPRETATION

The research sub-question, “*Can online MCQs without IFF replace CRQs in fairly assessing students’ ability to do problem-solving calculations in science?*”, is answered by the **first hypothesis** in Section 4.3.3. The null hypothesis was rejected on a 95% level of confidence and the analysis showed that the average for CRQs was significantly higher than the average of online MCQs without IFF. Online MCQs without IFF can therefore not be used as an effective alternative to CRQs to assess problem-solving calculations in science.

The next research sub-question, “*Are online MCQs with IFF, and the granting of partial credit, more reliable and fair than MCQs without IFF in assessing students’ ability to do problem-solving calculations in science?*”, is answered by the **second hypothesis** in Section 4.3.3. The null hypothesis was rejected on a 95% level of confidence and the analysis showed that the average for online MCQs with IFF and the granting of partial credit was significantly higher than the average of online MCQs without IFF. MCQs with IFF and the granting of partial credit are therefore more reliable than MCQs without IFF.

Finally, the research sub-question, “*Can online MCQs with IFF and the granting of partial credit be used to assess problem-solving calculations as a fair and valid alternative to CRQs?*”, is answered by the **third hypothesis** in Section 4.3.3. The null hypothesis was not rejected at a 95% level of confidence. This finding was checked to confirm that the difference was significantly small. A 95% confidence interval was therefore constructed to identify the value of the difference between the average of CRQs and online MCQs with IFF and the granting of partial credit, for which the null hypothesis would not be rejected. The null hypothesis could not be rejected for a difference of between -9.59 and 5.19 on a 95% confidence level and the subsequent probability of not rejecting the null hypothesis mistakenly was calculated at 7.8%. Online MCQs with IFF and the granting of partial credit could therefore be used as an effective alternative to CRQs.

#### 4.5 CONCLUDING REMARKS

The outcome of the baseline assessment was used to set up the distractors for the online MCQ with IFF and the granting of partial credit. The researcher is of the opinion that the

more lecturers available to contribute to the database, the better it would become, containing more examples of common mistakes made by students.

According to the analysis of the data, online MCQs with IFF and the granting of partial credit may be more reliable than MCQs without IFF in doing problem-solving calculations. The marks allocated for problem-solving calculations assessed through MCQs with IFF and the granting of partial credit could be an effective alternative to problem-solving calculations assessed by means of CRQs. The IFF allows students to choose a different answer based on the corrective feedback because partial credit is now granted for correct second, third or fourth attempts.

The students would be able to learn from common mistakes and improve their learning. MCQs with IFF and the granting of partial credit could also be used as a class assessment method.

The lecturer could also benefit if students could do problem-solving calculations with IFF and the granting of partial credit online because time and money would be saved with this type of assessment as it is marked by computer.

#### **4.6 SUMMARY**

In this chapter, the analysis and results of this study were explained. The research questions were analysed in detail and were answered satisfactorily. This led to the conclusion that MCQs could be designed to provide IFF and partial credit to students and could possibly be used as an effective alternative to CRQs in the assessment of problem-solving calculations for Engineering Science students.

The research process was discussed to explain the process followed in this research. The reliability and validity of this research were also discussed to show its trustworthiness.

The mistakes made by the students in the baseline assessment were used as the distractors for the MCQs. In the experiment, the average of the CRQs and the average of the MCQs with IFF and the granting of partial credit were close to each other, with 99% confidence. The MCQs with IFF and the granting of partial credit could therefore be used as an effective alternative to CRQs to assess students with problem-solving calculations.

# **CHAPTER 5: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS**

## **5.1 INTRODUCTION**

In this chapter, the conclusions regarding the topic under study are made in reference to the research aims and objectives, and in alignment with the findings. The limitations of this study are mentioned and the implications of this study for future research are presented. These relate to how online MCQs can be used to assess problem-solving calculations as long as they make provision for IFF and the granting of partial credit. Based on the findings of this study, recommendations are made for assessment practice.

## **5.2 SUMMARY OF THE RESEARCH FINDINGS**

The key findings from the literature review and the empirical study will be discussed in this section.

### **5.2.1 Key findings from the literature study**

In this study, the use of online MCQs with IFF and the granting of partial credit for problem-solving calculations were explored. From the literature study, it became evident that e-learning is becoming more popular with its associate reliance on technology. The use of online MCQ-assessments is an example of the utilisation of online technology in assessment.

The consequence is a shift from the use of traditional assessment to solve calculations towards the extended use of MCQs as assessment method. The literature indicated that many assessors are not yet comfortable using online MCQs to assess problem-solving calculations (Section 2.2.6). They support the necessity for an assessment to be set up consisting of both CRQs as well as MCQs.

It was found that the actual value of formative assessment is only visible when students lose marks and get feedback on how to improve their performance (Section 2.2.1).

There is evidence that feedback is the most positive single influence on learning, motivation and success. The sooner the feedback is given, the more the students can benefit from it (Section 2.2.1.2). The basic reason for the use of formative assessment is that feedback has a useful positive effect on student learning, inspiration and self-confidence. Students understand and remember the feedback the best when it is immediately given since each phase of the learning process builds on the previous phase. If no feedback is given, the next phase could be built on misconceptions. The process of moving from the traditional way of assessment towards online assessment would initially be time consuming until a database is set up (Section 2.2.1).

MCQs and CRQs were classified as the two main methods of formative assessment. MCQs are currently a very popular method of assessment, and possess several benefits, such as automated marking, reliable scores, reduced marking time, and consistency between different assessors. The most important advantage of MCQs is that it can be done online, which will save time and can be done anytime by many students (Section 2.2.3).

When MCQs are done online, cheating can be controlled by randomising the questions and there is no subjectivity from the examiner when marking. This makes MCQs fair and easy to mark. Constructing MCQs to assess the higher level of cognitive skills is a very difficult task and it needs investment in time to set up proper questions. When it comes to the marking of assessments, the return of this investment is seen (Section 2.2.3).

The design of effective distractors for different cognitive levels is also a huge challenge. If distractors are not functional, this could have an effect on the validity and reliability of the assessment. It could also lead to choosing the correct answer for the wrong reasons and can have a negative influence on the outcome. Those students who choose the correct alternative must be acknowledged. In an effort to overcome the guessing temptation, the two-tier test was introduced. The first tier is similar to traditional MCQs and the second tier gives a set of reasons for the given answer in the first tier. Negative marking was also suggested as an alternative. Another option is to set different marks for more challenging questions, or to give more distractors for each question (Section 2.2.2).

Online MCQs are used to give IFF without partial credit, e.g. Key train and Moodle. The students can get their results with feedback immediately after completion, but no partial credit is given. Immediate formative assessment techniques (IFAT) can currently be used to give partial credit to students, but without corrective IFF. This emphasises that students can learn from their wrong answer and that IFAT® forms contribute to significant learning (Section 2.2.8).

From the literature, it was established that there was a need for using online MCQs to assess problem-solving calculations. Such MCQs, however, have to be developed as they need to include immediate feedback during the assessment process. Recalculation by making use of multiple opportunities, corrective feedback, and the granting of partial credit lead students to choose the correct answer. This study was undertaken to investigate this possibility.

### **5.2.2 Key findings from the empirical study**

Three sets of findings will be put forward in this section:

CRQs were taken (for the purpose of this study) as an assessment method that measures students' ability accurately and fairly. Another popular way of assessing students' ability is using online MCQs without IFF. This assessment method has many advantages (Section 2.2.3) that make this an attractive assessment method to use due to its efficiency and objectivity. It is, however, indistinct from the literature if this assessment method can be used as an effective alternative for CRQs in assessing problem-solving calculations in science. This was tested under the first hypothesis (Section 4.3.3) where the null hypothesis was rejected and it was concluded that the average for CRQs was significantly higher than for online MCQs without IFF on a 95% confidence level. Online MCQs without IFF cannot replace CRQs in fairly assessing students' ability to do problem-solving calculations in science, and the second research sub-question was answered.

An assessment method was proposed by the researcher for the purpose of combining the efficiency and objectivity advantages of online MCQs without IFF with the accuracy and fairness of CRQs to assess problem-solving calculations in science. IFF was added to online MCQs without IFF, and partial credit was granted by allowing students to apply the IFF to the question with multiple opportunities to choose an answer for partial marks at every additional opportunity. Online MCQs with IFF and the granting of partial credit

were firstly compared to online MCQs without IFF to determine if the students' marks improved. The null hypothesis of the second hypothesis in Section 4.3.3 was rejected at a 95% confidence level and it was concluded that the average for MCQs with IFF and the granting of partial credit was indeed higher than that of MCQs without IFF. The third research sub-question therefore concluded that online MCQs with IFF and the granting of partial credit are a more reliable and a fair method to assess students' ability to do problem-solving calculations in science.

The researcher was satisfied that online MCQs without IFF could not replace CRQs when assessing students' ability to do problem-solving calculations in science and that online MCQs with IFF and the granting of partial credit were more reliable than online MCQs without IFF. In order to answer the last research sub-question, and hence also the main research question, the researcher needed to determine if online MCQs with IFF and the granting of partial credit could be an effective alternative to CRQs in assessing problem-solving calculations in science. The third hypothesis (Section 4.3.3) was used to test if there was a significant difference between the average marks for CRQs and online MCQs with IFF and the granting of partial credit. This hypothesis could not be rejected at a 95% confidence level and it was concluded for the fourth research sub-question that online MCQs with IFF and the granting of partial credit could be used as an effective alternative to CRQs.

The findings align with the theoretical framework as follows:

- It is necessary and possible to use online MCQs with IFF and partial credit to assess at higher cognitive levels, as suggested by Bloom (1968);
- Online MCQs with IFF and partial credit still make it possible for students with a higher ability to be awarded for finding the correct answer on their first try in alignment with the IRT model (Lord & Novic, 1968);
- Online MCQs with IFF and partial credit are particularly well-suited to the Feedback Interventions (FIs) Theory (Kluger & DeNisi, 1996:67-72) because IFF and the granting of partial credit will increase task motivation; and
- Online MCQs with IFF and partial credit were found to be suitable for assessing at the problem saving level, as suggested in the assessment component taxonomy of Huntley et al. (2011:3-16).

### 5.3 RESEARCH CONCLUSIONS

The research question formulated to address the purpose of this study is:

***Can online multiple-choice questions, designed to provide immediate formative feedback with the granting of partial credit, be used as an effective alternative for constructed written questions in the assessment of problem-solving calculations for Engineering Science students?***

The following sub-questions were formulated to answer the research question:

- **When setting online MCQs, how can an assessor determine which distractors would anticipate the mistakes that students could make so that formative feedback can be built in to indicate to them what they did wrong and how to correct it?**

A baseline assessment was done to identify the most common errors that students make when completing step-by-step constructed assessments for problem-solving calculations. Lecturers can also add to the identified common mistakes based on their past experience. The more times the calculations are repeated, the more refined the distractors may become when students do the same or similar calculations in answering online MCQs.

- **Can online MCQs without IFF replace CRQs in fairly assessing students' ability to do problem-solving calculations in science?**

Currently, online MCQs without IFF cannot replace CRQs because neither partial credit nor IFF are given. It was found that students answering CRQs do much better than when answering MCQs without IFF. To answer using MCQs without IFF, a student can only have 1 or 0 as an outcome. The analysis showed that the average for CRQs was significantly higher than the average of online MCQs without IFF. Online MCQs without IFF can therefore not be used as an effective alternative to CRQs to assess problem-solving calculations in science (Section 4.3.3).

- **Are online MCQs with IFF, and the granting of partial credit, more reliable and fair than MCQs without IFF in assessing students' ability to do problem-solving calculations in science?**

If immediate formative feedback is given, students can see what they did wrong during the assessment process. This could help students to learn from their mistakes immediately and to recalculate and correct their mistakes based on the corrective feedback they receive. The granting of partial credit for the corrected answers enables them to learn from their mistakes and to improve their learning gains. The analysis showed that the average for online MCQs with IFF and the granting of partial credit was significantly higher than the average of online MCQs without IFF. Online MCQs with IFF and the granting of partial credit are therefore more reliable and fair than online MCQs without IFF (Section 4.3.3).

- **Can online MCQs with IFF and the granting of partial credit be used to assess problem-solving calculations as a fair and valid alternative to CRQs?**

The researcher discovered that online MCQs without IFF cannot replace CRQs when assessing students' ability to do problem-solving calculations in science. It was also shown that online MCQs with IFF and the granting of partial credit were more reliable than online MCQs without IFF. The third hypothesis (Section 4.3.3) was used to test if there was a significant difference between the average marks for CRQs and online MCQs with IFF and the granting of partial credit. This hypothesis could not be rejected at a 95% confidence level, and it was concluded that online MCQs with IFF and the granting of partial credit could be used as an effective alternative to CRQs.

## **5.4 RECOMMENDATIONS**

For MCQ-assessments with IFF and the granting of partial credit to be successfully implemented as an alternative to MCQs in assessing problem-solving calculations, a few recommendations should be considered:

**Recommendation 1:** Lecturers must take note of the common errors that students make when completing step-by-step written problem-solving calculation assessments.

It is proposed that lecturers should collect these common mistakes over a period of time. It might even require that the same assessment be given for a few years to compare the different answers from the different cohorts to find common mistakes.

**Recommendation 2:** Lecturers must work together in a team and use the same assessments on similar calculations to be able to identify the most common errors that students make. This will enrich and refine the collection of common mistakes made when doing calculations.

**Recommendation 3:** All of the data gathered from different lecturers must be coordinated to set up a proper electronic database of common mistakes associated with different types of calculations.

**Recommendation 4:** This database can then be utilised to decide on a set of distractors for a specific MCQ that would be aligned with the most common mistakes that students make.

**Recommendation 5:** Formative feedback should be composed to be immediately available during the assessment process to assist the students in correcting the mistake that led them to arrive at a particular wrong answer. It should be available as soon as the wrong answer is chosen so that students can recalculate based on the corrective feedback. This would then enable them to arrive at a different answer. Multiple-choices should be available in case another common mistake is made, and students would be afforded additional opportunities to correct their mistakes based on the corrective feedback for each alternative answer.

**Recommendation 6:** The proposed methodology set out in the above recommendations could be utilised to develop viable, valid and fair online MCQ-assessments for science or mathematics calculations. Such assessments are much cheaper in terms of time and labour costs as they do not require marking to be done by lecturers.

**Recommendation 7:** Online MCQs with IFF and the granting of partial credit are eminently suitable for, and should be used when, assessing unlimited numbers of students. A further important benefit is that it is flexible in its application for different formative purposes, such as using it to diagnose understanding and do revision, class

tests and examinations. Its most important attribute, however, is that it allows learners to learn from their mistakes while doing the assessment (which is the best time of all to do so). It also assists in improving their understanding of calculations, which should eventually result in improved learning gains for students.

It is important to note that the proposed MCQ-assessments with IFF and the granting of partial credit to assess calculations are not intended as replacements for all CRQ-assessments, but can be an alternative in certain circumstances where classes are very big or where lecturers have a heavy workload that makes it difficult to also mark assessments. One can, of course, also combine MCQ and CRQ-assessments or use CRQ-assessments only where it is preferred.

## **5.5 AVENUES FOR FURTHER RESEARCH**

Although it was found that MCQ-assessments with IFF and the granting of partial credit can be utilised as a fair and valid alternative to CRQs, there are some challenges that remain and require research, such as:

- Guessing is still a factor that may influence results;
- It is not possible to make provision for all possible mistakes in the distractors, which limits formative feedback and learning from mistakes, to an extent; and
- A study designed to compare students of different academic levels is recommended in order to further investigate the effect of the students' level of knowledge on the use of the proposed MCQ-assessments. The use of the proposed MCQ-assessments with IFF and the granting of partial credit was only researched in terms of calculations and not in regard to other types of learning on higher cognitive levels of Bloom's Revised Taxonomy.
- Lastly, the usefulness of the feedback associated with each distractor on further student success in arriving at the correct answer was not investigated and is a recommended research topic for further studies.

## **5.6 LIMITATIONS OF THIS STUDY**

Online MCQ-assessment, as researched in this study, and immediate feedback to assist the learning of students regarding problem-solving calculations indicates promising

potential for use in practice. There are, unfortunately, important limitations in this study that must be taken into account:

- The sample of this study was only 40 students. The small sample size resulted in a large standard deviation of the means for this research. If this study could be done with a bigger sample group, the findings would be more reliable.
- The feedback provided is, unfortunately, limited because it is not possible to make provision for all possible calculation errors that students could make.
- Guessing still plays a part in MCQ-assessments, which can skew results.
- Only four questions were used to collect the data in one series of assessment. More questions would increase the reliability of the results.

It is therefore suggested that these limitations should be taken into account when replicating this study.

## **5.7 CONCLUDING REMARKS**

The research, although limited in scope, gave a definitive answer to the main research question. It is possible to use online MCQs with IFF and the granting of partial credit as an alternative to CRQs in problem-solving calculations.

From the data analyses, the research question could definitely be answered. Therefore, online MCQs designed to provide immediate formative feedback with the granting of partial credit to students, can be used as an effective alternative to CRQs in the assessment of problem-solving calculations for Engineering Science students.

It is the opinion of the researcher that this research made a contribution to the field of MCQ-assessments, and that the gap in knowledge regarding effective MCQ-assessments for use in assessing problem-solving calculations has been addressed.

## REFERENCES

- Alder, G.S. 2007. Examining the relationship between feedback and performance in a monitored environment: A clarification and extension of feedback intervention theory. *Journal of High Technology Management Research*, 17:157-174.
- Alford, G.H., Herbert, P. & Frangenheim, E. 2006. *Bloom's Taxonomy Overview, Innovative Teachers Companion*. Toowomba: ITC Publications.
- Ali, S.H., Carr, P.A. & Ruit, K.G., 2016. Validity and Reliability of Scores Obtained on Multiple-Choice Questions: Why Functioning Distractors Matter. *Journal of the Scholarship of Teaching and Learning*, 16(1):1-14.
- Al-Sadi, J. & Al-Halabi, D. 2010. MCQ Exams Correction in an Offline Network using XML. *GSTF International Journal on Computing*, 1(1):176-182.
- Aubrecht, I.I., Gordon, J. & Aubrecht, J. 1983. Constructing Objective Tests. *American Journal of Physics*, 51:613.
- Babbie, E.R. 2007. *The practise of social research*. 11<sup>th</sup> ed. Belmont: Wadsworth/Thomas.
- Barnard, J. 2013. Die beginpunt van meerkeusige vrae. *SA tydskrif vir Natuurwetenskap en Tegnologie*, 32(1):172.
- Benvenuti, S. 2010. Towards a more authentic approach to assessment. An innovative response to large class assessment.. In: D. L. Lange (Ed.) *Teaching and learning* (p.1). Pretoria: Council on Higher Education.
- Benvenuti, S. & Cohen, J. 2008. An In-Depth View of MCQ-based Testing in Information Systems. *South African Computer Journal*, 4:21.
- Berry, J.W. & Chew, S.L. 2008. Improving learning through interventions of student-generated questions and concept maps. *Teaching of Psychology*, 35(4):305-312.
- Biju, I. & Seibu, M. J. 2010. Effects of Online MCQ Tests on Students Learning. In: K.K. Iskander (Ed.) *Technological Developments in Education and Automation* (pp. 403-408). Netherlands: Springer.

- Bishop, R.L. 1990. Multiple-choice questions. *The teaching of Astronomy*: 83.
- Black, P., Harrison, C., Marshall, B. & Wiliam, D. 2004. Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan*, April 1986(1):8-21.
- Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assesment in Education: Principles, Policy and Practice*: pp. 7-74.
- Black, P. & Wiliam, D. 2003. In praise of educational research: Formative assessment. *British Educational Research Journal*, 29(5):623-637.
- Black, P. & Wiliam, D. 2004. The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education*, 103(2):20-50.
- Bless, C. & Higon-Smith, C. 1995. *Fundamentals of social research methods: an African perspective*, 2<sup>nd</sup> ed. Kenwyn: Juta & co.
- Bloom, B. 1969. Some theoretical issues relating to educational evaluation. In: *Educational evaluations: New roles, new means*. Chicago: University of Chicago Press.
- Boud, D. & Soler, R. 2016. Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41(3):400-413.
- Boud, D. 1995. Assessment and Learning: Contradictory or complementary? In: P. Knight (Ed.) *Assesment for Learning in higher education* (p.35). London: Kogan Page Ltd.
- Boud, D. 2000. Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22:151-167.
- Brame, C.J. 2014. *Writing Good Multiple-choice Test Questions*. [Online]. Available at: <http://cft.vanderBilt.edu/guides-subpages/writing-good-multiple-choice-test-questions>. [Accessed 01 December 2015].
- Brandl, K., 2005. Are you ready to "Moodle". *Language, Learning and Technology*, 9(2):16-23.

- Bridgman, B. 1992. A comparison of Quantitative questions in open-ended and Multiple-Choice format. *Journal of Education Measurement*, 29:253-271.
- Brosvic, G.M., Epstein, R.A. & Cook, M. 2006. Aquisition and retension of esperanto: the case for error correction and immediate feedback.. *The Psychological record*, 56:205-218.
- Buchanan, T. 2000. The efficacy of a world-wide web mediated formative assesment. *Journal of Computer Assisted Learning*, 16:193-200.
- Burns, N. & Grove, S.K. 2011. *The practice of nursing research: conduct, critique & utilization*. 4<sup>th</sup> ed. Philadelphia: Saunders.
- Bush, M.E. 2006. Quality assurance of multiple-choice tests.. *Quality Assurance in Education*, 14(4):398-404.
- Caledon, I. & Subramaniam, R. 2010. Development and Application of a Three-Tier Diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, 32(7):939-961.
- Canfield, W. 2001. ALEKS: a web-based intelligent tutoring system. *Mathematics and Computer Education*, 35(2):152-158.
- Chadwick, B.A., Bahr, H. & Albrecht, S.L. 1984. *Social science research methods*. New Jersey: Prentice-Hall.
- Chiheb, R., Faizi, R. & Afia, A.E. 2011. Using objective online testing tools to assess students' learning: Potentials and limitations. *Journal of Theoretical and Applied Information Tehnology*, 24(1):69-72.
- Clariana, R.B., Ross, S.M. & Morrison, G.R. 1992. The effects of different strategies using computer administered multiple-choice questions as instruction. *Educational Technology Research and Development*, 39(2): pp. 5-17.
- Clariana, R.B., Ross, S.M & Morrison, G.R. 1991. The effect of different feedback strategies using Computer-Administered Multiple-Choice Questions as Instruction. *Educational Technology Research and Development.*, 39(2):5-17.

- Coyne, K., 2015. Using technology to check for understanding. *Massachusetts Music News*, 63(4):24-26.
- Creswell, J.W. 2009. *Research Design*. 3<sup>rd</sup> ed. London: Sage publications.
- Creswell, J.W. 2012. *Research design: qualitative, quantitative and mixed methods approaches*. 4<sup>th</sup> ed. USA: Sage Publications Asia-Pacific Pty Ltd.
- De Lange, P., Suwardy, T. & Mavondo, F. 2003. Intergrating a virtual learning environment into an introductory accounting course; determinants of student motivation. *Accounting Education*, 12(1):1-14.
- De Vos, A.S., Strydom, H., Fouche, C.B. & Delport, C.S. 2011. *Research at grass roots*. 4<sup>th</sup> ed. Pretoria: Van Schaiks.
- Dempsey, J. & Litchfield, B. 1993. Feedback, retention, discrimination error, and feedback study time. *Journal of Research on Computing in Education*, 25(3):303-327.
- Dempsey, J.V. & Driscoll, M. 1989. *The effects of four methods of immediate corrective feedback on retention discrimination error, and feedback study time in computer based instruction*. San Francisco, Annual Meeting of the American Educational Research Association..
- Dempsey, J. & Wager, S. 1988. A taxonomy for the timing of feedback in computer-based instruction. *Educational Technology*, 28(10):20-25.
- Dempster, E.R. 2007. Textual strategies for answering multiple-choice questions among South African learners: what can we learn from TIMMS 2003? *African Journal of Research in SMT Education*, 11(1):47-60.
- DiBattista, D. & Gosse, L., 2006. Test anxiety and the immediate feedback assessment technique. *The Journal of Experimental Education*, 74(4):311-328.
- Einig, S. 2013. Supporting students' learning: The use of formative online Assessments. *Accounting Education*:847-848.
- Engelbrecht, J. & Harding, A. 2003. Online assessment in Mathematics: Multiple assesment formats. *New Zealand Journal of Mathematics*, 32:57-65.

- Engelbrecht, J. & Harding, A. 2004. Combining online and paper assessment in a web-based course in undergraduate Mathematics. *Journal of Computer, Mathematics and Science Teaching*, 23(3):217-231.
- Epstein, M.L., Lazarus, A.D., Calvano, T.B., Matthews, K.A., Hendel, R.A., Epstein, B.B. & Brosvic, G..M. 2002. Immediate feedback assessment technique promotes learning and corrects inaccurate first responses. *The Psychological Record*, 52:187-201.
- Epstein, M.L., Epstein, B.B. & Brosvic, G.M. 2001. Immediate feedback during academic testing. *Psycho-logical Reports*, 88:889-894.
- Escudier, M.P. 2011. University students' attainment and perceptions of computer delivered assessment; a comparison between computer-based and traditional tests in a 'high-stakes' examination. *Journal of Computer Assisted Learning*, 27(5): pp. 440-447.
- Fellenz, M.R. 2004. Using assessment to support higher level learning: The multiple-choice item development assignment. *Assessment & Evaluation in Higher Education*, 29(6):703-719.
- Foong, M.Y. 2015. Supplemental materials for use of constructed-response-questions to support learning of cell biology during lectures. *Journal of Microbiology & Biology Education*, 16(1):87-89.
- Friel, K. & Johnstone, A. 1978. Scoring systems which allow for partial knowledge. *Journal for Chemistry Education*, 55(11):717-719.
- Gay, S. & Thomas, M. 1993. Just because they got it right, does it mean they know it? *Assessment in the Mathematics classroom*. Reston, VA. The National Council of Teachers of Mathematics.
- Geyser, H. 2004. Learning from assessment. *Teaching and learning in higher education*, Pretoria: van Schaik.
- Gibbs, G. & Simpson, C. 2005. Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1:3-31.

- Gogri, M.H., Shaikh, A. & Venugopal, A. 2013. Evaluation of Students performance based on formative assessment using data mining. *International of Computer Applications*, 67(2):89-93.
- Gomes, A. & Mendes, A. 2009. Bloom's taxonomy based approach to learn basic programming. *World Conference on Educational Multimedia, Hypermedia and Telecommunication*, 2009(1):2547-2554.
- Gretton, H. & Challis, N. 1999. *Does the punishment fit the crime?* San Francisco, Proceedings of the International Conference on Technology in Education.
- Guo, R., Palmer-Brown, D., Lee, S.W. & Cai, F.F. 2014. Intelligent diagnostic feedback for online multiple-choice questions. *Artificial Intelligence Review*, 42(3):369-383.
- Hadsell, L. 2009. The effect of quiz-timing on exam performances. *Journal of Education for Business*, 84(3):135-141.
- Haladyna, T.M. 2011. *Developing and Validating Multiple-choice questions*. 3<sup>rd</sup> ed. London and New York: Routledge.
- Hambleton, R.K., Swaminathan, H. & Rogers, H. 1991. *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage Press ed.
- Hambleton, R. & Swaminathan, H. 2013. *Item response theory: Principles and applications*. Netherlands: Springer Science & Business Media.
- Hamilton, L., 1994. *An Investigation of students' affective responses to alternative assessment formats*. New Orleans: National Council on Measurement in Education.
- Harper, R. 2003. Correcting computer-based assessments for guessing. *Journal of Computer Assisted learning*, 19:3-8.
- Hattie, J.A. 1987. Identifying the salient facets of a model of student learning: a synthesis of meta-analyses. *International Journal of Educational Research*, 11:187-212.
- Hattie, J.A. & Timperley, H. 2007. The power of feedback. *Review of Educational Research*, 81:81-112.

- Heaslip, G., Donovan, P. & Cullen, J. 2014. Student response systems and learner engagement in large classes. *Active learning in higher education*, 15(1):11-24.
- Heirdsfield, A., Walker, S., Tambyah, M. & Beutel, D. 2011. Blackboard as an online learning environment: What do teacher education students and staff think? *Australian Journal of Teacher Education (Online)*, 36(7):1-16.
- Hewson, C. 2012. Can online course-based assessment methods be fair and equitable? Relationships between students' preferences and performance within online and offline assessments. *Journal of Computer Assisted Learning*, 28(5):488-498.
- Hift, R. 2014. Should essays and other "open-ended" type questions retain a place in written summative assessments in clinical medicine? *BMC Medical Education*: 249.
- Higgins, E. & Tatham, L. 2003. Exploring the potential of Multiple-Choice Questions in Assessment. Multiple-Choice questions in Assessment. *Learning and Teaching in Action*, 2(1):1-11.
- Htwe, T., Ismail, T.B. & Low, L. 2014. Comparative assessment of studnets' performance and perceptions on objective structured practical models in undergraduate pathology teaching. *Singapore Med*, 9(55):502.
- Hughes, S.J. & Quinn, F. 2013. *Quinns Principles and Practice of Nurse Education*. 4<sup>th</sup> ed. Andover: Cengage Learning.
- Huntley, B., Engelbrecht, J. & Harding, A. 2010. A Model for measuring the quality of a mathematics question. *Researchgate*, 5(2):141-171.
- Huntley, B., Engelbrecht, J. & Harding, A. 2011. Can multiple-choice questions be successfully used as an assessment format in undergraduate mathematics? *Pythagoras*, 69:3-16.
- Hyland, P. & Booth, P. 2000. Learning from feedback on assessment. *The Practice of University History Teaching*:233-247.

- Ilgen, D.R., Fisher, C.D. and Taylor, M.S., 1979. Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), p.349.
- Jonanovic, N. & Aburaghiem, L. 2006. Use Multiple-choice questions, short questions and exercises to assess student performance. *Acta Academia-University of the Free State*, 38(1):233-252.
- Jordan, B. 2012. *Advancing Ethnography in Corporate environments. Challenges and emerging opportunities*. 4<sup>th</sup> ed. Walnut Creek, CA: Left Coast Press.
- Karandikar, R.L. 2010. On multiple-choice tests and negative marking. *Current Science*, 99(8):1042-1045.
- Kastner, M. & Stangl, B. 2011. Multiple-choice and Constructed Response Tests: Do test format and scoring matter? *Procedia Social and Behavior Sciences*, 12:263-273.
- Kearns, L.R. 2012. Student assessment in Online learning. In: S. Alman, C. Tomer & M.L. Lincoln (eds.) *Designing online learning* (p.43). Oxford, England: Libraries Unlimited.
- Kluger, A. & DeNisi, A. 1996. The effects of feedback interventions on performance: a historical review, a meta-analysis and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2):256-284.
- Kluger, A.N. & DeNisi, A. 1998. Toward the understanding of a Double-Edge Sword. *Current Directions in Psychological Science*, 7(3):67-72.
- Kuechler, W.L. & Simkin, M.G. 2003. How well do Multiple-choice Tests evaluate student understanding in Computer Programming classes? *Journal of Information Systems Education*, 14(4):389.
- Kutluay, Y. 2005. *Diagnosis of eleventh grade students misconceptions*. Unpublished Master's dissertation. Istanbul: Middle East Technical University.
- Leedy, P.D. & Ormond, J. 2005. *Review of the relates literature*. New Jersey: Pearson.
- LeFebvre, M. & Buddin, R. 2014. *Keytrain*. [Online]. Available at: [www.keytrain.com](http://www.keytrain.com)

- Lesage, E., Valcke, M. & Sabbe, E., 2013. Scoring methods for multiple choice assessment in higher education—Is it still a matter of number right scoring or negative marking? *Studies in Educational Evaluation*, 39(3), pp.188-193.
- Lincoln, Y.S. & Gub, E.G. 1985. *Naturalistic inquiry*. 1<sup>st</sup> ed. California: Sage.
- Lister, R. 2000. On blooming first year programming, and its blooming assessment. *Proceedings of the Australasian conference on Computing education*, pp. 158-162.
- Loh, A.S.L., Subramaniam, R. & Tan, K. 2014. Exploring students' understanding of electrochemical cells using enhanced two-tier diagnostic instrument. *Science and Technological Education*, 3:229-250.
- López-Pastor, V. M., Pintor, P., Muros, A. & Webb, G. 2012. Formative assessment strategies and their effect on student performance and on student and tutor workload. *Journal of Further and Higher Education*, 37(2):163-180.
- Lord, F. M, Novick, M.R. & Birnbaum, A. 1968. *Statistical theories of mental test scores*.
- Lukhele, R., Thissen, D. & Wainer, H. 1994. On the relative value of Multiple-Choice, Constructed Response and Examinee-selected items on Two Achievement Tests. *Journal of Educational Measurement*, 31(3):234-250.
- Maier, U., Wolf, N. & Randler, C. 2016. Computers and Education. *Elsevier*, 95:85-98.
- Malau-Aduli, B.S., Assenheimer, D., Choi-Lundberg, D. & Zimitat, C. 2013. Using computer-based technology to improve feedback to staff and students on MCQ assessments. *Innovation in Education and Teaching International*, 51(5):510-522.
- Maree, K., Creswell, J.W., Eloff, I., Ferreira, R., Ivankova, N., Jansen, J.D., Nieuwenhuis, J., Pietersen, J., Plano Clark, V.L., 2011. *First steps in research*. 7<sup>th</sup> ed. Pretoria: Van Schaik.
- Marriott, P. & Lau, A. 2008. The use of online summative assessment in an undergraduate financial accounting course. *Journal of Accounting Education*, 26:73-90.

- Marx, J.D. 1998. *Creation of a diagnostic exam for introductory, undergraduate electricity and Magnetism*. New York: Dissertation Abstracts International.
- Mathews, J. 2006. Just whose idea was all this testing? *The Washington Post*:A6.
- Maxwell, J.A. 2013. *Qualitative Research Design. An Interactive Approach*. 3<sup>rd</sup> ed. Washington: Sage.
- Mc Lachlan, J.C. & Whiten, S.C. 2000. Marks, Scores and grades: scaling and aggregating student assessment outcomes. *Medical education*, 34(10):788-797.
- McCoubrie, 1990. Improving the fairness of Multiple questions: a literature review. *Medical Teacher*, 26(8):709-712.
- McMillan, J. & Schumacher, S. 2001. *Research in Education: a conceptual introduction*. New York: Addison Wesley Longman.
- Merrel, J.D., Cirillo, P.F., Schwartz, P.M. & Webb, J. 2015. Multiple-Choice Testing Using Immediate Feedback-Assessment Technique (IFAT) Forms: Second-Chance Guessing vs Second-Chance Learning? *Higher Education Studies*, 5(5):50-55.
- Moncrief, D. & Foster, R. 2012. *Well That Was Easy: Misdirecting Respondus LockDown Browser for Fun and Profit*. [Sound Recording] (Central Washington University).
- Morton, S., Xu, Y. & Joplin, R. 2012. Evaluation of a new automated spotter style exam for assessment of anatomical knowledge. *Clinical Anatomy*, 25(6):806.
- Mory, E. 2003. Feedback research revisited. In: D.J. Ed (Ed.) *Handbook of research for educational communications and technology* (pp. 745-783). London and New York: Routledge.
- Mzee, A. 2016. Effectiveness of School Administration in Zanzibar: Student's Achievement. *Journal of Asian and African Social Science and Humanities*, 1(3): 1-19.
- Nazeem, I., Vinayak, S., Bhaysar, H. & Sukhwant, B. 2016. Contemplation on marking scheme for Type X MCQs, and an illustration of a practically applicable scheme. *Indian J Pharmacol*, 2:114-121.

- Nicol, D.J. & Dick, D.M. 2007. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2):199-218.
- Nicol, D.J. & Macfarlane-Dick, D. 2005. Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 2005: 1-19.
- Palmquist, M. 2016. *writing@csu writing guide*. [Online]. Available at: <http://writing.colostate.edu/guides/pdfs/guide66.pdf>. [Accessed 25 August 2016].
- Payne, W.F. & Brinkman, A. 2007. Towards effective feedback in e-learning packages: The design of a package to support literature searching, referencing and avoiding plagiarism.. *Proceedings of HCI2007 workshop: design and use and experience of e-learning systems*, pp. 71-75.
- Pittenger, A.L. & Lounsbery, J.L. 2011. Student-generated questions to assess learning in an online orientation to pharmacy course. *American Journal of Pharmaceutical Education*, 75(5).
- Poutre, B., Hedlund, D. & Nau, W. 2015. *Combining Testing Software, Online Proctoring and Lockdown Browsers to Assure a Secure Assessment Environment for Students in Hybrid or Online Programs*. California: Creighton University.
- Pressey, S. 1926. A Siple apparatus which gives tests and scores and teaches. *School and Society*, 23:373-376.
- Pressey, S. 1950. Development and appraisal of devices providing self-instruction. *The Journal of Psychology*, 29:417-447.
- Ramsden, P. 1992. *Learning to Teach in Higher Education*. London: Routledge.
- Roediger, H.L. & Butler, A. 2011. The critical role of retrieval practice in long-term retention. *Trends in cognitive sciences*, 15(1):20-27.
- Rollnick, M., Brenner, E. & Molatsane, G. 2010. Item analysis as a useful diagnostic tool in the design and analysis of multiple-choice questions. In: D.L. Lange (Ed.) *Teaching and Learning* (pp. 51-65). Pretoria: Council of Higher Education.

- Sambell, K., Sambell, A. & Sexton, G. 1999. Student perception of the learning benefits of computer-assisted assessment. *Computer-Assisted Assessment in Higher Education*:179-191.
- Schnell, J., Lukoff, B. & Mazur, E. 2013. Catalyzing learner engagement using cutting-edge classroom response systems in higher education. *Technologies in higher education* :233-261.
- Schuwirth, L.W.T. & Van der Vleuten, C.P.M. 2004. Different written assessment methods:what can be said about their strengths and weaknesses? *Medical Education*, 38(9):974-979.
- Scriven, M. 1967. The Methodology of evaluation. *Perspectives of curriculum evaluations*: 39-83.
- Shute, V. 2008. Focus on formative feedback. *Review of Educational Research*, 78(1):153-189.
- Siddiqui, N.I., Bhaysar, V.H., Bhaysar, A.V. & Bose, S. 2016. Contemplation on marking scheme for Type X multiple-choice questions and an illustration of a practically applicable scheme. *Indian Journal of Pharmacology*, 48(2):114-121.
- Slepkov, A.D. 2013. Integrated testlets and the immediate feedback assessment technique. *American Journal of Physics*, 81(10):782-791.
- Smith, J.A. 2013. Immediate Feedback and Assessment Technique (IF-AT) testing forms: An overview of the tool and uses. *Developments in Business and Experimental Learning*, 40:108.
- Smith, G., Wood, L., Coupland, M., Stephenson, B., Crawford, K. & Ball, G. 1996. Constructing mathematical examinations to assess a range of knowledge and skills. *International Journal of Mathematical Education in Science and Technology*, 27(1):65-77.
- Snowball, J.D. 2014. Using interactive content and online activities to accommodate diversity in a large first year class. *Higher Education*, 10:823-838.

- Stiggins, R. 2002. Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10):758-767.
- Tamir, P. 1990. Justifying the selection of answers in Multiple-Choice items. *International Journal for Science Education*, 12:563-573.
- Toplis, R. 2015. *Learning to teach Science in the Secondary school*. 4<sup>th</sup> ed. New York: Routledge.
- Tsze, S.L. 2013. Multiple-choice Question exam an examination process. A Descriptive analysis. *Malaysian Online Journal of Educational Management (Mojem)*, 1(2):19-32.
- University of South Wales, 2014. *Teaching@UNSW/Assessment toolkit*. [Online]. Available at: <https://teaching.unsw.edu.au/assessing-multiple-choice-questions> [Accessed 20 November 2016].
- Velan, G.M., Kumar, R., Dziegliwski, M. & Wakefield, D. 2002. Web-based self-assessments in pathology with questionmark perception. *Pathology*, 34:282-284.
- Ventouras, E.E., Triantis, P., Tsiakas, P. & Stergiopoulos, C. 2011. Comparison of oral examination and electronic examination using paired multiple-choice questions. *Computers and Education*, 56(3):616-624.
- Vygotsky, L.S. 1978. *Mind in Society: The development of higher psychological process*. Cambridge: Harvard University Press.
- Wainer, H. & Thissen, D. 1993. Combining multiple-choice and constructed-response test scores: toward a Marxist theory of test construction. *Applied Measurement in Education*, 6(2):103-118.
- Webb, J.H. 1989. Multiple-choice questions in Mathematics. *Suid-Afrikaanse tydskrif vir Opvoedkunde*, 9(1):216-218.
- Wenning, C.J. & Vieyra, R.E. 2015. *Teaching high school Physics*. Released on Kindle March 25, 2015.

- Wiliam, D. 2014. Formative assessment and contingency in the regulation of learning processes. *Paper presented in a Symposium entitled Toward a Theory of Classroom Assessment as the Regulation of Learning at the annual meeting of the American Educational Research Association, Philadelphia, PA, April 2014.*
- Wiliam, D. 2006. Formative assessment: Getting the focus right. *Educational Assessment*, 3&4(11):283-289.
- Wiliam, D. 2011. What is Assessment For Learning? *Studies in Educational Evaluation*, 37:3-14.
- Wiliam, D., Lee, C., Harrison, C. & Black, P. 2010. Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, 11(1):49-65.
- Wiliam, D. & Thompson, M. 2007. Integrating assessment with instruction: what will it take to make it work? *In: The Future of Assessment: Shaping Teaching and Learning* (pp. 53-82). Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Williams, J.B. 2006. Assertion reason multiple-choice testing as a tool for deep learning: A qualitative analysis. *Assessment and Evaluation in Higher Education*, 31(3):287-301.
- Zendejas, B., Cook, D. & Farley, D. 2010. Teaching First or Teaching last: Does the Timing Matter in Simulation-Based Surgical Scenarios? *Journal of Surgical Education*, 67(6):432-438.

## APPENDIX A



COLLEGE OF EDUCATION

---

DEPARTMENT OF SCIENCE AND  
TECHNOLOGY

INFORMATION LEAFLET AND INFORMED CONSENT

PROJECT TITLE:

**The Use of Multiple Choice Questions by Students in a first year University Science Education Course at a residential University.**

Primary investigator: Mrs HA Terblanche,

Bed Hons (Science and Technology Education)

Study leader: Prof J Dreyer, College of Education, UNISA

Dear Potential research participant,

You are invited to participate in a research study that forms part of my formal MEd-studies. This information leaflet will help you to decide if you would like to participate. Before you agree to take part, you should fully understand what is involved. You should not agree to take part unless you are completely satisfied with all aspects of the study.

### **WHAT IS THE STUDY ALL ABOUT?**

Students need to be assessed in several ways.

Questions need to be asked in class to confirm knowledge after facts are repeated often that will cause unstructured discussion.

Some knowledge can be assessed by recognition – MCQ (Multiple Choice Questions)

Sometimes lecturers make use of construct response to test calculations.

Although the students do not have a choice in the decision of assessment by the lecturer, they need to accept the choice of assessment made by the lecturer.

The research that I am doing is in the interest you as the student.

I really believe that you will benefit out of it.

It is to find a way that calculations can also be tested with MCQ if immediate feedback is given. You will then not get zero for a question if you make a simple mistake during calculation. Marks allocated for a written test will be the same as marks allocated by using MCQ. The data will be collected from 40 TUT students in the same class to assure that all conditions will be the same.

Feedback will be given to inform the students of the findings.

### **WHAT WILL YOU BE REQUIRED TO DO IN THE STUDY?**

If you decide to take part in the study, you will be required to do the following:

- 1) Sign this informed consent form
- 2) Answer the questions traditionally in writing, in duplicate. The one copy must be handed in and will be marked.  
Choose the correct one-out-of-four distractors online from the other copy.

Repeat the same questionnaire online from your other copy with exactly the same questions and distractors. This time there will be hints for you to help.

At this time you will not know your results from the previous test.

The fifth distractor will be. “Can I have a hint please?”

- If that option is chosen, a hint in general will be given and you can still get **three** out of the five marks if the second try is correct.
- For trying a third time, you can still get **one** out of five marks
- If the correct answer is chosen at first, you will get **five** out of five for the question
- If you make the wrong decision by choosing the wrong distractor, a specific hint will be given that will tell you what you did wrong. If the correct answer is chosen the second time, you will get **four** out of five for the question.
- By trying another time, you can get **two** out of five marks

You can redo this to learn from your own mistakes (with or without feedback) or this method can be used for testing.

- 3) All 3 procedures will be taken in 1 period – 90 minutes. It will be taken once-off.
- 4) The location will be in our normal class venue – building 2-364
- 5) No specific requirements are necessary. It will be like a normal class test.
- 6) There will be no cost involve to participate in this study.

### **ARE THERE ANY CONDITIONS THAT MAY EXCLUDE YOU FROM THE STUDY?**

There is no condition of exclusion that is applicable on the study.

### **CAN ANY OF THE STUDY PROCEDURES RESULT IN PERSONAL RISK, DISCOMFORT OR INCONVENIENCE?**

There is no risk involved by participating in this research because it is actually just another class test to be written

### **WHAT ARE THE POTENTIAL BENEFITS THAT MAY COME FROM THE STUDY?**

The benefit of participating in this study is that you will make a contribution towards assessment of Multiple Choice Questions in teaching and learning.

### **WILL YOU RECEIVE ANY FINANCIAL COMPENSATION OR INCENTIVE FOR PARTICIPATING IN THE STUDY?**

Please note that you **will unfortunately not** be paid to participate in the study.

### **WHAT ARE YOUR RIGHTS AS A PARTICIPANT IN THIS STUDY?**

Your participation in this study is entirely voluntary. You have the right to withdraw after writing the written test without any penalty or future disadvantage whatsoever. You do not even have to provide the reason/s for your decision. Your withdrawal will in no way influence your continued relationship with the lecturer or Engineering Science marks. Note that you are not waiving any legal claims, rights or remedies because of your participation in this research study.

## **HOW WILL CONFIDENTIALITY AND ANONYMITY BE ENSURED IN THE STUDY?**

All information obtained during the course of this study is strictly confidential. The study data will be coded so that it will not be linked to your name. Your identity will not be revealed while the study is being conducted or when the study is reported in scientific journals. All the data sheets that have been collected will be stored in a secure place. Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. The information received during the project will only be used for research purposes and not be released for any academic assessment, study progress, disciplinary purposes and/or study permit-related matters.

## **IS THE RESEARCHER QUALIFIED TO CARRY OUT THE STUDY?**

The researcher is an adequately trained and qualified researcher in the study fields covered by this research project, specifically in education.

## **HAS THE STUDY RECEIVED ETHICAL APPROVAL?**

Yes. The Faculty Committee for Postgraduate Studies and the Research Ethics Committee of the Tshwane University of Technology and UNISA have approved the formal study proposal. All parts of the study will be conducted according to internationally accepted ethical principles.

## **WHO CAN YOU CONTACT FOR ADDITIONAL INFORMATION REGARDING THE STUDY?**

The primary investigator, Mrs H Terblanche, can be contacted during office hours at Tel (012) 382-5531, or on her cellular phone at 082 888 5769. The study leader, Prof J Dreyer, can be contacted during office hours at Tel (012) 542 3562. Should you have any questions regarding the ethical aspects of the study, you can contact the chairperson of the TUT Research Ethics Committee, Dr WA Hoffmann, during office hours at Tel (012) 382-6265/46, E-mail [hoffmannwa@tut.ac.za](mailto:hoffmannwa@tut.ac.za). Alternatively, you can report any serious unethical behaviour at the University's Toll Free Hotline 0800212341.

## **DECLARATION: CONFLICT OF INTEREST**

This research study is funded by UNISA. No publication prohibitions, conditions or limitations are placed on the researcher.

## **A FINAL WORD**

Your co-operation and participation in the study will be greatly appreciated. Please sign the informed consent below if you agree to participate in the study. In such a case, you will receive a copy of the signed informed consent from the researcher.

## **CONSENT**

I hereby confirm that I have been adequately informed by the researcher about the nature, conduct, benefits and risks of the study. I have also received, read and understood the above written information. I am aware that the results of the study will be anonymously processed into a research report. I understand that my participation is voluntary and that I may, at any stage, without prejudice, withdraw my consent and participation in the study. I had sufficient opportunity to ask questions and of my own free will declare myself prepared to participate in the study.

Research participant's name: \_\_\_\_\_ (Please print)

Research participant's signature: \_\_\_\_\_

Date: \_\_\_\_\_

Researcher's name: \_\_\_\_\_ (Please print)

Researcher's signature: \_\_\_\_\_

Date: \_\_\_\_\_



Tshwane University  
of Technology

*We empower people*

FACULTY OF Q

Dept of X

**FIELDWORKER AND/OR RESEARCH ASSISTANT CONFIDENTIALITY  
AGREEMENT**

I, ITANI BRENDA NGOMA, hereby agree to:

1. Abide by the confidentiality requirements of this study, as approved by the Research Ethics Committee (REC) of Tshwane University of Technology, by ensuring that the identities and information of the participants are not revealed during and after the course of study;
2. Keep all the research information shared with me confidential by not discussing or sharing the research information in any form or format with anyone other than the Principal Investigator(s);
3. Keep all research information in any form or format securely stored while it is in my possession;
4. Return all research information in any form or format to the Principal Investigator(s) when I have completed the research tasks;
5. After consulting with the Principal Investigator(s), erase or destroy all research information in any form or format regarding this research project that is not returnable to the Principal Investigator(s) (e.g. information sorted on computer hard drive).

If you have any questions or concerns about this study, please contact:

Dr. Z

Phone: (012) 382 1111

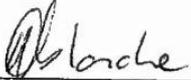
E-mail: [drz@tut.ac.za](mailto:drz@tut.ac.za)

This study has been reviewed and approved by the Research Ethics Committee, Tshwane University of Technology. For questions regarding your rights and/or the ethical conduct of research, contact the REC Chairperson, Dr. WA Hoffmann, at (012) 382 6246 ([hoffmannwa@tut.ac.za](mailto:hoffmannwa@tut.ac.za)).

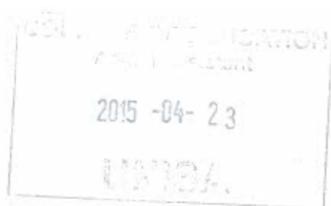
Research Assistant or Fieldworker:

<u>ITANI BRENDA</u>	<u></u>	<u>02-06-2015</u>
Print name	Signature	Date

Principal Investigator:

<u>H A Terblanche</u>	<u></u>	<u>02/06/2015</u>
Print name	Signature	Date

## APPENDIX B



### COLLEGE OF EDUCATION RESEARCH ETHICS REVIEW COMMITTEE

15 April 2015

Ref#: **2015/04/15/06697097/07/MC**  
Student #: Mrs HA Terblanche  
Student number#: 06697097

Dear Mrs Terblanche,

**Decision: Approved**

#### Researcher

Mrs HA Terblanche  
+2782 8885769  
[hettiet@gmail.com](mailto:hettiet@gmail.com)

#### Supervisor

Prof J Dreyer  
Department of Science and Technology Education  
College of Education  
+27828885769  
[dreyer@unisa.ac.za](mailto:dreyer@unisa.ac.za)

**Proposal:** The use of multiple choice questions (MCQ) by students in a first year University Science Education course at a Residential University

**Qualification:** M Ed in Natural Science Education

Thank you for the application for research ethics clearance by the College of Education Research Ethics Review Committee for the above mentioned research. Final approval is granted for 2 years.

**For full approval:** The application was reviewed in compliance with the Unisa Policy on Research Ethics by the CEDU ERC on 15 April 2015.

The proposed research may now commence with the proviso that:

- 1) The researcher/s will ensure that the research project adheres to the values and principles expressed in the UNISA Policy on Research Ethics.
- 2) Any adverse circumstance arising in the undertaking of the research project that is relevant to the ethicality of the study, as well as changes in the methodology, should



University of South Africa  
Preller Street, Muckleneuk Ridge, City of Tshwane  
PO Box 392 UNISA 0003 South Africa  
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150  
[www.unisa.ac.za](http://www.unisa.ac.za)

*be communicated in writing to the College of Education Ethics Review Committee.*

*An amended application could be requested if there are substantial changes from the existing proposal, especially if those changes affect any of the study-related risks for the research participants.*

- 3) *The researcher will ensure that the research project adheres to any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study.*

**Note:**

*The reference number [top right corner of this communiqué] should be clearly indicated on all forms of communication [e.g. Webmail, E-mail messages, letters] with the intended research participants, as well as with the College of Education RERC.*

Kind regards,



**Dr M Claassens**  
**CHAIRPERSON: CEDU ERC**

[mcdtc@netactive.co.za](mailto:mcdtc@netactive.co.za)



**Prof VI McKay**  
**ACTING EXECUTIVE DEAN**



University of South Africa  
Preller Street, Muckleneuk Ridge, City of Tshwane  
PO Box 392 UNISA 0003 South Africa  
Telephone: +27 12 429 3111 Facsimile: +27 12 429 4150  
[www.unisa.ac.za](http://www.unisa.ac.za)

16 February 2015

Dear Mrs Viljoen

Permission for data collection, during class time, for master study done by HA Terblanche

As you know, I am currently busy with my Master Study in Science Technology Education.

The need for my study comes from my teaching experience in Engineering Science. Therefore, it will be better for me to do the data collection for my study during scheduled class times because the info actually need to be taken in the most normal circumstances possible.

The test will be on work that the students are currently busy with. My students need to write a class test on the work tested anyway.

Some knowledge can be assessed by recognition – Multiple Choice Questions (MCQ). Sometimes lecturers make use of construct response to test calculations

Although the students don't have a choice in the decision of assessment by the lecturer, they need to accept the choice of assessment made by the lecturer.

Because the same test will be written 3 times for the research purposes, students who don't want to participate in the research, will be excused from class after the written test is written. The other students will then do the MCQ test and the test with the immediate feedback anonymously as well, for the data collection. No new work will be done the rest of the period. No student will be negatively influenced at all.

Students in class who wish not to participate in the study, will therefore be ensured and respected.

I will appreciate it if you can take up this matter seriously.

Regards

*HA Terblanche*

HA Terblanche

*W. Approved  
25/2/2015*

2015-02-25
TEL: (012) 382-5037 FAX: (012) 382-3577
ENTITY/ACCOUNT: T018/2425

## Research Ethics Committee

---

*The TUT Research Ethics Committee is a registered Institutional Review Board (IRB 00005968) with the US Office for Human Research Protections (IORG# 0004997) (Expires 9 Jan 2017). Also, it has Federal Wide Assurance for the Protection of Human Subjects for International Institutions (FWA 00011501) (Expires 22 Jan 2019). In South Africa it is registered with the National Health Research Ethics Council (REC-160509-21).*

---

May 26, 2015

Ref #: REC/2015/05/002 Name: Terblanche HA Student #: 6697097, UNISA
--

Terblanche HA  
C/o Prof J Dreyer  
Department of Science and Technology Education  
College of Education  
UNISA

Dear Ms Terblanche,

<b>Decision: Final Approval</b>
---------------------------------

---

**Name:** Terblanche HA

**Project title:** *The use of multiple choice questions by students in a first year university Science Education course at a residential university*

**Qualification:** M Ed Science and Technology Education, UNISA

**Supervisor:** Prof J Dreyer

---

Thank you for submitting the project documents for ethics clearance by the Research Ethics Committee (REC), Tshwane University of Technology (TUT). In reviewing the documents, the comments and notes below are tabled for your consideration, attention and notification:

- **University of South Africa (UNISA), Ethics Letter**
  - The REC took note of the ethical clearance granted by the UNISA College of Education Research Ethics Review Committee (Ethical Clearance Number: 2015/04/15/6697097/07/MC; dated 15 April 2015).
  
- **TUT, Head of Dept, Permission Letter**
  - The REC took note of the permission granted by the relevant Head of Dept on 25 Feb 2015.



---

*We empower people*

- **Proposal**

- **Sampling & Data Collection Strategies.** The REC strongly recommends that an independent data collector should be used to administer and collect the questionnaire data; this is to ensure voluntary participation in a context where the researcher is also the target participants' lecturer. As such, the REC requires that a **formal confidentiality agreement** be in place with the data collector (i.e. fieldworker) before the onset of data collection activities (see the attached TUT template *Fieldworker Confidentiality Agreement*). Also, note that it is the researcher's responsibility to ensure that the data collector fully adheres to and complies with the ethical principles applicable to the project.

- **Information Leaflet & Informed Consent**

- The Information Leaflet and Informed Consent document is in order and duly noted.

The Research Ethics Committee (REC), Tshwane University of Technology (TUT) reviewed the project documents at its meeting on May 25 2015. **Final approval** is granted to the project.

The proposed research project may now continue with the proviso that:

- 1) The researcher/s will conduct the study according to the procedures and methods indicated in the approved proposal, particularly in terms of any undertakings and/or assurances made regarding the confidentiality of the collected data.
- 2) The proposal will again be submitted to the Committee for prospective ethical clearance if there are any substantial changes from the approved proposal.
- 3) The researcher will act within the parameters of any applicable national legislation, professional codes of conduct, institutional guidelines and scientific standards relevant to the specific field of study.
- 4) The current ethics approval expiry date for this project is **December 31, 2017**. No research activities may continue after the ethics approval expiry date. Submission of a duly completed Research Ethics Progress Report (available at: <http://www.tut.ac.za/Other/rminew/ResearchEthicsCommittees/Pages/default.aspx>) will constitute an application for renewal of SCRE ethics approval.

Note:

*The reference number [top right corner of this communiqué] should be clearly indicated on all forms of communication [e.g. Webmail, E-mail messages, letters] with the intended research participants.*

Yours sincerely,



WA HOFFMANN (Dr)  
Chairperson: Research Ethics Committee  
[Ref#2015=05=002=TerblancheHA]



---

*We empower people*

## APPENDIX C

Engineering Science Class test 7 Work and Energy - 2015

Initials and Surname: \_\_\_\_\_

Group number \_\_\_\_\_

Class list number \_\_\_\_\_

Use  $g = 9.81 \text{ ms}^{-2}$

Round of to 2 decimal places

### Question 1

A skip of mass 500kg is raised from a depth of 400m to the surface by a rope of mass 5kg/m. Calculate the work done to raise the skip for 100m above the surface.

### Question 2

A horse pulls a car 100kg along a railway track by means of a rope inclined at  $15^\circ$  to the track. The pull in the rope is 500N and the coefficient of friction is 0.2. Determine the amount of work done to move the car 15m.

### Question 3

A truck is travelling at 15m/s down a hill when the brakes on all 4 wheels lock. The hill makes an angle of  $10^\circ$  with respect with the horizontal. The coefficient of kinetic friction between the tires and the road is 0.5. How far does the truck skid before coming to a stop? Use conservation of energy.

### Question 4

The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the useful energy supplied by the engine.

Engineering Science Class test 7M Work and Energy - 2015

Initials and Surname: \_\_\_\_\_

Group number \_\_\_\_\_

Class list number \_\_\_\_\_

Use  $g = 9.81 \text{ ms}^{-2}$

Round of to 2 decimal places

**Question 1**

A skip of mass 500kg is raised from a depth of 400m to the surface by a rope of mass 5kg/m. Calculate the work done to raise the skip for 100m above the surface.

- A) 735.75kJ
- B) 833.85 kJ
- C) 5150.25 kJ
- D) 1716.75 kJ
- E) 2207.25 kJ

**Question 2**

A horse pulls a car 100kg along a railway track by means of a rope inclined at  $15^\circ$  to the track. The pull in the rope is 500N and the coefficient of friction is 0.2. Determine the amount of work done to move the car 15m.

- A) 4.66 kJ
- B) 4.69 kJ
- C) 4.3 kJ
- D) 4.56 kJ
- E) 7.24 kJ

**Question 3**

A truck is travelling at 15m/s down a hill when the brakes on all 4 wheels lock. The hill makes an angle of  $10^\circ$  with respect with the horizontal. The coefficient of kinetic friction between the tires and the road is 0.5. How far does the truck skid before coming to a stop? Use conservation of energy.

- A) 2.40 m
- B) 35.94 m
- C) 23.64 m
- D) 25.32 m
- E) 35.05 m

**Question 4**

The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the useful energy supplied by the engine.

- A) 1261.29 kJ
- B) 290.1 kJ
- C) 4704.6 kJ
- D) 1243.35 kJ
- E) 1315.29 kJ

## Engineering Science Class test 7N Work and Energy - 2015

Initials and Surname: \_\_\_\_\_

Group number \_\_\_\_\_

Class list number \_\_\_\_\_

Use  $g = 9.81 \text{ ms}^{-2}$

Round of to 2 decimal places

### Question 1

A skip of mass 500kg is raised from a depth of 400m to the surface by a rope of mass 5kg/m. Calculate the work done to raise the skip for 100m above the surface.

F) 735.75kJ - Always from the top and use the correct part of the graph.

G) 833.85 kJ - Use the 5kg/m

H) 5150.25 kJ - Always from the top!

I) 1716.75 kJ - Add the 500 kg

J) 2207.25 kJ - Well done!

K) Give me a hint please - Draw the sketch and the graph. Always from the top!

### Question 2

A horse pulls a car 100kg along a railway track by means of a rope inclined at  $15^\circ$  to the track. The pull in the rope is 500N and the coefficient of friction is 0.2. Determine the amount of work done to move the car 15m.

F) 4.66 kJ - Not on the incline and use the horizontal component of the force.

G) 4.69 kJ - Well done!

H) 4.3 kJ - Recalculate  $F_N$

I) 4.56 kJ - Use the horizontal component of the force

J) 7.24 kJ - Use friction

K) Give me a hint please - The car is not on an incline

### Question 3

A truck is travelling at 15m/s down a hill when the brakes on all 4 wheels lock. The hill makes an angle of  $10^\circ$  with respect with the horizontal. The coefficient of kinetic friction between the tires and the road is 0.5. How far does the truck skid before coming to a stop? Use conservation of energy.

F) 2.40 m -  $E_k = \frac{1}{2} mv^2$

G) 35.94 m - Well done!

H) 23.64 m - Recalculate h

I) 25.32 m - Recalculate h

J) 35.05 m -  $F_N = mg\cos\theta$

K) Give me a hint please - Use the formula for energy conservation

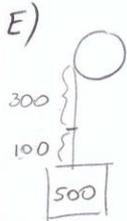
**Question 4**

The speed of a car of mass 1.2 ton increases from 10 m/s to 45 m/s while travelling 300 m up an incline of 1 in 40. The tractive resistance 50 N per ton. Determine the useful energy supplied by the engine.

- F) 1261.29 kJ - Well done!!
- G) 290.1 kJ -  $E_{pi} = 0$
- H) 4704.6 kJ - Recalculate  $h$
- I) 1243.35 kJ - Recalculate  $E_{loss}$
- J) 1315.29 kJ - Recalculate  $E_k$
- K) Give me a hint please - Use conservation of energy

Memo of assessment M & N

Question 1



$$m_1 = 500 \text{ kg} \checkmark$$

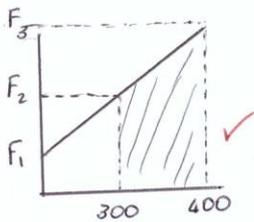
$$F_1 = 4905 \text{ N} \checkmark$$

$$m_2 = (300 \times 5) + 500 \checkmark$$

$$F_2 = 19620 \text{ N} \checkmark$$

$$m_3 = (400 \times 5) + 500 \checkmark$$

$$F_3 = 24525 \text{ N} \checkmark$$

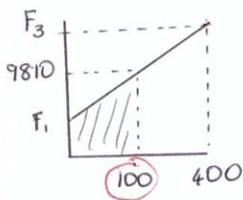


$$W = \frac{(24525 + 19620) \times 100}{2}$$

$$= \underline{2207,25 \text{ kJ}} \checkmark$$

(S) (E)

A)  $F_2 = [(100 \times 5) + 500] \times 9,81$   
 $= 9810 \text{ N}$



$$W = \frac{(4905 + 9810) \times 100}{2}$$

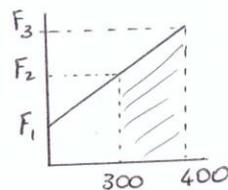
$$= \underline{735,75 \text{ kJ}}$$

Always from the top +  
 choose correct part of graph!

B)  $F_2 = (300 + 500) \times 9,81$   
 $= 7848 \text{ N}$

$$F_3 = (400 + 500) \times 9,81$$

$$= 8829 \text{ N}$$



$$W = \frac{(8829 + 7848) \times 100}{2}$$

$$= \underline{833,85 \text{ kJ}}$$

5 kg/m ?

C)  $F_2 = [(100 \times 5) + 500] \times 9,81$   
 $= 9810 \text{ N}$

$$W = \frac{(24525 + 9810) \times 300}{2}$$

$$= \underline{5150,25 \text{ kJ}}$$

Always from the top!

D)  $F_1 = 4905 \text{ N}$

$$F_2 = 1500 \times 9,81 = 14715 \text{ N}$$

$$F_3 = 2000 \times 9,81 = 19620 \text{ N}$$

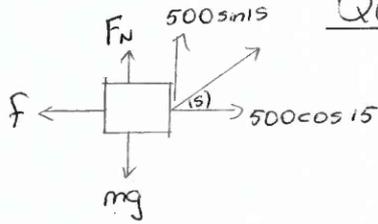
$$W = \frac{(14715 + 19620) \times 100}{2}$$

$$= \underline{1716,75 \text{ kJ}}$$

Add 500 kg!

Question 2

B)



$$F_N + 129,41 = mg$$

$$F_N = 851,59 \text{ N} \checkmark$$

$$F = 170,32 \text{ N} \checkmark$$

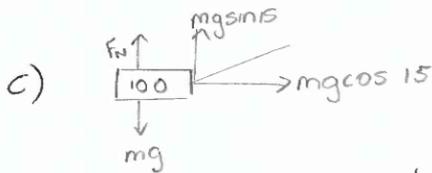
$$F_{\text{appl}} = 500 \cos 15 - 170,32 \checkmark$$

$$= 312,64 \text{ N} \checkmark$$

$$W = F \times s$$

$$= 4689,64 \text{ J}$$

$$= \underline{4,69 \text{ kJ}} \checkmark$$



$$F_N = 100 \times 9,81 = 981 \text{ N}$$

$$F = 981 \times 0,2 = 196,2 \text{ N}$$

$$F_{\text{appl}} = 500 \cos 15 - 196,2 \text{ N}$$

$$= 482,96 - 196,2$$

$$= 286,76 \text{ N}$$

$$W = F \times s$$

$$= 286,76 \times 15$$

$$= \underline{4,3 \text{ kJ}}$$

Recalculate  $F_N$

D)  $F = 196,2 \text{ N}$

$$F_{\text{appl}} = 500 - 196,2$$

$$= 303,8 \text{ N}$$

$$W = F \times s$$

$$= \underline{4,56 \text{ kJ}}$$

Take horizontal comp of 500 N

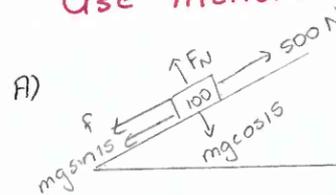
E)  $F_{\text{appl}} = 500 \cos 15$

$$= 482,96 \text{ N}$$

$$W = F \times s$$

$$= \underline{7,24 \text{ kJ}}$$

Use friction



$$F_N = mg \cos 15$$

$$F = 189,51 \text{ N}$$

$$F_{\text{appl}} = 500 - 189,51$$

$$= 310,49 \text{ N}$$

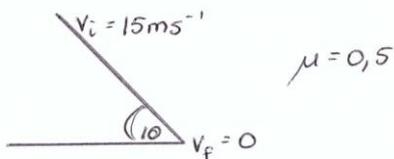
$$W = F \times s$$

$$= \underline{4,66 \text{ kJ}}$$

Not on incline +  
Use horizontal comp of  
force, too!

### Question 3

B)



$$E_{ki} + E_{pi} + E_{add} = E_{kf} + E_{pf} + E_{loss}$$

$$\frac{1}{2}m(15)^2 + mg\Delta x \sin 10 + 0 = 0 + 0 + \Delta x(0,5mg \cos 10)$$

$$\therefore 112,5m + 1,7m\Delta x + 0 = 4,83m\Delta x$$

$$112,5m = 3,13m\Delta x$$

$$\Delta x = 35,9m \quad \checkmark \quad (5)$$

C)  $112,5m + 1,7m = 4,83m\Delta x$

$$114,2m = 4,83m\Delta x$$

$$\Delta x = 23,64m$$

$$h \neq \sin 10$$

D)  $112,5m + 9,81m = 4,83m\Delta x$

$$122,31m = 4,83m\Delta x$$

$$\Delta x = 25,32m$$

Recalculate h

E)  $F_N = 9,81m \quad F = 4,91m$

$$112,5m + 1,7m\Delta x = 4,91m\Delta x$$

$$\Delta x = 35,05m$$

$$F_N = mg \cos 10$$

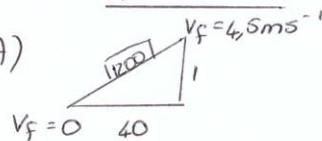
A)  $\frac{1}{2}m(15) + 1,7m\Delta x = 4,83m\Delta x$

$$7,5m = 3,13m\Delta x$$

$$\Delta x = 2,40m \quad E_k = \frac{1}{2}mv^2$$

### Question 4

A)



$$E_{ki} + E_{pi} + E_{add} = E_{kf} + E_{pf} + E_{loss}$$

$$\frac{1}{2}(1200)(10)^2 + 0 + E_{add} =$$

$$\frac{1}{2}(1200)(4,5)^2 + (1200 \times 9,81 \times \frac{300}{40})$$

$$+ 300 \times (1,2 \text{ ton} \times \frac{50N}{\text{ton}})$$

$$\therefore 60000 + E_{add} = 1215000 + 88290 + 18000$$

$$E_{add} = 1261,29 \text{ kJ} \quad \checkmark \quad (5)$$

B)  $\frac{1}{2}(1200)(10)^2 + 882900 + E_{add} =$

$$1215000 + 18000 + 0$$

$$60000 + 882900 + E_{add} = 1233000$$

$$E_{add} = 290,1 \text{ kJ}$$

$$E_{pi} = 0$$

C)  $60000 + E_{add} = 121500 +$

$$(mg \times 300) + 18000$$

$$E_{add} = 4704,6 \text{ kJ}$$

$$h = \frac{300}{40} \neq 300$$

D)  $60000 + E_{add} = 121500 + 88290 +$

$$1,2 \text{ ton} \times \frac{50N}{\text{ton}}$$

$$E_{add} = 1243,35 \text{ kJ}$$

$$E_{loss} \neq F$$

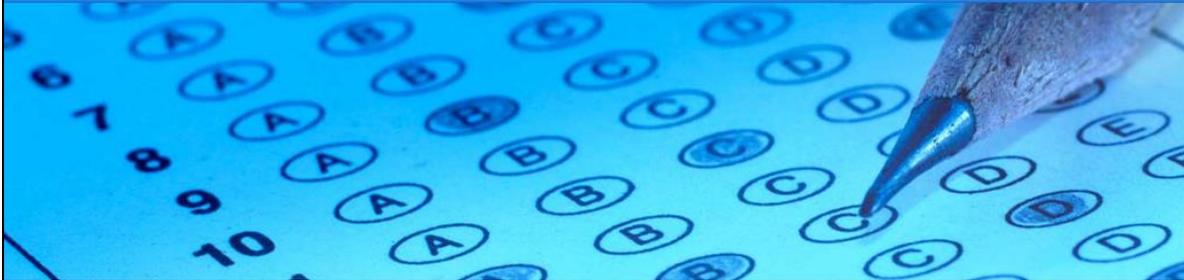
E)  $\frac{1}{2}(1200)(10) + E_{add} = 1215000 +$

$$88290 + 18000$$

$$E_{add} = 1315,29 \text{ kJ} \quad E_k = \frac{1}{2}mv^2$$

## APPENDIX D

Adaptive Testing



### Test Login

Login with your student code

Login

---

© 2017 - Hettie Terblanche

Adaptive Testing

### Normal Test

Your test is ready. You will have **10** minutes to answer **4** questions for **20** marks.

Do not use the browser's back or refresh buttons, because it will cause you to fail the test

Start Test

---

© 2017 - Hettie Terblanche

## Question 2

A horse pulls a car 100kg along a railway track by means of a rope inclined at  $15^\circ$  to the track. The pull in the rope is 500N and the coefficient of friction is 0.2. Determine the amount of work done to move the car 15m.

4.66 kJ
4.69 kJ
4.3 kJ
4.56 kJ
7.24 kJ

[Next Question »](#)

## Adaptive Test

Your test is ready. You will have **25** minutes to answer **4** questions for **20** marks.

Do not use the browser's back or refresh buttons, because it will cause you to fail the test

[Start Test](#)



## Test Login

Login with your student code

You need to do test M first

Login

© 2017 - Hettie Terblanche

## Question 2

Time: 00:24:09  
Chances: 3/3

A horse pulls a car 100 kg along a railway track by means of a rope inclined at  $15^\circ$  to the track. The pull in the rope is 500 N and the coefficient of friction is 0.2. Determine the amount of work done to move the car 15 m.

 4.66 kJ 4.69 kJ 4.3 kJ 4.56 kJ 7.24 kJ Give me a hint please

Well done!!!!

Next Question »

© 2017 - Hettie Terblanche

## Question 3

Time: 00:23:46  
Chances: 0/3

A truck is travelling at 15 m/s down a hill when the brakes on all 4 wheels lock. The hill makes an angle of  $10^\circ$  with respect with the horizontal. The coefficient of kinetic friction between the tires and the road is 0.5. How far does the truck skid before coming to a stop? Use conservation of energy.

2.40 m

35.94 m

23.64 m

25.32 m

35.05 m

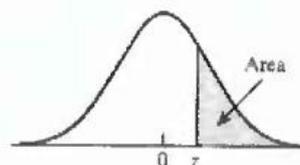
Give me a hint please

$$E_k = \frac{1}{2}mv^2.$$

Next Question »

## APPENDIX E

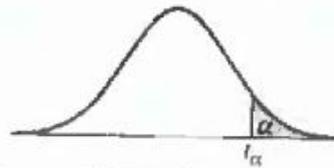
**Table 4 Normal Curve Areas**  
**Standard normal probability in right-hand tail**  
 (for negative values of  $z$ , areas are found by symmetry)



$z$	Second decimal place of $z$									
	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0722	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0352	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.00135									
3.5	.000233									
4.0	.0000317									
4.5	.00000340									
5.0	.000000287									

From R. E. Walpole, *Introduction to Statistics* (New York: Macmillan, 1968).

Table 5 Percentage Points of the  $t$  Distributions



$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$	df
3.078	6.314	12.706	31.821	63.657	1
1.886	2.920	4.303	6.965	9.925	2
1.638	2.353	3.182	4.541	5.841	3
1.533	2.132	2.776	3.747	4.604	4
1.476	2.015	2.571	3.365	4.032	5
1.440	1.943	2.447	3.143	3.707	6
1.415	1.895	2.365	2.998	3.499	7
1.397	1.860	2.306	2.896	3.355	8
1.383	1.833	2.262	2.821	3.250	9
1.372	1.812	2.228	2.764	3.169	10
1.363	1.796	2.201	2.718	3.106	11
1.356	1.782	2.179	2.681	3.055	12
1.350	1.771	2.160	2.650	3.012	13
1.345	1.761	2.145	2.624	2.977	14
1.341	1.753	2.131	2.602	2.947	15
1.337	1.746	2.120	2.583	2.921	16
1.333	1.740	2.110	2.567	2.898	17
1.330	1.734	2.101	2.552	2.878	18
1.328	1.729	2.093	2.539	2.861	19
1.325	1.725	2.086	2.528	2.845	20
1.323	1.721	2.080	2.518	2.831	21
1.321	1.717	2.074	2.508	2.819	22
1.319	1.714	2.069	2.500	2.807	23
1.318	1.711	2.064	2.492	2.797	24
1.316	1.708	2.060	2.485	2.787	25
1.315	1.706	2.056	2.479	2.779	26
1.314	1.703	2.052	2.473	2.771	27
1.313	1.701	2.048	2.467	2.763	28
1.311	1.699	2.045	2.462	2.756	29
1.282	1.645	1.960	2.326	2.576	inf.

From "Table of Percentage Points of the  $t$ -Distribution." Computed by Maxine Merrington, *Biometrika*, Vol. 32 (1941), p. 300.