CHAPTER 1 INTRODUCTION

1.1 Background

The landscapes in the South African education system are changing. This is due to the introduction and implementation of Outcome-Based Education (OBE) in the South African education system. Universities used the previous education systems as basis for structuring their teaching practices and syllabi. The universities were well informed about the educational background, demographic background, and skills, which their prospective students possessed. Students taught according to OBE have different attitudes and skills (amongst others) as compared to students taught according to the previous provincial education systems (Potgieter *et al.*, 2005a). Since the teaching practices and syllabi at universities are informed by the proficiencies of students coming from secondary schools, and with the changing landscapes, more students taught according to the OBE format will register at universities, therefore the universities need to be prepared to accommodate these changes. The universities need to be well informed about conceptual understanding, skills development, and educational background of their prospective students (Potgieter *et al.*, 2005b).

1.2. The South African Context

Zaaiman *et al.*, (2000) documented in their study that there are imbalances in terms of the educational background of students at the secondary school level in South Africa. One of the imbalances is the quality of secondary school the students are coming from. The quality of schools that students come from differs in terms of infrastructure

and the competency of teachers offering physical science at secondary schools. Most rural schools in South Africa are faced with overcrowded classrooms, making it harder for teachers to give individual attention to students, than in the schools in more privileged environments. These types of schools have as such limited resources for experimental activities. According to social constructivists, the student has to "see', "read" and "hear" in order for him/her to construct his/her own understanding of concepts (Thanasoulas, http://www3.telus.net/linguisticsissues/constructivist.html)). Seeing, hearing and reading are part of the interactions that a student has with the environment. Lev Vygotsky (1962), a proponent of the constructivist theory of learning, emphasized the importance of social interaction on learning and understanding of concepts (Scanlon *et al.*, 2002).

Shortage of qualified teachers of physical science is another contributing factor to the poor conceptual understanding of students in mechanics. Most of the South African physical science teachers, at secondary schools, went to colleges of education for their professional diplomas. At colleges of education they studied the textbooks they will be using when teaching and are normally not trained to be subject specialists. The teachers are therefore not fully equipped with much content knowledge of the subject. Such teachers end up teaching only the sections they understand, or rushing through the syllabus not paying attention to the understanding of students. On the other hand, learning for understanding, is regarded as a social activity. Interactions of the student with his/her environment and social agents, such as teachers, contribute to a greater extent to the learning and understanding of the student. If teachers are poorly equipped with the content knowledge of the subject, "scaffolding" (McKenzie, 1999) as an effective form of teaching will not be possible. According to McKenzie (1999)

scaffolding means that with guided and appropriate help from a teacher, a student can be able to perform difficult tasks he or she is not able to perform while working alone. The teacher at the same time regulates his/her degree of help, so that the student can develop understanding and skills necessary for applications of concepts in future.

Lack of resources mostly in rural schools also contributes to the limited knowledge and understanding of concepts. The students from these schools only get textbooks, but no additional learning materials to study. Experimental activities are hardly done because of lack of experimental facilities. The students rely on what the teacher tells them. The teacher, on the other hand, ends up "performing" the experiment theoretically, and thus telling the students what they are supposed to observe while performing the experiment. The students are thus not exposed to experimental learning or observation, and hence cannot infer on the observations made. Understanding of the learning material is thus not achieved, because one understands better what one observes and gets involved with, rather than what one is told will happen, or has happened. Some of these students end up registered at universities. The universities thus have a huge task of accommodating the imbalances in the educational background of the prospective students. In addition, the universities have to be well informed about the content knowledge and conceptual understanding the students have, upon entry and whether common sense ideas about physical processes are accepted scientific ideas or not.

1.3 Alternative Conceptions

Physics education research explores common sense ideas, alternative conceptions and preconceptions which refer to the students' intuitive ideas that are acquired before formal instruction has taken place. These are the ideas that students have as a result of their observations and interactions with the environment they live in. These ideas that students have, about physical processes, are sometimes different from acceptable scientific ideas. Hasan *et al.* (1999:294) defines misconceptions as "strongly held cognitive structures that are different from the accepted understanding in the field that are presumed to interfere with the acquisition of new knowledge", and the Longman Dictionary of Contemporary English (1995: 906) defines misconceptions as "An idea which is wrong or untrue, but which people believe because they do not understand it properly". However, misconception is a rather strong word, and Ramaila (2000) indicated that the use of the word misconceptions is inappropriate as it ignores the rational basis of the students' conceptions about physical processes prior to formal instructions. For the purpose of this study alternative conceptions, common sense ideas and preconceptions will be regarded as synonymous to misconceptions, and common sense ideas and alternative conceptions will be used in this study.

Physics education research documented that students have alternative conceptions about physical processes, which are developed from very young ages. These alternative conceptions developed through their observation and interactions with the environment, and from past experiences. These ideas form a foundation for the learning and understanding of basic mechanics, since mechanics is a part of physics that is closer to

the students' daily life experiences about motion and forces (Planinic *et al.*, 2006; Ramaila, 2000).

A minimum level of conceptual understanding in mechanics is taken for granted by lecturers at universities, especially at first year level. Over the past years, there have been a number of studies focusing on the students' understanding of physics concepts. Jimoyiannis & Komis (2003) and Knight (1995) documented that secondary and university physics students have limited basic knowledge and thus have difficulties in the understanding of mechanics. This lack of basic knowledge and understanding has an impact on the understanding of other more complex topics at higher levels of physics.

Hasan *et al.* (1999) and Planinic *et al.* (2006) developed techniques to reveal the presence of alternative conceptions by studying whether students are making correct judgment about their understanding of physics concepts. Hasan *et al.* (1999) used the certainty of response index (CRI) to classify students into those having lack of knowledge and those having strongly held alternative conceptions. The study by Planinic *et al.* (2006) indicated that students show higher confidence levels on Newtonian mechanics than on electromagnetism, and that the higher confidence levels associated with incorrect answers indicated strongly held alternative conceptions. The students were confident about their responses, even though they gave incorrect answers. This thus indicated that the students made incorrect judgments about their understanding of basic concepts.

This study investigated the students' knowledge and understanding of basic concepts of mechanics upon entry into some of the South African universities. The study was conducted at the University of Limpopo (UL) Turfloop Campus, the University of

Pretoria (UP) and the University of Cape Town (UCT). Existing standardized tests were used in this study to probe the students' knowledge and understanding of basic mechanics concepts. Items from the Force Concepts Inventory (FCI) (Hestenes *et al.*, 1992) and Mechanics Baseline Test (MBT) (Hestenes & Wells, 1992) were selected for the study. Items in these test instruments probed for conceptual understanding of basic mechanics. The mechanics part was chosen because it is the part of physics that is closer to the daily experiences and activities that the students encounter (Planinic, *et al.*, 2006; Ramaila, 2000).

This study follows a correlational research design, which explores relationships between variables. The variables investigated in the study are conceptual understanding, test performance and confidence level. Descriptive statistics were used to analyse the quantitative data collected for the study. The analysis of conceptual understanding was done by coding of the students' written explanations of their chosen options.

1.4 Rationale

The South African education system is in the process of phasing out the previous systems of teaching and learning employed by various provincial education departments, and replacing it with the OBE system of teaching and learning. This will have some bearing on the teaching activities and syllabi at secondary schools and universities, and these changes will affect the profiles of students entering the universities. This study is thus aimed at investigating the students' profiles upon entry to tertiary education. Do students know and understand basic mechanics concepts? Are students able to make accurate judgment about their knowledge and understanding of these concepts?

Galili (1995) conducted an investigation about alternative conceptions in both electromagnetism and mechanics. The author documented that alternative conceptions in mechanics influenced alternative conceptions in electromagnetism. Mechanics forms the basis for the learning and understanding of the other topics in physics. Learning and understanding in higher levels of physics will thus become a struggle for a student having difficulties in understanding mechanics.

Students having poor knowledge and understanding of basic concepts will experience failure somewhere in their study at universities, and will ultimately take a longer time to complete their degrees programmes (Santiago & Einarson, 1998), a fact that has some bearing on the finances of a number of stakeholders, i.e. the parents and the government. Over and above, this has also some bearing on the students' academic progress; the students may not complete their programmes on time or could possibly drop out without completing their programmes. This, together with other factors, affects the retention and throughput rates of universities, which has become a serious concern due to changes in government funding formulae.

The results of study will inform lecturers, particularly at first year level about the conceptual understanding and misconceptions present among the students. Knowing what the students do not understand and how they think, will require commitment and active intervention from the lecturers. Such intervention and commitment by the lecturers could possibly lead to improvement in the students' understanding of the concepts. This could have positive consequences in that the students' problem solving abilities improve and students' performance, which is linked to understanding, could also improve, thus an improvement in the throughput rate at tertiary institutions.

The science education literature has a range of studies that investigated aspects of students' confidence and how such confidence impacts on the students' test performance (Metcalfe, 1998; Pallier *et al.*, 2002). Hasan *et al.* (1999) used a certainty of response index to differentiate between students' strongly held alternative conceptions and their lack of knowledge. However, they did not have the means of checking their hypothesis. The significance of this study is that students' written explanations of answers given for multiple-choice questions were analysed to check this hypothesis. This study provides results that could be of use to the science education research community as well as to tertiary educators in physics.

CHAPTER 2 LITERATURE REVIEW

2.1. Background Literature

Research work on self-efficacy and confidence levels has been conducted mainly in the social sciences (Bandura, 1982). In the sciences a number of studies have been conducted on some aspects of student confidence and how such confidence impact on the students' test performance (Metcalfe, 1998; Pallier *et al.*, 2002).

2.1.1. Self-efficacy

Bandura (1986: 391) defines self-efficacy as "The judgment of one's capabilities to organize and execute courses of action required to obtain designated types of performance." Self-efficacy refers to students' confidence about their ability to accomplish a task. For example, if students have high, positive self-efficacy about mastering chemistry or physics, then they believe that they have the power and abilities to pass chemistry or physics and ultimately succeed in their studies. On the other hand, students with low self-efficacy feel that they do not have the power and abilities to master components of a discipline, thus admitting failure from the start. Students with high self-efficacy are more likely to succeed at learning and also to be more motivated to seriously study. Highly motivated students work hard, persevere in the face of difficulties, and find satisfaction in the successful accomplishment of a learning task (Barnhardt, 1997)

A number of studies have found that self-efficacy and academic performance are related (Andrew, 1998; Chacko & Huba, 1991). These studies have been conducted at different academic settings. Sottile *et al.* (2002) conducted a study on the school culture,

science achievement and mathematics achievement; and found achievement and self-efficacy are related among in-service teachers. Several studies have been undertaken about relationships between confidence and academic achievement (Cavallo & Rozman, 2004; Gillibrand *et al.*, 1999) and between self-efficacy and performance (Kranzler & Pajares, 1997). In these studies it is shown that a student whose performance is high has high self-efficacy, and the one with low performance has low self-efficacy.

2.1.2. Confidence

Longman Dictionary of Contemporary English (1995: 284) defines confidence as "The belief that you have the ability to do things well or deal with situations successfully." For the purpose of the study self-efficacy, confidence levels and confidence would be used synonymously.

Potgieter *et al.* (2005b) conducted a study on the initial understanding and knowledge of basic chemical concepts of students upon entry into some of the South African universities. The authors documented that students do not make correct judgments about their understanding of basic chemical concepts. Students tend to be overconfident about their competencies. The students' poor judgment about their competencies could indicate lack of knowledge or strongly held alternative conceptions. They attributed the fact that students in the University of Pretoria Foundation Year Programme (UPFY) have very high levels of confidence to be partly due to their experiences leading to selection in the programme. As a result, experiences of failure do not impact significantly on their confidence. Students with high levels of confidence may dismiss messages of failure because it does not fit into their self-perception. Another

factor could be that the student has never failed or repeated a grade before, hence their overconfidence. Therefore when such a student enters the university and meets those whose performance is above his/hers, the environment thus proves that there is a lot that he/she does not know, in other words his/her knowledge is limited. However, if this message is not accepted and acted upon in time, failure may be inevitable.

2.1.3 Alternative Conceptions

Spencer Kagan (1992) indicated in his book "Cooperative Learning" that students do not come to institutions as "empty" vessels. They have knowledge based on their observations and interactions with the environment they live in and from past experiences. These experiences with mechanics developed from a very young age. Even before formal instruction takes place, children have observed donkeys pulling a cart in their villages. They have observed their mothers balancing buckets of water on their heads. They have observed busses speeding up the roads past their villages. This implies that children have observed and experienced motion and forces from very young ages. Therefore even students at secondary schools have developed their own understanding of mechanical problems before formal instructions in mechanics have taken place. Upon entry to university, students bring with them alternative conceptions about physical processes, and have meanings attached to these physical processes. Physics education research concur with Spencer Kagan (1992), and has, over a number of years, documented that students have some ideas about physical processes in physics, and already have common sense explanations for these processes. However, these ideas that students have, are in many instances, incompatible with accepted scientific ideas

(Halloun & Hestenes, 1985a; Maloney *et al.*, 2001). The alternative conceptions that students bring along to class are the major influences of what they (students) will learn in the physics course. Mundalamo & Grayson (2006) conducted a study on the performance of foundation and non-foundation physics I students, at some South African universities. They reported that, even though the students came from different educational backgrounds, there were some indications that the students had little knowledge of certain concepts in mechanics. Their study indicated that the students performed badly in items that related to gravitational acceleration. The concepts of gravitational forces and acceleration are regarded as basic in elementary mechanics. The success of students in their physics first year of study is determined, among others, by their conceptual understanding and knowledge of basics mechanics principles. Galili (1995) conducted a study to investigate the presence of alternative conceptions in mechanics and in electromagnetism. The author documented that the presence of alternative conceptions is stronger in mechanics than in electromagnetism.

2.1.4 South African School Situation

In South Africa there are unique problems which negatively affect the acquisition of knowledge and understanding in physics and the development of justified confidence in physics, e.g. lack of discipline, insufficient resources, poor morale, inadequate parental involvement, inadequate transfer of relevant skills, classroom environment and student-teacher ratio, amongst others. Zaaiman *et al.* (2000) indicated that the majority of the secondary schools in South Africa, have inadequate or non-existent physical infrastructures. The majority of the teachers are not qualified to teach physical science.

These teachers are not enthusiastic, confident and competent in their teaching, and are unable to fulfill one of their responsibilities, helping students to learn in meaningful ways. Barlia & Beeth (1999) seem to concur with Zaaiman *et al.* (2000), and reported in their research that these are some of the reasons why students tend to have negative attitudes towards sciences, and low motivation in learning sciences for understanding. The students' awareness of affection and individual needs are not taken into consideration, since the teacher and the teaching environment is not supportive and conducive to learn science for understanding.

2.2. Research Questions

The study is aimed at answering the following questions:

1. What are the performance and associated confidence levels of first entering physics students registered at selected South African universities?
2. Is there a correlation between the confidence and performance of students in mechanics?
3. Can the relationship between confidence and performance be used to reliably identify the presence of misconceptions in mechanics?

CHAPTER 3 THEORETICAL FRAMEWORK

3.1 Introduction

The theoretical framework within which this study is situated is social constructivism. According to Kim (2001), social constructivism emphasizes the importance of culture and the society in which knowledge is constructed. This theory is closely associated with Vygotsky's constructivist theory of learning.

Vygotsky, (1962) viewed learning as a social process. Through observations and interactions with the environment, a student is able to acquire knowledge. Vygotsky emphasizes that learning cannot take place in isolation, but takes place within individuals who interact with each other and as well as the world around them. Students must be more actively involved in activities with each other and their teacher to generate new understanding (Epstein & Ryan, 2002).

Proponents of social constructivism believe that knowledge is constructed through human activities, and can be constructed socially and culturally (Kim, 2001). Learning is an activity that takes place within students who are engaged cooperatively with each other, and thus learning can only take place among a number of individuals interacting with each other.

3.2 Alternative Conceptions

This study explores students' prior knowledge and their ideas about the physical processes, upon entry into the first year physics programmes at universities. The students' prior knowledge in mechanics includes a number of alternative conceptions. Many of

these alternative conceptions were formed outside physics classrooms through their everyday experiences of moving objects, donkeys pulling carts, bulls pushing each other during their fights for dominance, etc.

The study will utilize the Hasan et al. (1999) model in order to interpret and analyze the students' responses. This model employs the Certainty of Response Index (CRI) as a way of requesting students to provide their own assessment of the confidence they have in the correctness of their answers. If the certainty of response is low, the implication is that guesswork played a major role in the determination of the answer, which in turn implies that the student has lack of knowledge. If the certainty of response is high, it means that the student has high confidence in his/her choice of the laws and methods used to arrive at the answer. Hasan *et al.*'s (1999) certainty of response will be used in conjunction with the answer to an item to differentiate between lack of knowledge and the presence of alternative conceptions.

Four possible combinations of the answer (correct or incorrect) and certainty of response (high or low) are shown in Table 3.1, on the part of an individual student. For a given item, if a student chose the correct answer and reported a low certainty of response, it would indicate lack of knowledge; this implies that the student obtained the correct answer through guessing. If a student chose an incorrect answer and reported a low certainty of response, this would imply that the student lacks knowledge of the concept. If a student chose a correct answer and reported a high certainty of response, this would signify knowledge of the correct concepts. The student is classified as having adequate knowledge and understanding of the concept. If a student chose an incorrect answer and reported a high certainty of response, this would signify the presence of alternative

conceptions. The student is confident about his/her choice; however his/her confidence is misplaced. When dealing with a group of students the identification of alternative conceptions will be done in the same manner as the analysis for the individual student, but the average value for the certainty of response will be used and the percentage of students choosing the correct answer will be utilized instead of an answer given by single student.

Analysis of groups of students' responses rather than individual responses would require the use of averages for both the performance and confidence, which could render the application of the Hasan et al. (1999) model to groups rather than individual respondents to be less reliable. Whether this is indeed the case, will be investigated in this study. The manner in which Hasan et al. (1999) applied this model to a group of students as an aid to teaching activities in physics classrooms will now be presented.

A 36-item multiple-choice test was administered to students. Each item was allocated a mark of 1 if correct and a mark of 0 if incorrect. For each item, a student was requested to provide the degree of certainty he/she has in his/her ability to apply and use scientific laws and principles to arrive at an answer. The sum of all the students who answered an item correctly was divided by the total number of responses, and this gave the number of correct answers as a fraction of the total number of students. Students were also requested to express their certainty of response to each item on a six-point scale of zero (0) to five (5). For a given item, the sum of the CRI for correct answers was divided by the number of students who had given a correct answer and this gave the average CRI for a correct answer. Similarly the sum of the CRI for incorrect answers was divided by the number of students who had given an incorrect answer and this gave the average CRI

for incorrect answers, for a given item. In order to decide whether a CRI value is low or high, the authors adopted a threshold value of 2.5. A CRI value above or below 2.5 was considered high or low, respectively.

Table 3.1. Decision matrix for an individual student and for a given question, based on combinations of correct or wrong answers and of low or high average CRI (Hasan et al., 1999: 296).

	Low CRI (<2.5)	High CRI (>2.5)
Correct answer	Correct answer and low CRI. Lack of knowledge (lucky guess)	Correct answer and high CRI. Knowledge of correct concepts
Wrong answer	Wrong answer and low CRI. Lack of knowledge.	Wrong answer and high CRI. Misconceptions

Average CRI values for correct and incorrect answers were utilized in conjunction with a fraction of students choosing the correct answer, to decide whether students have alternative conceptions or are lacking knowledge and understanding of principles and concepts. For example, a high average CRI for a correct answer and a high average CRI for an incorrect answer coupled with a low fraction of students choosing the correct answer is interpreted to suggest the presence of alternative conceptions. While a low average CRI for a correct answer and low average CRI for an incorrect answer coupled with low fraction of students choosing the correct answer is interpreted to suggest lack of knowledge of the principles and scientific laws. In situations where, for a given item, the average CRI value for correct and incorrect answers were very close to 2.5, the authors

utilized the fraction of correct answers in order to decide whether the CRI value is high or low, and hence decide on the presence of alternative conceptions. For example, if the CRI for incorrect answers is very close to 2.5, and a large fraction of students have chosen the incorrect answers, then the implication is the large number of students who have chosen the incorrect answers were quite confident about their choices. This situation thus signals that a large fraction of students have alternative conceptions. The suggestion is that teaching should be geared toward addressing the specified alternative conception. On the hand, if the CRI for correct answers is very close to 2.5, and a large fraction of students have chosen the correct answers, then the implication is that a large number of students have chosen the correct answers and are not quite confident about their choices. This would indicate that only a small fraction of students has alternative conceptions. In this case the authors suggest that nothing special needs to be done, since only a small fraction of students seem to have alternative conceptions.

In this study the Hasan et al. (1999) model as described above will be used for individual students as well as for groups of students. The percentage of students choosing the incorrect answer in conjunction with their average confidence levels will be used to gauge the strength of the alternative conceptions among the groups of students. Item difficulty, in terms of the percentage of students choosing the correct answer, and the difference in average confidence associated with the correct and incorrect answer will be used also to determine the accuracy in the judgment of the knowledge and understanding of correct concepts. Positive and large differences would imply the justified confidence in the laws and methods used to arrive at the answer. Negative and smaller differences would signify poor judgment in the knowledge and understanding of correct concepts. In

addition to the four possible combinations used by Hasan et al. (1999), coded written responses and the percentage of students, from the UL and UPmaj cohorts, will be utilized (i) to distinguish between the students' lack of knowledge and the presence of alternative conceptions, and (ii) to determine the strength of the misconceptions among the students.

CHAPTER 4 RESEARCH METHODOLOGY

4.1. Research Design

The study followed a correlational research design, which explored relationships between variables. The variables in this study are performance in the test instrument; knowledge and understanding of basic mechanics concepts, and confidence levels. This study attempts to determine whether relationships exist between these variables. The study is meant to establish the presence of a relationship(s), and not to establish a cause-effect relationship (Gall *et al.*, 1996; Gay, 1987).

4.2. Test Design

A test instrument consisting of thirty (30) multiple-choice items was developed from existing standardized tests from the literature. The test was first given to grade 12 physical science teachers in the Limpopo Province and physics lecturers from UL during the month of August 2005, and then piloted to the UNIFY students at UL, during October 2005. Based on the results of the pilot study and the comments of the educators, the test was modified and refined, and five items were removed from the test.

The final paper-and-pencil test consisting of twenty-five (25) items had two sections. Section A, consisting of 5 items, required the students to report on their educational background. In this section the students had to indicate gender, language used at home (mother tongue), language of instruction used at the secondary school he/she attended, the language used by their grade 12 physical science teacher, and the type of secondary school he/she attended. Section B, consisting of 20 items, probed for students'

conceptual understanding six of the concepts in mechanics. Items in Section B were obtained from existing sources in the literature. Twelve (12) items were taken from the Force Concept Inventory (FCI) (Hestenes *et al.*, 1992) and seven (7) items taken from the Mechanics Baseline Test (MBT) (Hestenes & Wells, 1992). In order to get a reasonable number of items covering the six chosen concepts in mechanics, one item was obtained and adapted from a question in "The Physics Classroom" from the Internet (www.http://physicsclassroom.com/Newtonlaws/html). This item addresses Newton's Second Law of Motion.

The two tests, FCI and MBT, were chosen because they are complementary to each other. They probe for the students' understanding of the most basic concepts in mechanics, and the scope of the tests is limited to the concepts that are addressed in elementary physics at levels starting from high school. Items in the test instrument focused on the students' conceptual understanding of basic mechanics, no mathematics was required (Hestenes *et al.*, 1992).

Each item in Section B of the test instrument had three parts. The first part was a statement in the form of a question followed by five options (A, B, C, D, and E) to choose from. The second part required that the students give written explanations for their chosen options. This part was included so that the student's knowledge and understanding of certain concepts could be explored (Planinic *et al.*, 2006). The third part required that the students indicate their confidence levels, given by a certainty of response index (CRI) (Hasan *et al.*, 1999). The student indicated on a scale of A to D their confidence in the correctness of their answer [certain (D), almost certain (C), almost a guess (B), or a totally guessed answer (A)].

Table 4.1 Basic mechanical conceptual dimensions included in the test instrument (adapted from Hestenes *et al.*, 1992 (FCI); Hestenes & Wells, 1992 (MBT))

	Concept	Item number (correct option)
1	Kinematics: Velocity discrimination from position, Acceleration discrimination from velocity Constant acceleration	15(E) 16(D) 9(D), 19(B), 20(D), 23(E)
2	Newton's First Law: No force Canceling forces	24(B), 25(B) 10(B), 11(B), 13(B), 18(C)
3	Newton's Second Law: Constant acceleration, Direction of acceleration and the resultant force, Dependence on mass	23(E) 21(A,D) 12(D), 17(A)
4	Newton's Third Law: Impulsive forces	7(E), 11(B)
5	Superposition Principle: Vector sum Canceling forces	14(B), 10(B), 11(B) 13(B), 18(C), 22(C)
6	Gravitation Acceleration independent of weight Parabolic trajectory	8(D) 6(C) 9(D)

The understanding and interpretation of forces form a foundation for basic mechanics in physics. Once a student has good understanding of forces, the interpretation and application of forces in the other dimensions of mechanics and physics in general do not pose a problem. The test items chosen for this study address concepts like: linear and

circular motions; Newton's laws of motion; superposition principle; and gravitational free-fall. Table 4.1 lists the mechanical concepts involved, the corresponding item number in the test, and the scientifically acceptable answer for each item is given in brackets.

4.3. Pilot Study

A test instrument consisting of 30 items was first given to twenty educators, i.e. sixteen grade 12 physical science teachers in four circuits of the Limpopo Province, and four physics lecturers at the University of Limpopo, during the month of August in 2005. The educators were chosen because of their willingness to participate in improving and refining the test instrument. The test instrument was given to: three teachers from the Bolobedu circuit, seven teachers from Mankweng circuit, two teachers from Thabammoopo circuit, three teachers from Mahwelereng circuit, three lecturers from mainstream physics and one lecturer from the physics section of the University Foundation Year (UNIFY) programme at the University of Limpopo (Turfloop campus). The purpose of giving the test instrument to the educators was to verify whether the test items are based on concepts that are within the physics content of the grades 11 to 12 physical science curriculum as prescribed by the South African Department of Education, and also to check for clarity of language and presentation of the test items. The UNIFY Programme was initiated in 1992 at the then University of the North (now known as the University of Limpopo, as a result of the merger between the University of the North and the Medical University of South Africa, which took place on the 1st of January 2005) to improve the number and quality of learners from disadvantaged backgrounds admitted

into the science faculties. It is a pre-degree foundation year programme in Mathematics, Sciences, and English and Study Skills. The students entering UNIFY do not satisfy the necessary requirements to be admitted into science degree programmes directly (Smith & Cantrell, 1995; Mabila *et al.*, 2006).

The test instrument was then piloted at UNIFY during the month of October 2005. A sample of 115 UNIFY students wrote the test. Based on the comments from the educators, and the results of the pilot study some items were removed, others were refined and improved for inclusion in the final test to be used in the main study. In its first version the test instrument had a section where the students were required to rephrase the questions in their own simple words. The results of the pilot study indicated that the majority of the students skipped the section without answering it, and it was felt that it (the section) does not serve the intended purpose, and was therefore removed. After these modifications and refining, the final test instrument had 25 items. The final version of the test instrument is included as Appendix B.

4.4. Sample

The final version of the test instrument was administered to first entering physics students at three South African universities, at the beginning of the year, during the month of February 2006, before any formal instruction could take place. The participants in the main study were 982 first entering students registered for physics at the three universities (UL, UP and UCT), in South Africa. These are the students who study physics either as a major or as an additional course. At UL three groups of first entering physics students took part in the study. The groups were: 102 students

registered for the foundation physics module, PHYS 010, as part of the University of Limpopo Foundation Year (UNIFY) programme; 43 mainstream students who registered for the physics module, PHYS 111, and may opt to take it as a major course in their degree programmes; and 79 students who registered for the physics module, PHYS 151, offered as a service course for science professional degrees (e.g. nursing, pharmacy, nutrition, etc.).

At UP four groups of students took part in the study. The groups were: 33 mainstream students who registered for the physics module PHY171, and may opt to take it as a major course in their degree programmes; 31 students registered for the physics module JFK110, as offered for professional science teachers degrees, 68 students registered for module PHY101, as part of the Extended Programme for students who are under-prepared for mainstream science, and 483 students registered for module PHY131, offered as a service course for biological sciences degrees.

One group of students at UCT took part in the study: 143 students who registered for the physics module, as part of their Academic Development Programme.

For the purpose of this study the student modules will be given different codes, as shown in Table 4.2 below. The capital letters in the code refer to the tertiary institution and the lower case letters describe the cohort as follows: teach refers to teachers in training; adp refers to academic development programme; sc refers to a service course; maj refers to a cohort who may select physics as major for their B.Sc degree and fy refers to a foundation programme.

Table 4.2 Codes of student groups participating in the study

Name of Module	Description	Code to be used in the study
JFK110	Physics course for pre-service teachers at UP	UPteach
PHY101	Physics course in Extended Programme at UP	UPadp
PHY131	Physics service course at UP	UPsc
PHY171	Physics for science majors at UP	UPmaj
UCT ADP	Physics course in Academic Development Programme at UCT	CTadp
PHYS010	Foundation physics course at UL	ULfy
PHYS151	Physics service course at UL	ULsc
PHYS111	Physics for science majors at UL	ULmaj

4.5. Test Validity and Reliability

4.5.1. Content Validity

Validity is concerned with the degree to which a test measures what it is supposed to measure (Blaikie, 2004; Gall *et al.* 1996; Gay 1987). Sixteen grade 12 physical science teachers from four circuits in the Limpopo Province, and four physics lecturers from UL, were given the test instrument and asked to evaluate the test individually before it was administered to the students. The educators were asked to assess the test and make judgment concerning how well the items represented the intended content area. The intended content area in this case is mechanics in the physics section of the grade 12 physical science curriculum in the South African context.

4.5.2. Reliability

Reliability is defined by Gay (1987) as the degree to which a test consistently measures whatever it is supposed to measure. Reliability is usually expressed as a numerical coefficient; a high coefficient indicates high reliability and a low coefficient indicates low reliability. Different methods are used to assess the reliability of a test. For the purpose of this study the split-half and the Cronbach alpha methods were used to assess the reliability of the test.

4.5.2.1. Split-Half Method

The test was divided into two halves, i.e. in odd numbered and even numbered items. The scores for the individual students in each half were computed. Each student had two scores, a score for the odd numbered items and a score for the even numbered items. The two sets of scores were correlated. Using data obtained in this study and the computer software SPSS, the split-half reliability of the test was found to be 0.57. However, literature indicates that longer tests tend to be more reliable, and the split-half reliability represents the reliability of a test that is half as long as the actual test. The Spearman-Brown prophecy formula (Gall *et al.* 1996; Gay 1987) was thus applied to the correlation. For example, the split-half reliability coefficient for a 20-item test was found to be 0.57. The 0.57 was based on the correlation between scores on 10 even items and 10 odd items, not on a 20-item test. The Spearman-Brown prophecy formula (Gall *et al.* 1996; Gay 1987) was needed to estimate the reliability of the 20-item test. The formula is given as

$$r_{total-test} = \frac{2 \times r_{split-half}}{1 + r_{split-half}}$$

Where r represents the split-half reliability coefficient, and applying the formula

$$r_{total} = \frac{2 \times 0.57}{1 + 0.57} = 0.73$$

Therefore the split-half estimate for 20 items was calculated to be 0.73. The test reliability coefficient for the test is therefore 0.73.

4.5.2.2. Cronbach Alpha Coefficient

Cronbach alpha coefficient is another approach to measure the internal reliability of a test. According to Blaikie (2004), Gall *et al.* (1996) and Gay (1987), the value of this coefficient ranges from 0 to 1; a high value indicates high level of consistency among the test items, while a low value indicates low level of consistency among the test items. Using the data obtained in this study the Cronbach alpha coefficient was calculated using computer software package called the SPSS. The Cronbach alpha coefficient for the test items was calculated as 0.68. Comparing the value obtained when using the split-half method and the value obtained when using the Cronbach alpha method, one realizes that the two values, 0.68 and 0.73, are comparable. The values revolve around a value of about 0.7. A correlation coefficient of about 0.7 is regarded as a high value. This value indicates that there is consistency among the test items and that test reliability is high.

4.6. Ethical Issues

The participants were informed about the study and had an option of not participating in the study (Onwuegbuzie, 2001). They were, in addition to this

information, given a consent form (Appendix A) to read and complete before the start of the test (Witt-Rose, 2003). Information about students' responses and participation is kept strictly confidential, and the researcher is the only one having access to this information. Acronyms and codes were used instead of the students' names and student numbers. Data is stored in both hard copy and electronic form (CDs and flash discs) in a safe place at the offices of the UNIFY programme.

CHAPTER 5 RESULTS AND ANALYSIS

5.1. Introduction

The results of the first five questions on the school background of the participants will be reported both in a narrative form as well as in a table form (Appendix E). This study was aimed at investigating the relationship between performance and confidence levels, and to investigate the presence and the strength of alternative conceptions among first entering students at some universities. The influence of factors such as school background, gender, language of instruction, etc. on performance and confidence levels of students is however, beyond the scope of the study, and probably may be revisited for future work. The conceptual understanding of the students will be analysed using item by item analysis which will also be given in both narrative and tabular form.

5.2. Educational Background

Participants in the study were 468 male students and 514 female students. The students had different home languages i.e. 434 students had an African language as their home language, 264 students were Afrikaans speaking, and 257 were English speaking, 15 had "Another European language" as their home language, while 12 had their home language classified as "Other". In the South African Education System, initially the official medium of instruction at secondary schools was either English or Afrikaans. However, this has since changed; all the eleven official languages can be used as medium of instruction at secondary schools. From the results it was found that 50 students had their medium of instruction as an African language, and Afrikaans was a medium of

instruction for 252 students, while 679 students had English as the medium of instruction, and one student was taught in "Another European language" at the secondary school. The language of instruction used by the Physical Science teacher at secondary schools is found to differ among the students. In this study 34 students indicated that their teachers used an African language as a medium of instruction, and 231 students indicated that their teachers used Afrikaans as the medium of instruction, while 717 students indicated that their teachers used English as the medium of instruction. Of these students who participated in the study, 196 students attended grade 12 at a private school, 110 students attended grade 12 at a township secondary school, 14 students attended grade 12 at a farm secondary school, and 208 students attended grade 12 at a secondary school in rural areas, while 454 students attended grade 12 at a secondary school in a town or city. A detailed demographic background of the students who participated in the study is given as Appendix E.

5.3. Conceptual Understanding

The performance in Section B of the test was calculated by allocating a score of one (1) for the correct option and a score of zero (0) for the incorrect option chosen for an item by a student. The scores for the individual students were added to obtain an average performance score for each cohort. There were twenty items in Section B, which gives a potential maximum performance score of 20. For the purpose of this study, and in deciding whether a student's score is high or low, a threshold value of 10 was adopted. A total score above or below 10 is considered to be high or low, respectively. Raw scores indicating the performance of the individual students obtained in the test are shown in

Appendix C, for each of the eight groups. 36% of students from the eight groups have their performance scores above the threshold, while 64% of the students have performance scores below the threshold.

The confidence levels of the students were calculated by allocating a score of zero (0) for choosing option A (a totally guessed answer), a score of one (1) for option B (almost a guess), a score of two (2) for option C (almost certain), and a score of three (3) for option D (certain). A score for the confidence level of an individual student was obtained by calculating the average of the scores obtained by a student in all twenty items. For the purpose of this study, and in order to decide whether a student's confidence level is high or low, a threshold value of 1.5 was adopted. An average score above or below 1.5 is considered to be high or low, respectively. The average confidence levels of individual students are shown in Appendix C, for each of the eight groups. 87.5% of the students have an average confidence level which is above the threshold, and 12.5% of the students have an average confidence levels below the threshold.

The overall performance and confidence levels of students participating in the study are shown below in tabular form. From Table 5.1, one realizes that the test performances of students from the eight groups differ. Students from groups UPsc and UPmaj have a higher average test performance as compared to the rest of the students, their average test performances are above the threshold of 10, i.e. 10.1 and 10.9, respectively. The other six groups are regarded as having obtained low scores, because their average test performances are below the threshold of 10. Students from group ULfy have the lowest average test performance, 5.6. The students in all the eight groups have

high confidence levels. The average confidence levels of students in the eight groups are above the threshold of 1.5.

Table 5.1 Average performance and average confidence levels of students from the eight student cohorts

Student groups	**Number of students**	**Average Test Performance** (Maximum 20)	**Average Confidence Level**	**Correlation coefficient**
UPteach	31	6.5	1.9	0.44
UPadp	68	7.6	2.0	0.30
UPsc	483	10.1	2.1	0.42
UPmaj	33	10.9	2.2	0.57
CTadp	143	7.5	1.9	0.25
ULfy	102	5.6	1.9	0.23
ULsc	79	6.9	2.0	0.07
ULmaj	43	7.3	2.0	0.05

The nature of the relationship between test performance and average confidence levels in the eight cohorts that took part in the study is shown in a tabular form. The test performance and average confidence levels of individual students, from the eight groups, were correlated and the results are given in the last column of Table 5.1, and also in the scatter plots shown in Appendix D. The correlation coefficient of 0.57 between performance and confidence levels for students from group UPmaj is fairly strong. This

indicates that a high confidence level is usually associated with a high score in performance. The correlations between performance and confidence levels for UPteach and UPsc are both weak and the correlation for UPadp, CTadp and ULfy are all very weak. In the case of groups ULsc and ULmaj there is virtually no correlation between performance and confidence levels. The correlation coefficient in the region of 0.06 that was obtained for these groups implies that a high (or low) confidence level is equally likely to be associated with a high or low test performance. This lack of correlation is also shown in the scatter plots of performance versus confidence level for the two groups of students, shown in Appendix D ((g) and (h)).

According to Hasan *et al.* (1999) model, the students were divided into four groups according to their performance and confidence levels. Table 5.2 reports the population of the four groups.

Table 5.2 Overall performance matrix of all the students and for all the test items (N = 982).

	Low Confidence Level ($CL^b < 1.5$)	High Confidence Level ($CL^b > 1.5$)
High Test Performance ($TP^a > 10$)	19 (1.9%) Lack of knowledge (lucky guess)	335 (34.1%) Knowledge of correct concepts
Low Test Performance ($TP^a < 10$)	104 (10.6%) Lack of knowledge	524 (53.4%) Misconceptions

[a] TP represents test performance; [b] CL represents confidence level.

From Table 5.2 above, 19 (1.9%) of the students had high performance and low average confidence levels. According to Hasan *et al.* (1999) these students are regarded as having made lucky guesses. Low average confidence levels indicate that the students have reported that they have guessed or almost guessed the answer. 104 (10.6%) of the students scored low on the test and had low average confidence levels. These students are classified as having lack of knowledge and understanding of the concepts. They have performed poorly and they were aware that their knowledge and understanding were inadequate. 524 (53.4%) of the students scored low on the test and have high average confidence levels. According to Hasan *et al.* (1999) this is an indication that these students have strong alternative conceptions about concepts in mechanics. The students are confident about the choices they are making, even though their choices are incorrect. 335(34.1%) of the students scored high on the test and have high confidence levels. These students are classified as having acceptable knowledge and understanding of concepts, which are associated with justified confidence.

5.4. Item by Item Analysis

Item analysis is one of the important activities in test development. Gall *et al.* (1996) describe item analysis as a set of procedures that is used to determine the difficulty, the validity and reliability of each item in the test. One can also use item analysis to find out whether distractors for a particular item are good or bad. The specific procedures are determined by the purpose of the test. In this study item analysis was carried out in order to determine the performance and confidence levels of the different

groups of students, and to judge item difficulty in terms of the number and percentage of students answering an item correctly.

Hasan *et al.* (1999) used test performance and the certainty of response index (CRI) to determine the presence of alternative conceptions among students. Planinic *et al.* (2006) used linear measures of item difficulty and student confidence to asses the strength of known alternative conceptions. They documented that poor performance in the test (low scores) coupled with high confidence indicate that the student has alternative conceptions for a particular concept. This instrument has been designed so that one can pick up alternative conceptions not just indirectly from specific distractors included in the multiple-choice component of each item, but also from the written explanations that the students provided. Chase (1999) documented that multiple-choice tests can be used to assess factual levels of learning, but they are poor at assessing higher order of cognition.

Item analysis was done with respect to the general performance of the whole student population who sat for the test. Item by item analysis in this case was carried out in two ways, i.e. analysis of the multiple-choice section and analysis of the student's written responses. The analysis of the multiple-choice was carried out for all the students who participated in the study, and analysis of the written responses was carried out for a selected sample of the students who participated in the study. The students from UL were selected for the analysis of the written responses because of their geographical location, which was convenient for the researcher. The UPmaj cohort was used as a benchmark in this analysis for two reasons: (i) This group of students achieved the best performance in the test and (ii) was able to judge their answers more accurately than any of the other

groups, as reflected by the correlation coefficient of 0.57 between average performance and average confidence level.

5.5. Analysis of the Multiple-Choice Section

This analysis was carried out by comparing and interpreting the percentage of students choosing the different options for each item. The analysis was carried out for all eight groups of students involved in the study. The analysis of each item is given in a table form as well as in a narrative form below. The table indicates the percentage of students choosing each of the five options for each item, and the average confidence levels, for each group. The correct option in each item is given in brackets and underlined. Refer to Appendix B for the items in the test instrument.

There are three average confidence levels in each of Tables 5.3(a – t) below. The average confidence levels marked "All" indicates the average confidence levels of all the students in a group. This value is calculated by dividing the sum of the confidence levels for all the students by the number of students. The average confidence level marked "Correct Option" indicates the average confidence levels of the students who answered the item correctly. This value is calculated by dividing the sum of the confidence levels for the students who answered the item correctly by the number of students who chose the correct option. The average confidence level marked "Incorrect Option" indicates the average confidence level of the students who answered the item incorrectly. This average confidence level is calculated by dividing the sum of confidence levels of students who answered the item incorrectly by the number of students choosing the incorrect options.

The analysis is done this way in order to compare the confidence levels of the students choosing the correct option and those choosing the incorrect option.

A comparison of the average confidence levels associated with correct and incorrect answers, respectively, provides an indication of the quality of judgment made within a specific cohort about the correctness of answers provided for a specific test item.

5.5.1. Item 6

This item deals with two metal balls, of different masses, dropped simultaneously from the top of a two-storey building. Students were to indicate which ball would reach the ground first. The item is found in the conceptual dimension of Gravitation (Table 4.1), and addresses the conception that acceleration of falling objects is independent of the weight of objects. Distractors A and D, the most prevalent alternative conception documented in literature, reflect the idea that "Heavier objects fall faster than lighter objects" (Gunstone & White, 1981; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1992). The less prevalent distractors B and E reflect the alternative conception that lighter objects fall faster than heavier objects. Table 5.3(a) below indicates the percentage of students, for each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(a) Performance and Confidence levels of all students for item 6

Student Group	Number of Students	Options (%)							Average Confidence Level		
		A	B	(C)	D	E	A+D	B+E	All	Correct Option	Incorrect Options
UPteach	31	19.4	3.2	45.2	19.4	12.9	39.8	16.1	2.2	2.2	2.1
UPadp	68	5.9	13.2	63.2	13.2	4.4	19.1	17.6	2.3	2.4	2.2
UPsc	483	5.4	6.0	74.2	9.4	5.0	14.8	20.8	2.4	2.5	2.0
UPmaj	33	15.2	6.1	78.8	0.0	0.0	15.2	6.1	2.6	2.7	2.0
CTadp	143	13.3	7.7	51.7	22.4	4.9	35.7	12.6	2.1	2.3	1.9
ULfy	102	13.7	21.6	34.3	18.6	11.8	32.3	33.4	2.4	2.6	2.3
ULsc	79	15.2	12.7	35.4	19.0	17.7	34.2	30.4	2.3	2.4	2.3
ULmaj	43	18.6	11.6	48.8	14.0	7.0	32.6	18.6	2.4	2.4	2.5

The correct option in this item was option C, that the two metal balls will reach the ground simultaneously. More than 50% of the students from UPadp, UPsc, UPmaj and CTadp have chosen the correct option; while less than 50% of the students from UPteach, ULfy, ULSC and ULmaj have chosen the correct option. However, a number of students have the belief that the heavier metal ball will reach the ground first. This is indicated by the high percentage of students in all the eight groups choosing options A and D. The responses obtained for ULfy are randomly distributed between the options C (34%), A + D (32.3%) and B + E (33.4%). Distractors A and D are similar in the sense

that they reflect the conception that a heavier ball will reach the ground significantly faster than the lighter ball. In the same way distractors B and E reflect the conception that the lighter ball will reach the ground significantly faster than the heavier ball.

The average confidence levels for the students in all the eight groups are found to be high, ranging from 2.1 to 2.6, indicating that the students are very confident about their choices. The students who chose the correct option are confident about their answers; this is indicated by the confidence levels ranging from 2.2 to 2.7, and the students who chose the incorrect options are also confident about their choice, with their confidence levels ranging from 1.9 to 2.5 (shown in Table 5.3(a) above). The difference between the average confidence values associated with a correct answer and with the combination of wrong answers for a specific group is indicative of the quality of judgment about the correctness of the answer provided that the students of that group are capable of. Students in cohorts UPsc and UPmaj showed both the highest performance and the best quality of judgment on this item.

5.5.2. Item 7

The item deals with a head-on collision between a large truck and a small car. The students were to indicate the forces involved during this interaction between the two vehicles. The item is located in the conceptual dimension of Newton's third law (Table 4.1) for impulsive forces. The alternative conceptions documented in the literature are: Distractors A and D, that "a greater mass implies a greater force", and distractor C, that "obstacles exert no force" (Halloun *et al.*, 1985b; Maloney, 1984). Table 5.3(b) below

indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(b) Performance and Confidence levels of all students for item 7

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	D	(E)	All	Correct Option	Incorrect Options
UPteach	31	64.5	6.5	3.2	0.0	25.8	2.4	2.4	2.4
UPadp	68	58.8	1.5	0.0	1.5	38.2	2.2	2.2	2.2
UPsc	483	40.3	1.0	0.8	0.4	57.4	2.4	2.5	2.3
UPmaj	33	39.4	3.0	0.0	0.0	57.6	2.5	2.7	2.2
CTadp	143	42.7	4.2	0.0	1.4	51.7	2.3	2.3	2.2
ULfy	102	66.7	3.9	1.0	2.0	26.5	2.4	2.4	2.3
ULsc	79	59.5	2.5	6.3	3.8	27.8	2.4	2.5	2.4
ULmaj	43	51.2	2.3	0.0	2.3	44.2	2.4	2.5	2.3

The correct answer for this item was option E, i.e. both vehicles exert equal force on each other. The table above indicates that, in general the performance on this item was poor, more than 50% of the students from UPsc, UPmaj and CTadp have chosen the correct option, while less than 50% of the students from UPteach, UPadp, ULfy, ULsc and ULmaj have chosen the correct option. The concept assessed in this item is clearly more difficult to grasp. Option A was the only meaningful distractor for all of the eight cohorts, i.e. the bigger vehicle exerts a greater amount of force on the smaller vehicle,

while the smaller vehicle exerts a smaller amount of force on the bigger vehicle. Distractors B, C and D were very weak and attracted less than 10% of the responses in seven of the eight groups.

The students in all eight groups were very confident about their chosen options. This is evident from the table above, which reflects average confidence levels ranging from 2.3 to 2.7. The table also reflects that the students who chose the correct option and those students, who chose the incorrect options, are both confident about their choice. The students who chose the correct option have confidence levels ranging from 2.2 to 2.7, while those who have chosen the incorrect options have confidence levels ranging from 2.2 to 2.4. The only exception to this trend is UPmaj where the majority of students either knew that their answers were correct or realized that they may be incorrect.

5.5.3. Item 8

This item deals with a steel ball being thrown vertically upwards, with the effect of air resistance being ignored. Students had to identify the force(s) exerted on the ball during the course of its flight. The item is found in the conceptual dimension of gravitation (from Table 4.1). The most common alternative conceptions found in physics education literature are, distractor A "impetus dissipation", distractor B "gravity increases as object falls, gravity acts after impetus wears down", distractor C "delayed impetus build-up" and distractor E "gravity is intrinsic to mass" (Gunstone *et al.*, 1981; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1985). Table 5.3(c) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(c) Performance and Confidence levels of all students for item 8

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	6.5	35.5	45.2	12.9	0.0	1.9	2.0	1.9
UPadp	68	5.9	27.9	50.0	14.7	1.5	2.2	2.3	2.2
UPsc	483	3.5	8.1	58.2	29.9	0.2	2.4	2.6	2.3
UPmaj	33	3.0	9.1	51.5	36.4	0.0	2.4	2.7	2.2
CTadp	143	4.2	21.8	56.3	16.9	0.7	2.2	2.4	2.1
ULfy	102	14.7	28.4	44.1	11.8	1.0	2.2	2.3	2.2
ULsc	79	13.9	19.0	35.4	27.8	3.8	2.2	2.2	2.2
ULmaj	43	9.3	11.6	48.8	27.9	2.3	2.2	2.3	2.2

The correct answer for the item was option D, that is only a constant gravitational force is acting on the ball until it returns to the ground. From the table above it is evident that less than 50% of students in all eight groups have chosen the correct option D, while a higher percentage (ranging from 35.4% to 58.2%) of students in all eight groups have chosen option C, i.e. the forces acting on the ball until it returns to the ground are a constant gravitational force together with an upward force that decreases as the ball goes up. The alternative conception reflected by distractor B is important for all cohorts, except UPsc and UPmaj. Distractor E is too weak to contribute to the analysis. Distractor A is weak for all groups except for the UL groups. According to Halloun & Hestenes

(1985a) the students believe that as the ball goes up, the upward force wears down. This misconception is more prevalent within UL cohorts than in the other cohorts.

The students in all the groups were confident about their chosen options. This is evident from the table above, which reflects average confidence levels from 1.9 to 2.4. Table 5.3(c) also reflects that the students who have chosen the correct option and those students, who have chosen the incorrect options, were confident about their choice. The confidence levels of students choosing the correct option ranges from 2.0 to 2.7, while those students who have chosen the incorrect options have their confidence levels ranging from 1.9 to 2.3. Significantly, the largest difference between average confidence associated with correct answers and average confidence associated with incorrect answers was observed in the cohort UPmaj, a group that also achieved the highest performance. This result is interpreted to mean that the better performing students were able to make better quality judgments about the correctness of their answers.

5.5.4. Item 9

This item deals with a bowling ball accidentally falling from the cargo of an airliner which is flying in a horizontal direction. Students had to identify the path that will most likely be followed by the ball, as seen by an observer on the ground. The item is found in the conceptual dimensions of kinematics and gravitation (Table 4.1), and addresses the conception that "constant acceleration entails parabolic trajectory". The most common alternative conceptions documented in the literature are: Distractors A and B that "mass makes objects stop", distractor C that "force compromise determines motion" and distractor E that "gravity acts after impetus wears down, impetus

dissipation" (Halloun & Hestenes, 1985b; Jimoyiannis *et al.*, 2001). Table 5.3(d) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(d) Performance and Confidence levels of all students for item 9

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	35.5	29.0	6.5	25.8	3.2	2.0	2.0	2.0
UPadp	68	33.8	8.8	16.2	36.8	4.4	2.3	2.5	2.1
UPsc	483	27.4	13.9	10.8	45.3	2.5	2.2	2.4	2.0
UPmaj	33	27.3	0.0	15.2	57.6	0.0	2.2	2.4	2.0
CTadp	143	42.1	15.0	10.0	27.1	5.7	2.0	2.1	1.9
ULfy	102	52.9	15.7	9.8	14.7	6.9	2.0	2.5	2.0
ULsc	79	16.5	19.0	10.1	49.4	5.1	2.4	2.6	2.2
ULmaj	43	14.0	4.7	16.3	58.1	7.0	2.2	2.2	2.3

The correct option was D, that as seen from the ground the bowling ball will follow a parabolic path forward while falling down. The two mainstream cohorts, UPmaj and ULmaj, have shown the best average performance on this item, i.e. 57.6% and 58.1%, respectively, while less than 50% of the students from UPteach, UPadp, UPsc, CTadp, ULfy and ULsc have chosen the correct option. Distractors A – C feature with varying prominence for the eight groups and distractor E is too weak to be significant for

the analysis. As compared to distractors B and C, a higher percentage of students in six of the eight groups have chosen the incorrect option A, that as seen from the ground the bowling ball will follow a parabolic path backward while falling down.

Students from all the eight groups are very confident about their chosen options; the table above indicates that the confidence levels of the students ranges from 2.0 to 2.4. Table 5.3(d) also reflects that the students who have chosen the correct option and those students, who have chosen the incorrect options, are confident about their choice. The confidence levels of students choosing the correct option ranges from 2.0 to 2.6, while those students who have chosen the incorrect options have their confidence levels ranging from 1.9 to 2.3. A comparison of the difference between average confidence values for correct and incorrect answers indicates that four groups showed reasonable accuracy of judgment. Significantly poorer accuracy of judgment is observed for UPteach and ULmaj.

5.5.5. Item 10

The item deals with two blocks of equal masses hanging from the ceiling of an elevator by means of two strings. Students were to determine the magnitude of the force exerted by rope 1 on block I when the elevator goes upwards at constant velocity. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1). Distractor A includes a common alternative conception that "$F = m \times v$" (Clement J., 1982; Hestenes *et al.*, 1992). The mistake that students make arises from confusing speed and acceleration, which they use interchangeably. Distractor C, D and E are about the direction of motion and the magnitude of the force. Table 5.3(e)

below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(e) Performance and Confidence levels of all students for item 10

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	32.3	45.2	12.9	9.7	0.0	1.9	2.2	1.8
UPadp	68	20.6	42.6	7.4	27.9	1.5	1.8	2.1	1.6
UPsc	483	14.6	54.3	12.7	16.7	1.7	2.0	2.2	1.6
UPmaj	33	15.2	63.6	9.1	12.1	0.0	2.0	2.2	1.8
CTadp	143	18.3	59.2	8.5	12.0	2.1	1.8	1.8	1.7
ULfy	102	52.0	35.3	2.0	10.8	0.0	2.0	2.1	1.8
ULsc	79	25.6	50.0	10.3	14.1	0.0	2.2	2.3	2.2
ULmaj	43	20.9	72.1	4.7	2.3	0.0	2.1	2.2	1.9

More than 50% of the students in each of UPsc, UPmaj, CTadp, ULsc and ULmaj have chosen the correct option, option B, that the force exerted by rope 1 on block I has a magnitude of 10 N. Significantly the best performance is observed for the two mainstream cohorts, UPmaj and ULmaj. Distractor A is most prominent, that the forces exerted by rope 1 on block I has a magnitude of 2 N. This is a common alternative conception where students use acceleration and speed interchangeably. Distractors C, D and E, are about motion being in the direction of bigger force. Students have this

conception that motion is always in the direction of a bigger force (Maloney, 1984). Since the downward force is 10 N, the students therefore make a mistake that the force causing the object to go up must be greater than 10 N. Distractor D was more prominent than distractor A for UPadp and UPsc. However, distractors C and D vary in prominence for the different groups and distractor E is too weak to be meaningful.

Students from each of the eight groups are confident about their chosen options, their confidence levels range from 1.8 to 2.2. Table 5.3(e) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. All students are confident about their correct option, (confidence levels ranging from 1.9 to 2.3). Students choosing the incorrect options are also confident about their choices; this is indicated by the confidence levels ranging from 1.6 to 2.2 in all the groups. Noteworthy is the fact that a higher quality of judgment is observed for all UP groups compared to the others. The students were clearly more accurate with their judgment about the correctness of their answers to this item.

5.5.6. Item 11

The item deals with two blocks of equal masses hanging from the ceiling of an elevator by means of two strings. Students were to determine the magnitude of the force exerted by rope 1 on block II when the elevator is stationary. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1). Table 5.3(f) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

that they reflect the conception that a heavier ball will reach the ground significantly faster than the lighter ball. In the same way distractors B and E reflect the conception that the lighter ball will reach the ground significantly faster than the heavier ball.

The average confidence levels for the students in all the eight groups are found to be high, ranging from 2.1 to 2.6, indicating that the students are very confident about their choices. The students who chose the correct option are confident about their answers; this is indicated by the confidence levels ranging from 2.2 to 2.7, and the students who chose the incorrect options are also confident about their choice, with their confidence levels ranging from 1.9 to 2.5 (shown in Table 5.3(a) above). The difference between the average confidence values associated with a correct answer and with the combination of wrong answers for a specific group is indicative of the quality of judgment about the correctness of the answer provided that the students of that group are capable of. Students in cohorts UPsc and UPmaj showed both the highest performance and the best quality of judgment on this item.

5.5.2. Item 7

The item deals with a head-on collision between a large truck and a small car. The students were to indicate the forces involved during this interaction between the two vehicles. The item is located in the conceptual dimension of Newton's third law (Table 4.1) for impulsive forces. The alternative conceptions documented in the literature are: Distractors A and D, that "a greater mass implies a greater force", and distractor C, that "obstacles exert no force" (Halloun *et al.*, 1985b; Maloney, 1984). Table 5.3(b) below

indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(b) Performance and Confidence levels of all students for item 7

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	D	(E)	All	Correct Option	Incorrect Options
UPteach	31	64.5	6.5	3.2	0.0	25.8	2.4	2.4	2.4
UPadp	68	58.8	1.5	0.0	1.5	38.2	2.2	2.2	2.2
UPsc	483	40.3	1.0	0.8	0.4	57.4	2.4	2.5	2.3
UPmaj	33	39.4	3.0	0.0	0.0	57.6	2.5	2.7	2.2
CTadp	143	42.7	4.2	0.0	1.4	51.7	2.3	2.3	2.2
ULfy	102	66.7	3.9	1.0	2.0	26.5	2.4	2.4	2.3
ULsc	79	59.5	2.5	6.3	3.8	27.8	2.4	2.5	2.4
ULmaj	43	51.2	2.3	0.0	2.3	44.2	2.4	2.5	2.3

The correct answer for this item was option E, i.e. both vehicles exert equal force on each other. The table above indicates that, in general the performance on this item was poor, more than 50% of the students from UPsc, UPmaj and CTadp have chosen the correct option, while less than 50% of the students from UPteach, UPadp, ULfy, ULsc and ULmaj have chosen the correct option. The concept assessed in this item is clearly more difficult to grasp. Option A was the only meaningful distractor for all of the eight cohorts, i.e. the bigger vehicle exerts a greater amount of force on the smaller vehicle,

while the smaller vehicle exerts a smaller amount of force on the bigger vehicle. Distractors B, C and D were very weak and attracted less than 10% of the responses in seven of the eight groups.

The students in all eight groups were very confident about their chosen options. This is evident from the table above, which reflects average confidence levels ranging from 2.3 to 2.7. The table also reflects that the students who chose the correct option and those students, who chose the incorrect options, are both confident about their choice. The students who chose the correct option have confidence levels ranging from 2.2 to 2.7, while those who have chosen the incorrect options have confidence levels ranging from 2.2 to 2.4. The only exception to this trend is UPmaj where the majority of students either knew that their answers were correct or realized that they may be incorrect.

5.5.3. Item 8

This item deals with a steel ball being thrown vertically upwards, with the effect of air resistance being ignored. Students had to identify the force(s) exerted on the ball during the course of its flight. The item is found in the conceptual dimension of gravitation (from Table 4.1). The most common alternative conceptions found in physics education literature are, distractor A "impetus dissipation", distractor B "gravity increases as object falls, gravity acts after impetus wears down", distractor C "delayed impetus build-up" and distractor E "gravity is intrinsic to mass" (Gunstone *et al.*, 1981; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1985). Table 5.3(c) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(c) Performance and Confidence levels of all students for item 8

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	6.5	35.5	45.2	12.9	0.0	1.9	2.0	1.9
UPadp	68	5.9	27.9	50.0	14.7	1.5	2.2	2.3	2.2
UPsc	483	3.5	8.1	58.2	29.9	0.2	2.4	2.6	2.3
UPmaj	33	3.0	9.1	51.5	36.4	0.0	2.4	2.7	2.2
CTadp	143	4.2	21.8	56.3	16.9	0.7	2.2	2.4	2.1
ULfy	102	14.7	28.4	44.1	11.8	1.0	2.2	2.3	2.2
ULsc	79	13.9	19.0	35.4	27.8	3.8	2.2	2.2	2.2
ULmaj	43	9.3	11.6	48.8	27.9	2.3	2.2	2.3	2.2

The correct answer for the item was option D, that is only a constant gravitational force is acting on the ball until it returns to the ground. From the table above it is evident that less than 50% of students in all eight groups have chosen the correct option D, while a higher percentage (ranging from 35.4% to 58.2%) of students in all eight groups have chosen option C, i.e. the forces acting on the ball until it returns to the ground are a constant gravitational force together with an upward force that decreases as the ball goes up. The alternative conception reflected by distractor B is important for all cohorts, except UPsc and UPmaj. Distractor E is too weak to contribute to the analysis. Distractor A is weak for all groups except for the UL groups. According to Halloun & Hestenes

(1985a) the students believe that as the ball goes up, the upward force wears down. This misconception is more prevalent within UL cohorts than in the other cohorts.

The students in all the groups were confident about their chosen options. This is evident from the table above, which reflects average confidence levels from 1.9 to 2.4. Table 5.3(c) also reflects that the students who have chosen the correct option and those students, who have chosen the incorrect options, were confident about their choice. The confidence levels of students choosing the correct option ranges from 2.0 to 2.7, while those students who have chosen the incorrect options have their confidence levels ranging from 1.9 to 2.3. Significantly, the largest difference between average confidence associated with correct answers and average confidence associated with incorrect answers was observed in the cohort UPmaj, a group that also achieved the highest performance. This result is interpreted to mean that the better performing students were able to make better quality judgments about the correctness of their answers.

5.5.4. Item 9

This item deals with a bowling ball accidentally falling from the cargo of an airliner which is flying in a horizontal direction. Students had to identify the path that will most likely be followed by the ball, as seen by an observer on the ground. The item is found in the conceptual dimensions of kinematics and gravitation (Table 4.1), and addresses the conception that "constant acceleration entails parabolic trajectory". The most common alternative conceptions documented in the literature are: Distractors A and B that "mass makes objects stop", distractor C that "force compromise determines motion" and distractor E that "gravity acts after impetus wears down, impetus

dissipation" (Halloun & Hestenes, 1985b; Jimoyiannis *et al.*, 2001). Table 5.3(d) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(d) Performance and Confidence levels of all students for item 9

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	35.5	29.0	6.5	25.8	3.2	2.0	2.0	2.0
UPadp	68	33.8	8.8	16.2	36.8	4.4	2.3	2.5	2.1
UPsc	483	27.4	13.9	10.8	45.3	2.5	2.2	2.4	2.0
UPmaj	33	27.3	0.0	15.2	57.6	0.0	2.2	2.4	2.0
CTadp	143	42.1	15.0	10.0	27.1	5.7	2.0	2.1	1.9
ULfy	102	52.9	15.7	9.8	14.7	6.9	2.0	2.5	2.0
ULsc	79	16.5	19.0	10.1	49.4	5.1	2.4	2.6	2.2
ULmaj	43	14.0	4.7	16.3	58.1	7.0	2.2	2.2	2.3

The correct option was D, that as seen from the ground the bowling ball will follow a parabolic path forward while falling down. The two mainstream cohorts, UPmaj and ULmaj, have shown the best average performance on this item, i.e. 57.6% and 58.1%, respectively, while less than 50% of the students from UPteach, UPadp, UPsc, CTadp, ULfy and ULsc have chosen the correct option. Distractors A – C feature with varying prominence for the eight groups and distractor E is too weak to be significant for

the analysis. As compared to distractors B and C, a higher percentage of students in six of the eight groups have chosen the incorrect option A, that as seen from the ground the bowling ball will follow a parabolic path backward while falling down.

Students from all the eight groups are very confident about their chosen options; the table above indicates that the confidence levels of the students ranges from 2.0 to 2.4. Table 5.3(d) also reflects that the students who have chosen the correct option and those students, who have chosen the incorrect options, are confident about their choice. The confidence levels of students choosing the correct option ranges from 2.0 to 2.6, while those students who have chosen the incorrect options have their confidence levels ranging from 1.9 to 2.3. A comparison of the difference between average confidence values for correct and incorrect answers indicates that four groups showed reasonable accuracy of judgment. Significantly poorer accuracy of judgment is observed for UPteach and ULmaj.

5.5.5. Item 10

The item deals with two blocks of equal masses hanging from the ceiling of an elevator by means of two strings. Students were to determine the magnitude of the force exerted by rope 1 on block I when the elevator goes upwards at constant velocity. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1). Distractor A includes a common alternative conception that "F = m × v" (Clement J., 1982; Hestenes *et al.*, 1992). The mistake that students make arises from confusing speed and acceleration, which they use interchangeably. Distractor C, D and E are about the direction of motion and the magnitude of the force. Table 5.3(e)

below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(e) Performance and Confidence levels of all students for item 10

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	32.3	45.2	12.9	9.7	0.0	1.9	2.2	1.8
UPadp	68	20.6	42.6	7.4	27.9	1.5	1.8	2.1	1.6
UPsc	483	14.6	54.3	12.7	16.7	1.7	2.0	2.2	1.6
UPmaj	33	15.2	63.6	9.1	12.1	0.0	2.0	2.2	1.8
CTadp	143	18.3	59.2	8.5	12.0	2.1	1.8	1.8	1.7
ULfy	102	52.0	35.3	2.0	10.8	0.0	2.0	2.1	1.8
ULsc	79	25.6	50.0	10.3	14.1	0.0	2.2	2.3	2.2
ULmaj	43	20.9	72.1	4.7	2.3	0.0	2.1	2.2	1.9

More than 50% of the students in each of UPsc, UPmaj, CTadp, ULsc and ULmaj have chosen the correct option, option B, that the force exerted by rope 1 on block I has a magnitude of 10 N. Significantly the best performance is observed for the two mainstream cohorts, UPmaj and ULmaj. Distractor A is most prominent, that the forces exerted by rope 1 on block I has a magnitude of 2 N. This is a common alternative conception where students use acceleration and speed interchangeably. Distractors C, D and E, are about motion being in the direction of bigger force. Students have this

conception that motion is always in the direction of a bigger force (Maloney, 1984). Since the downward force is 10 N, the students therefore make a mistake that the force causing the object to go up must be greater than 10 N. Distractor D was more prominent than distractor A for UPadp and UPsc. However, distractors C and D vary in prominence for the different groups and distractor E is too weak to be meaningful.

Students from each of the eight groups are confident about their chosen options, their confidence levels range from 1.8 to 2.2. Table 5.3(e) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. All students are confident about their correct option, (confidence levels ranging from 1.9 to 2.3). Students choosing the incorrect options are also confident about their choices; this is indicated by the confidence levels ranging from 1.6 to 2.2 in all the groups. Noteworthy is the fact that a higher quality of judgment is observed for all UP groups compared to the others. The students were clearly more accurate with their judgment about the correctness of their answers to this item.

5.5.6. Item 11

The item deals with two blocks of equal masses hanging from the ceiling of an elevator by means of two strings. Students were to determine the magnitude of the force exerted by rope 1 on block II when the elevator is stationary. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1). Table 5.3(f) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(f) Performance and Confidence levels of all students for item 11

Module Code	Number of Students	Options (%)					Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	20.0	63.3	0.0	13.3	3.3	1.9	2.1	1.7
UPadp	68	10.3	75.0	4.4	10.3	0.0	1.8	2.0	1.4
UPsc	483	9.0	79.7	2.3	8.8	0.2	1.9	2.1	1.2
UPmaj	33	12.5	75.5	6.3	6.3	0.0	2.0	2.2	1.5
CTadp	143	12.3	73.9	2.2	10.9	0.7	1.8	1.9	1.3
ULfy	102	39.2	47.1	2.9	8.8	2.0	1.6	1.8	1.4
ULsc	79	26.2	55.7	10.1	6.3	1.3	2.0	2.0	2.0
ULmaj	43	18.6	53.5	2.3	25.6	0.0	2.1	2.0	2.1

More than 50% of the students in each of the eight groups, except for ULfy, have chosen the correct option, which is option B, that the force exerted by rope 1 on block II has a magnitude of 10 N. 47.1% of the students from ULfy have chosen the correct option. A high percentage (ranging from 9.0% to 39.2%) of students from each of the eight groups have chosen option A, that the forces exerted by rope 1 on block II has a magnitude of 2 N. Despite this being an easy item, the best performance is not observed in the two major course cohorts (UPmaj and ULmaj). At both institutions (UP and UL) the service course cohorts marginally outperformed the major course cohorts. The academic development groups at UCT and UP also performed well on this item. Options A and D are the most important distractors, with distractor A featuring more prominently

than distractor D for almost all the cohorts. Options C and E are weak distractors for all cohorts, with ULsc being the only exception.

Students from all eight groups were confident about their chosen options; this is indicated by their average confidence levels ranging from 1.6 to 2.1. Table 5.3(f) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from all eight groups were confident about their correct option (confidence levels ranging from 1.8 to 2.2. Students from UPadp, UPsc, CTadp and ULfy were not confident about their incorrect options; this is indicated by the confidence levels ranging from 1.2 to 1.4, while students from UPteach, UPmaj, ULsc and ULmaj have shown higher confidence levels, 1.5 and 2.1, respectively. The largest difference between the average confidence associated with the correct answers and the average confidence associated with the incorrect answers was observed in the UP mainstream cohorts, UPmaj and UPsc. In general, students were clearly more accurate in their judgments about the correctness of their answers in this item, with the exceptions being students from ULsc and ULmaj.

5.5.7. Item 12

This item deals with a car having a maximum acceleration of 3.0 m/s^2. The students were to determine what the maximum acceleration of the car would be when it tows a second car twice its mass. The item is found in the conceptual dimension of Newton's second law (from Table 4.1), and addresses the conception of the inverse proportion between mass and acceleration of objects at constant forces, i.e. when the mass of an object increases the acceleration decreases proportionally. All of the

distractors reflect the conception that when mass increases the acceleration decreases. Table 5.3(g) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(g) Performance and Confidence levels of all students for item 12

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	0.0	6.5	67.7	19.4	6.5	1.9	2.2	1.8
UPadp	68	1.5	1.5	61.2	32.8	3.0	1.6	1.7	1.6
UPsc	483	0.6	2.5	45.1	49.5	2.3	1.8	1.9	1.7
UPmaj	33	3.0	0.0	57.6	36.4	3.0	2.2	2.3	2.2
CTadp	143	0.7	1.4	71.1	21.1	5.6	1.9	1.8	1.9
ULfy	102	0.0	1.0	85.3	8.8	4.9	1.8	2.3	1.8
ULsc	79	8.9	7.6	57.0	24.1	2.5	2.0	2.1	2.0
ULmaj	43	0.0	0.0	86.0	7.0	7.0	1.8	2.3	1.8

Less than 50% of the students in each of the eight groups have chosen the correct option, option D. If a car is pulling a second car twice its mass, its new acceleration will be one-third (⅓) of the initial acceleration, because the combined mass of the two cars is three times that of the first car. A higher percentage (ranging from 45.1% to 86.0%) of students in each of the eight groups have chosen option C, that if the car is now pulling a second car twice its mass then the new acceleration will be half (½) of the initial

acceleration, clearly forgetting that the mass of the car pulling the second car has to be taken into consideration. Options A, B and E are very weak distractors for all cohorts, which can be interpreted to mean that the students understood the question well, but that flawed reasoning resulted in poor performance. The item, however, can be viewed as being difficult, because even the mainstream cohorts performed poorly.

Students from all eight groups were confident about their chosen options; this is shown by their average confidence levels which range from 1.6 and 2.2. Table 5.3(g) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from all eight groups were confident about their correct option; this is indicated by the confidence levels, in the above table, ranging from 1.7 to 2.3. Students, from all eight groups, choosing the incorrect options are confident about their choices (confidence levels of 1.6 to 2.2). However, a small difference between the average confidence levels for correct answers and the average confidence levels for incorrect answers can be observed in the majority of cohorts, except for UPteach, ULfy and ULmaj. The small percentage of correct answers obtained for these three groups should be noted.

The poor accuracy of judgment shown by all groups, but especially by the benchmark group, UPmaj, is a reason for concern. It may indicate that this item does not assess depth of conceptual understanding. Respondents may have thoroughly understood the concepts involved, but made a simple error due to inaccurate analysis of the problem situation as described.

5.5.8. Item 13

This item deals with an elevator that is being lifted up an elevator shaft by means of a cable. The students were to compare the magnitudes of the forces acting on the elevator, while moving up at constant velocity. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1), and addresses the conception of "canceling forces". The alternative conceptions documented in the literature are: distractor A that "largest force determines motion", distractor D which is "only active agents exert forces", and distractor E "air pressure assisted gravity" (Gunstone, 1981; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1992). Table 5.3(h) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

56.5% and 63.6% of the students from UPsc and UPmaj, respectively, have chosen the correct option, option B, while from 27.5% to 45.1% of the students from UPteach, UPadp, CTadp, ULfy, ULsc and ULmaj have also chosen the correct option, that the force by the cable on the elevator is equal in magnitude to the downward force by gravity on the elevator. A high percentage (ranging from 33.3% to 52.0%) of students from each of the eight groups have chosen A, the strongest distractor, that the upward force by the cable on the elevator is greater than the downward force by gravity on the elevator. This is a common alternative conception, as noted from above. Significantly, options C, D and E are weak distractors, except for ULsc and ULmaj, where distractors C and E feature slightly more prominently.

Table 5.3(h) Performance and Confidence levels of all students for item 13

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	45.2	38.7	3.2	3.2	9.7	2.3	2.4	2.3
UPadp	68	50.0	41.2	2.9	1.5	4.4	2.0	2.1	2.0
UPsc	483	38.9	56.5	2.3	2.1	0.2	2.3	2.5	2.1
UPmaj	33	33.3	63.6	0.0	0.0	3.0	2.4	2.6	2.1
CTadp	143	42.3	45.1	7.0	2.1	3.5	2.2	2.3	2.1
ULfy	102	52.0	27.5	9.8	5.9	4.9	2.0	1.8	2.1
ULsc	79	34.2	35.4	13.9	5.1	11.4	2.1	2.2	2.0
ULmaj	43	34.9	34.9	11.6	2.3	16.3	2.1	2.5	1.9

The average confidence levels in each of the eight groups is high (from 2.0 to 2.4), indicating that the students were confident about their chosen options. Table 5.3(h) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students choosing the correct option were confident about their correct option, their confidence levels ranging from 1.8 to 2.6. Students, from all the eight groups, choosing the incorrect options were also quite confident about their choice. This is indicated by confidence levels ranging from 1.9 to 2.3. There is a large difference between the average confidence levels for the correct answer and the average confidence levels for incorrect answers for students in UPsc, UPmaj and ULmaj. Interesting to note is the fact that the accuracy of judgment of ULmaj, a poorly

performing cohort, was on par with those of UPsc and UPmaj, the two best performing cohorts in this item. The accuracy of judgment from ULfy, the lowest performing cohort, was observed to be very poor, as the students choosing the incorrect options displayed higher confidence levels than those choosing the correct option.

5.5.9. Item 14

This item deals with a large man and a boy pulling as hard as possible on two ropes attached to a crate. The students were to identify the path that will be followed by the crate as the two people pull it along. The item is found in the conceptual dimension of the superposition principle (Table 4.1), and addresses the conception about the vector sum of forces. The alternative conceptions documented in the physics education literature are: distractor A which is that "the largest force determines the motion", distractors C and D that "force compromise determines motion" (Clement, 1982; Halloun & Hestenes, 1985b; Hestenes *et al.*, 1992). Table 5.3(i) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(i) Performance and Confidence levels of all students for item 14

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	9.7	22.6	64.5	3.2	0.0	2.1	2.4	2.0
UPadp	68	2.9	39.7	55.9	1.5	0.0	2.2	2.1	2.1
UPsc	483	1.0	62.1	35.8	0.8	0.2	2.3	2.3	2.2
UPmaj	33	3.0	57.6	39.4	0.0	0.0	2.3	2.3	2.2
CTadp	143	9.2	39.4	49.3	0.7	1.4	1.9	2.1	1.8
ULfy	102	11.8	35.3	50.0	2.0	1.0	2.1	2.4	2.0
ULsc	79	15.2	30.4	48.1	3.8	2.5	2.0	2.0	2.0
ULmaj	43	18.6	44.2	32.6	4.7	0.0	2.1	2.0	2.1

62.1% and 57.6% of students from UPsc and UPmaj, respectively, have chosen the correct option, while less than 50% (from 22.6% to 44.2%) of the students from the other six groups have also chosen the correct option, which is option B, that if both the man and the boy are pulling the crate, then the crate will follow a path closer to the man's pulling path than that of the boy. Option C, the strongest distractor, has attracted some attention from all the cohorts, with the percentage frequency ranging from 32.6% to 64.5%. This option says if both the man and the boy are pulling the crate, the crate will follow a path midway between the man's and the boy's pulling paths. Distractor A displays the thinking that "the winner takes all" and is more prevalent in the UL cohorts, than in the rest of the cohorts. Options D and E are weak distractors.

The table above indicates that all the students were confident about their chosen options (the average confidence levels ranging from 1.9 to 2.3). Table 5.3(i) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students in each of the eight groups were very confident about their correct option (their confidence levels ranging from 2.0 to 2.4). Students from all eight groups were also confident about their chosen incorrect options, their confidence levels range from 1.8 to 2.2. Noticeable is that there is very little difference between the average confidence levels for correct answers and average confidence levels for incorrect answers. Students therefore showed poor accuracy of judgment about their answers to this item in almost all cohorts, which seems to point to the presence of firm alternative conceptions amongst all cohorts.

5.5.10. Item 15

This item deals with two blocks moving to the right, and the positions of the two blocks represented by numbered squares at successive 0.20-second time intervals. The students were to indicate whether the two blocks ever had the same speed. If they choose a "yes" they had to indicate the instant at which the two blocks had the same speed. The item is found in the conceptual dimension of kinematics (Table 4.1), and addresses the conception about the differentiation between velocity and position. Distractors B, C and D address an alternative conception which is documented in the literature that "position and velocity are undiscriminated" (Clement, 1982; Halloun & Hestenes, 1985b). Table 5.3(j) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(j) Performance and Confidence levels of all students for item 15

Student Group	Number of Students	Options (%)						Average Confidence Level		
		A	B	C	D	(E)	B+C+D	All	Correct Option	Incorrect Options
UPteach	31	25.8	12.9	9.7	41.9	9.7	64.5	2.0	2.3	1.9
UPadp	68	40.3	4.5	16.4	28.4	10.4	49.3	2.2	2.6	2.1
UPsc	483	25.4	3.7	8.5	25.2	37.2	37.8	2.0	2.3	1.9
UPmaj	33	21.2	3.0	3.0	15.2	57.6	21.2	2.1	2.5	1.5
CTadp	143	44.8	5.6	9.8	21.7	18.2	37.1	1.9	1.8	1.9
ULfy	102	36.0	2.0	10.0	35.0	17.0	47.0	2.0	2.2	1.9
ULsc	79	38.0	7.6	13.9	26.6	13.9	48.1	2.2	1.9	2.2
ULmaj	43	46.5	4.7	4.7	32.6	11.6	42.0	2.1	2.6	2.1

57.6% of the students from UPmaj have chosen the correct option, and less than 50% of the students from each of remaining seven groups have chosen the correct option, which is option E, that the two blocks will have the same speed, at some time during the interval between 3 and 4. Options A and D are the most important distractors, with A featuring more prominently for almost all the cohorts. Distractor A states that the two blocks will never have the same speed. Distractors B, C and D state that the blocks are at the same position at instances 2 and/or 5. The students therefore make the mistake that the same position represents the same speed, a common alternative conception of being

unable to discriminate between speed and position (Clement, 1982; Halloun & Hestenes, 1985b; Hestenes *et al.*, 1992). The prevalence of this alternative conception is reflected by the sum of the responses B+C+D in Table 5.3(j).

The students were confident about their chosen options. The average confidence levels of students from all the eight groups are high (ranging from 1.9 to 2.2. Table 5.3(j) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Confidence levels ranging from 1.8 to 2.6 for the correct option indicate that the students from all eight groups were confident about their correct options. Students from all eight groups were also confident about their incorrect options, their confidence levels range from 1.5 to 2.2. This item displays a reasonable difference between the average confidences associated with correct answers as compared to the average confidence associated with the incorrect answers. This is evident from all the cohorts except for CTadp and ULsc cohorts, where the average confidence levels of the incorrect answers was higher than the average confidence levels for the correct answers. This is an excellent conceptual question. Despite its difficulty students who understood the concept and the graphical representation, analysed and answered it with confidence.

5.5.11. Item 16

The item deals with the positions of blocks "a" and "b", at successive time intervals. The time intervals were represented by means of numbered squares. The students were to interpret the visual representation in order to compare the acceleration of the two blocks. The item is found in the conceptual dimension of kinematics (Table 4.1),

and addresses the conception that acceleration is discriminated from velocity. Distractors A, B and C reflect the alternative conception that "velocity and acceleration are indiscriminate" (Clement, 1982; Hestenes *et al.*, 1992). Students were expected to realize that the numbered squares for blocks a and b are an equal distance apart, which is an indication of constant velocity. Table 5.3(k) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(k) Performance and Confidence levels of all students for item 16

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	35.5	0.0	19.4	29.0	16.1	2.2	2.3	2.1
UPadp	68	17.6	1.5	36.8	38.2	5.9	2.2	2.5	2.1
UPsc	483	10.0	2.5	25.2	56.1	6.2	2.4	2.8	2.0
UPmaj	33	12.1	0.0	21.2	66.7	0.0	2.2	2.7	1.2
CTadp	143	15.4	5.6	29.4	31.5	18.2	2.1	2.5	1.8
ULfy	102	24.0	7.0	43.0	15.0	11.0	1.8	2.1	1.8
ULsc	79	25.3	11.4	30.4	20.3	12.7	2.2	2.4	2.2
ULmaj	43	25.6	11.6	20.9	30.2	11.6	2.1	2.2	2.0

56.1% and 66.7% of the students from UPsc and UPmaj, respectively have chosen the correct option, while 15.0% to 38.2% of the students from UPteach, UPadp, CTadp, ULfy, ULsc and ULmaj have also chosen the correct option, option D, that the

acceleration of both blocks is equal to zero. Options A and C are the strongest distractors, with distractor C featuring more prominently in almost all the cohorts. Distractor C is about the acceleration of block "b" being greater than the acceleration of block "a", an answer that may be based on the larger distance between the squares for block b. Distractor A is about the acceleration of block "a" being greater than the acceleration of block "b". Distractors B and E were less prominent for all the cohorts as compared to distractors A and C. The graphical representation in this item is similar to the one used in item 15. Similar conceptual thinking and graphical interpretation are required. However, the performance is generally higher in this item than item 15.

The students were confident about their chosen options, with the average confidence levels of students from all the groups ranging from 1.8 to 2.4. Table 5.3(k) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students choosing the correct option, in all eight groups, were confident about their option, this was indicated by the confidence levels ranging from 2.1 to 2.8. These average confidence values are amongst the highest observed in this test. Similar values were only observed for items 6, 7 and 8. Students, from UPteach, UPadp, UPsc, CTadp, ULfy, ULsc and ULmaj, choosing incorrect options were confident about their choices. Their confidence levels range from 1.8 to 2.2. Students from groups UPmaj were not confident about their incorrect options, their average confidence level 1.2. The difference between the average confidence levels for correct answers as compared to the average confidence levels for incorrect answers is larger than in item 15, especially in the UPsc, UPmaj and CTadp cohorts. Accuracy of judgment for UPmaj, the best performing cohort, is exceptionally high, the highest

observed for any of the test items. According to the reasoning of Hasan *et al.* (1999) this could be interpreted to mean that the relatively poor performance in almost all of the cohorts can be ascribed to a lack of knowledge rather than to the presence of strong alternative conceptions.

5.5.12. Item 17

This item deals with two pucks of different masses on a frictionless table. The two pucks were pushed simultaneously across the table by means of equal forces. The students were to choose the puck that would reach the finish line first. The item is found in the conceptual dimension of Newton's second law (Table 4.1), and addresses the conception of the inverse proportion between mass and acceleration of objects upon application of constant forces. The strongest distractor is option C which is based on the alternative conception that "same amount of forces implies equal acceleration" (Clement, 1982). Distractors B and D reflect alternative conceptions that are less common among students. Distractor B states that the heavier puck will reach the finish line first, while distractor D is implying that additional information is required to enable one to provide an answer. Table 5.3(l) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(l) Performance and Confidence levels of all students for item 17

Student Group	Number of Students	Options (%)					Average Confidence Level		
		(A)	B	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	67.7	9.7	22.6	0.0	0.0	2.2	2.4	1.7
UPadp	68	72.1	7.4	20.6	0.0	0.0	2.3	2.3	2.1
UPsc	483	68.4	10.8	17.0	3.7	0.0	2.2	2.3	1.9
UPmaj	33	87.9	3.0	6.1	3.0	0.0	2.3	2.4	2.0
CTadp	143	68.3	6.3	18.3	7.0	0.0	2.2	2.3	1.8
ULfy	102	71.6	5.9	19.6	2.9	0.0	2.4	2.5	2.1
ULsc	79	64.6	5.1	12.7	13.9	3.8	2.1	2.3	1.6
ULmaj	43	65.1	9.3	16.3	9.3	0.0	2.2	2.2	2.2

More than 50% of the students in each of the eight groups have chosen the correct option, which is option A, saying that the puck having a smaller mass will reach the finish line first. Option C has attracted some attention from students. The percentage of students choosing this option ranges from 12.7% to 22.6%. Option C says that the two pucks will reach the finish line at the same time, since they received the same amount of force. Students make the mistake of ignoring the effect of mass on the acceleration of an object, for constant force. Distractors B and D are weak. The students probably realized that option D is an unlikely answer.

Students in each of the eight groups were confident about their chosen options: the table above shows the average confidence levels ranging from 2.1 to 2.4. Table 5.3(l) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students who chose the correct option were confident about their option, their confidence levels range from 2.2 to 2.5. Students from groups all the eight groups, who chose the incorrect options, were also confident about their incorrect options, their confidence levels range from 1.6 to 2.2. The difference in the average confidence levels of the correct answers and the average confidence levels for incorrect answers is large in almost all the cohorts, except for ULmaj and UPadp.

5.5.13. Item 18

This item is about a large box being pushed across the floor at constant speed. The students were to compare the magnitude of the forces acting on the box. The item is found in the conceptual dimensions of Newton's first law and superposition principle (Table 4.1), and addresses conception of "canceling forces." Distractor A address the alternative conception "speed is proportional to the applied force", distractors B and D address the alternative conception "motion when force overcomes resistance" and distractor E addresses the alternative conception "resistance opposes force" (Clement, 1982; Hestenes *et al.*, 1992; Minstrell, 1982). Table 5.3(m) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(m) Performance and Confidence levels of all students for item 18

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	(C)	D	E	All	Correct Option	Incorrect Options
UPteach	31	16.1	9.7	22.6	41.9	9.7	1.9	1.8	1.9
UPadp	68	13.2	5.9	16.2	60.3	4.4	2.0	2.3	2.0
UPsc	483	5.4	6.0	45.9	39.7	2.9	2.1	2.4	1.8
UPmaj	33	3.0	12.1	54.5	27.3	3.0	2.2	2.4	2.0
CTadp	143	7.0	8.4	30.1	48.3	6.3	1.9	2.1	1.8
ULfy	102	23.8	17.8	14.9	29.7	13.9	1.8	1.8	1.9
ULsc	79	20.3	13.9	30.4	27.8	7.6	1.9	1.8	1.9
ULmaj	43	30.2	7.0	25.6	30.2	7.0	2.1	2.5	1.9

The relatively poor performance on this item indicates that students found this to be a difficult question. 54.5% of the students in UPmaj have chosen the correct option, while less than 50% (ranging from 14.9%to 45.9%) of the students in each of the other seven groups have also chosen the correct option, which is option C. Option C states that the amount of force applied to move the box at a constant speed is equal to the amount of the frictional forces that resist the box's motion. Higher percentages (from 27.3% to 60.3%) of students in each of the groups have chosen D. Options B and D reflect a common alternative conception that the amount of force applied to move the box at a constant speed must be more than the frictional forces that resist the box's motion. The students make the mistake that the applied force must overcome the weight or the

frictional force for the box to actually move. Option E is relatively a weak distractor for all the cohorts, and is about the alternative conception that "resistance opposes force". According to Hestness *et al.* (1992); Hestenes & Wells (1992) and Minstrell (1982) the students do not regard friction as a "real" force; they take it that friction is just there to resist motion. However, this alternative conception is rare within the respondents in this study. Distractor A assumes that the speed is dependent on the applied force. Students make the mistake that increasing the force applied on the box would result in the speed of the box also increasing.

Students from all eight groups were confident about their chosen options, their average confidence levels range from 1.8 to 2.2. Table 5.3(m) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from all the eight groups, who have chosen the correct option, were confident about their choice. Their confidence levels range from 1.8 to 2.5. Students, from all the eight groups, who have chosen the incorrect options, also have high confidence levels ranging from 1.8 to 2.0. There is significantly poor accuracy of judgment for the UPteach, ULfy and ULsc. The best judgment on this item is recorded for UPsc and ULmaj.

5.5.14. Item 19

This item deals with a diagram representing a multiflash of an object moving to the right along a horizontal surface. The students were to identify the graph that best represented the object's velocity as a function of time. The item is located in the conceptual dimension of kinematics (from Table 4.1). The interpretation of a diagram of multiflash,

and transforming the multiflash into a graphical representation are assessed in this item. Table 5.3(n) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(n) Performance and Confidence levels of all students for item 19

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	26.7	43.3	13.3	13.3	3.3	2.0	1.9	2.1
UPadp	68	36.8	55.9	4.4	2.9	0.0	2.0	2.0	1.9
UPsc	483	26.2	63.2	4.6	2.5	3.5	2.1	2.2	1.9
UPmaj	33	24.2	66.7	9.1	0.0	0.0	2.3	2.5	1.9
CTadp	143	30.3	38.7	12.0	9.9	9.2	1.7	2.0	1.5
ULfy	102	28.7	28.7	14.9	4.0	23.8	1.4	1.4	1.4
ULsc	79	20.3	44.3	10.1	13.9	11.4	1.9	2.3	1.6
ULmaj	43	18.6	30.2	25.6	9.3	16.3	2.0	2.1	1.9

55.9%, 63.2% and 66.7% of the students from UPadp, UPsc and UPmaj, respectively, have chosen the correct option, and less than 50% of the students from UPteach, CTadp, ULfy, ULsc and ULmaj have also chosen the correct option, option B, that the speed of the object increases constantly for a longer period of time, the speed remains constant, and then the speed decreases constantly for a short period of time. Option A has also received some attention, the percentage frequency of students choosing

this option ranges from 18.6% to 36.8% in the eight groups. Option A shows the speed of the object to increase constantly for a short period of time, the speed remains constant, and then the speed decreases constantly for a longer period of time. Options C, D and E were weak distractors for UPadp, UPsc and UPmaj. Options C to E show a distinct break in the velocity of the object. It suggests that the UL groups, CTadp and UPteach may have confused the concepts of velocity and acceleration.

The table above indicates the average confidence levels of students for all the eight groups. Students from all the groups, except for students from ULfy, are confident about their chosen options; this is shown by their average confidence levels ranging from 1.7 to 2.3. Students from ULfy were not confident about their chosen options; their average confidence level is 1.4. Table 5.3(n) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from the seven groups (UPteach, UPadp, UPsc, UPmaj, CTadp, ULsc and ULmaj) are confident about their correct option, their confidence levels range from 1.9 to 2.5, while students from ULfy are not confident about their correct options, their confidents level is 1.4. Students from the seven groups, except for ULfy, have shown confidence levels ranging from 1.5 to 2.1, which indicate that they are fairly confident about their incorrect options. Students from ULfy are not confident about their incorrect options, their confidence level is 1.4. There is a moderately strong difference between the accuracy of judgment between students choosing the correct answers and those choosing the incorrect answers, especially in the UPmaj and ULsc. However the accuracy of judgment for the UPteach cohort is poor, the students answering the item incorrectly are more confident than those providing the correct answer.

5.5.15. Item 20

This item deals with the diagram representing a multiflash of an object moving to the right along a horizontal surface that was shown in item 19. Unfortunately the description of the problem did not contain any direct reference to item 19. The students were to identify the graph that best represented the object's acceleration as a function of time. The item is located in the conceptual dimension of kinematics (from Table 4.1). The graphical representation in this item is similar to the one used in item 19. Similar conceptual thinking and graphical interpretation is required, except that students were required to interpret the multiflash diagram in terms of acceleration rather than velocity. Table 5.3(o) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

More than 50% of the students from UPteach, UPadp, UPsc and UPmaj and less than 50% of the students from CTadp, ULfy, ULsc and ULmaj have chosen the correct option, option D. Option D is about the object accelerating constantly for a longer period of time, it maintains constant speed, and then decelerates constantly for a short period of time. Option E has also attracted some attention from students. The option is about the object accelerating constantly for a short period of time, maintaining a constant speed, and then accelerating for a longer period of time.

Table 5.3(o) Performance and Confidence levels of all students for item 20

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	3.3	10.0	3.3	60.0	23.3	1.3	1.2	1.5
UPadp	68	10.3	2.9	10.3	51.5	25.0	1.9	2.2	1.5
UPsc	483	5.8	3.3	7.9	62.0	21.0	2.2	2.3	2.0
UPmaj	33	6.1	3.0	6.1	69.7	15.2	2.2	2.3	1.9
CTadp	143	5.0	6.4	15.0	48.6	25.0	1.5	1.7	1.3
ULfy	102	8.8	17.6	23.5	35.3	14.7	1.1	1.3	0.9
ULsc	79	13.9	20.3	8.9	43.0	13.9	1.7	2.0	1.5
ULmaj	43	9.3	14.0	16.3	48.8	11.6	1.8	2.0	1.6

Students seem to be less confident about their answers to this item than about answers to other items. The average confidence levels of 1.1 and 1.3, respectively, for students from UPteach and ULfy, indicate that the students are not confident about their chosen options, while students from UPadp, UPsc, UPmaj, CTadp, ULsc and ULmaj are more confident about their chosen options, the average confidence levels range from 1.5 to 2.2. Table 5.3(o) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from UPadp, UPsc, UPmaj, CTadp, ULsc and ULmaj are confident about their correct option, their confidence levels range from 1.7 to 2.3, while students from UPteach and ULfy are not confident about their correct options, their confidence levels are 1.2 and 1.3, respectively.

Students from UPteach, UPadp, UPsc, ULsc and ULmaj have average confidence levels ranging from 1.5 to 2.0, this indicates that they are confident about their incorrect options, while students from CTadp and ULfy, have confidence levels of 1.3 and 0.9, respectively, this indicates that they are not confident about their incorrect options. A moderately large difference between average confidence associated with the correct answer and the average confidence associated with the incorrect answers was observed in almost all cohorts, except for UPteach whose accuracy of judgment was poor.

5.5.16. Item 21

This item deals with ticker tape trace which represents the motion of a car, moving to the right. The students were to study the tape and indicate the direction of the acceleration as well as the direction of the net force on the car. The item is located in the conceptual dimension of Newton's second law (Table 4.1), and addresses the conception that the directions of the acceleration and the net force are the same. Distractors B and C include the alternative conception that motion is in the direction of force (Clement, 1982; Halloun & Hestenes, 1985a). Table 5.3(p) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

The frequency distribution for this item could not be interpreted because of flaws in the problem presentation. It was not mentioned in the statement, for this item, where the first dot was made, a fact that was not picked up during the analysis of the pilot study. If one studies the tape carefully, there are two possible correct options. It is possible that the motion of the car could have been accelerated, and it is also possible that the motion could have been decelerated. In the case of the car accelerating, the acceleration and the

net force could both be directed to the right, and hence option A could be the correct one. For a decelerated motion, the acceleration and the net force could both be directed to the left, and hence option D could be the correct option.

Table 5.3(p) Performance and Confidence levels of all students for item 21

Student Group	Number of Students	Options (%)					Average Confidence Level		
		(A)	B	C	(D)	E	All	Correct Option	Incorrect Options
UPteach	31	22.6	29.0	29.0	6.5	12.9	1.7	3.0	1.6
UPadp	68	23.5	26.5	29.4	14.7	5.9	1.8	2.1	1.7
UPsc	483	23.3	11.0	41.4	20.8	3.5	2.0	2.3	1.9
UPmaj	33	27.3	3.0	30.3	36.4	3.0	2.4	2.4	2.4
CTadp	143	34.0	19.1	22.7	13.5	10.6	1.7	2.0	1.7
ULfy	102	30.4	42.2	15.7	4.9	6.9	1.8	2.6	1.7
ULsc	79	39.2	19.0	17.7	10.1	13.9	2.0	2.1	2.0
ULmaj	43	39.5	34.9	14.0	4.7	7.0	1.8	3.0	1.8

5.5.17. Item 22

This item deals with a person pulling a block across a rough horizontal surface. The person pulls the block at constant speed by applying a force F. The directions of the forces acting on the block were indicated on the diagram. The students were to choose the correct relationship of the magnitudes of the various forces acting on the block. The item is located in the conceptual dimension of superposition principle (Table 4.1), and

addresses the conception "canceling forces" (Clement, 1982; Halloun & Hestenes, 1985a; Maloney, 1984; Minstrell, 1982). Table 5.3(q) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(q) Performance and Confidence levels of all students for item 22

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	(C)	D	E	All	Correct Option	Incorrect Options
UPteach	31	12.9	3.2	9.7	61.3	12.9	2.1	1.7	2.1
UPadp	68	5.9	0.0	8.8	80.9	4.4	2.3	2.3	2.3
UPsc	483	16.7	1.5	11.9	65.1	4.8	2.3	2.5	2.3
UPmaj	33	15.2	3.0	6.1	75.8	0.0	2.6	3.0	2.6
CTadp	143	11.9	0.7	7.0	79.0	1.4	2.2	1.5	2.2
ULfy	102	4.9	5.9	13.7	63.7	11.8	1.9	2.1	1.9
ULsc	79	21.5	10.1	13.9	44.3	10.1	2.0	2.2	2.0
ULmaj	43	14.0	9.3	16.3	53.5	7.0	1.9	2.0	1.9

The performance of students in all groups on this item was exceptionally poor. Less than 17% of the students in each of the eight groups have chosen the correct option, which is option C. Option C is about the forces acting on the crate that is being pulled at constant speed, the normal force is less than the weight, and the applied force is greater than the frictional force. A high percentage (ranging from 44.3% to 80.9%) of students from each of the eight groups have chosen option D. Option D is about the normal force

being equal in magnitude to the weight and the pulling force being greater than the frictional force.

The students from all eight groups were confident about their chosen options, their average confidence levels range from 1.9 to 2.6. Table 5.3(q) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from all the eight groups were confident about their correct options, their confidence levels range from 1.5 to 2.5. Students, from all eight groups, choosing the incorrect options were also confident about their choices. Their confidence levels range from 1.9 to 2.6. Accuracy of judgment for the UPteach and CTadp is observed to be very poor as compared to the other cohorts. Their average confidence levels were higher for incorrect answers than for correct answers. This item seems to be associated with strong alternative conceptions as judged by the generally poor accuracy of judgment.

5.5.18. Item 23

This item deals with a rocket drifting sideways in outer space, with no outside force subjected to it. At position "b" the engine of the rocket produces a constant thrust at right angles to its original direction, and the engine is later switched off on reaching a certain point "c". The students were to select the path that best represented the movement of the rocket between "b" and "c". The item is found in the conceptual dimensions of kinematics and Newton's second law (Table 4.1), and addresses the conception "constant acceleration entails parabolic trajectory". Distractor A includes the alternative conception "loss of original force", distractor B includes the alternative conception "last force to act

determines motion", distractor C includes the alternative conception "force compromise determines motion" and distractor D includes the alternative conception "delayed impetus build-up" (Clement, 1982; Hestenes *et al.*, 1992; Jimoyiannis *et al.*, 2001;). Table 5.3(r) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(r) Performance and Confidence levels of all students for item 23

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	B	C	D	(E)	All	Correct Option	Incorrect Options
UPteach	31	9.7	41.9	32.3	3.2	12.9	1.3	1.3	1.3
UPadp	68	11.8	30.9	25.0	14.7	17.6	1.8	1.9	1.7
UPsc	483	10.4	23.6	26.5	14.6	24.8	1.7	1.8	1.6
UPmaj	33	3.0	21.2	27.3	18.2	30.3	1.8	2.2	1.7
CTadp	143	14.8	35.9	20.4	14.1	14.8	1.5	1.6	1.4
ULfy	102	10.8	41.2	21.6	7.8	18.6	1.6	1.2	1.6
ULsc	79	13.9	22.8	35.4	12.7	15.2	1.6	1.9	1.6
ULmaj	43	23.3	27.9	23.3	14.0	11.6	1.3	0.8	1.4

This appears to have been a difficult item, even the best performing cohort, UPmaj, performed poorly here. 30% or less of the students from each of the eight groups has chosen the correct option, which is option E. The two most prominent distractors are options B and C, with the percentage of students from the eight groups, ranging from

21.2% to 41.9% for option B and from 20.4% to 35.4% for option C. It is possible that the word "thrust" in the problem statement may not have been understood correctly. The relatively low confidence levels associated with most responses to this item is either an indication of a poor conceptual understanding or of a lack of clarity of the problem statement.

Students from UPadp, UPsc, UPmaj, CTadp, ULfy and ULsc, were marginally confident about their chosen options, their average confidence levels range from 1.5 to 1.8. The average confidence level, of 1.3, of the students from UPteach and ULmaj indicates that the students are not confident about their chosen options. Students from UPadp, UPsc, UPmaj, CTadp and ULsc, who have chosen the correct option, are confident about their chosen option. Their average confidence levels range from 1.6 to 2.2. The students who have chosen the correct options, from UPteach, ULfy and ULmaj are not confident about their choices. Their confidence levels range from 0.8 to 1.3. The average confidence of ULmaj is so low that it may indicate a lucky guess for most of the 11.6% correct responses. Students choosing the incorrect options, from UPadp, UPsc, UPmaj, ULfy and ULsc, were marginally confident about their incorrect choices. Their average confidence levels range from 1.6 to 1.7. Students from UPteach, CTadp and ULmaj were not confident about their incorrect choices. Their confidence levels range from 1.3 to 1.4. The largest difference between average confidence associated with the correct answers and the average confidence associated with the incorrect answers is observed for UPmaj. For the cohorts ULfy and ULmaj the accuracy of judgment is poor because average confidence associated with the correct answers is lower than average confidence associated with the incorrect answers.

5.5.19. Item 24

This item is about a heavy ball, attached to a string, and swung in a circular path in a horizontal plane. At a certain point the string breaks at the ball. The students were to predict the path the ball would follow as viewed from directly above the plane. The item is found in the conceptual dimension of Newton's first law (Table 4.1). Distractors A and D include the alternative conception "circular impetus", distractor C and E include the alternative conception "centrifugal force" (Clement, 1982; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1992). The correct answer is option B which shows the ball to be flying along the tangent of the circle at the time of separation. Table 5.3(s) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(s) Performance and Confidence levels of all students for item 24

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	25.8	35.5	16.1	6.5	16.1	1.6	1.8	1.5
UPadp	68	35.3	30.9	7.4	5.9	20.6	1.9	2.0	1.8
UPsc	483	28.5	46.3	7.7	8.8	8.8	1.8	1.9	1.7
UPmaj	33	30.3	42.4	12.1	6.1	9.1	1.8	2.1	1.6
CTadp	143	44.1	30.1	7.0	4.2	14.7	1.6	1.7	1.6
ULfy	102	34.3	28.4	11.8	5.9	19.6	1.7	1.8	1.6
ULsc	79	29.1	35.4	12.7	7.6	15.2	1.8	1.9	1.7
ULmaj	43	27.9	27.9	20.9	9.3	14.0	1.8	1.7	1.9

Less than 50% of the students from each of the eight groups have chosen the correct option, which is option B. Option A was a good distracter, because it has attracted some attention from students, the percentage ranges from 25.8% to 44.1%. Options C and E (centrifugal force) accounted for most of the remaining answers in all of the cohorts.

Students from each of the eight groups have shown that they were reasonably confident about their chosen options, their average confidence levels range from 1.6 to 1.9. Table 5.3(s) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from each of the eight groups have shown that they were confident about their correct option; their confidence levels are between 1.7 and 2.1. Students who have chosen the incorrect options in each of the eight groups have shown that they were also confident about their incorrect options. Their confidence levels range from 1.5 to 1.9.

5.5.20. Item 25

This item deals with the movement of the ball on a semicircular channel fixed to a table top. The ball entered and left the channel at the indicated points. The students were to select the path that would most nearly correspond to the path of the ball as it exited the channel and rolled across the tabletop. The item is found in the conceptual dimension of Newton's first law (Table 4.1), and addresses the conception of "with no force" (Clement, 1982; Halloun & Hestenes, 1985a; Hestenes *et al.*, 1992). Distractor A includes the alternative conception "circular impetus", distractors C and E includes the alternative conception "centrifugal force" and distractor D includes the alternative

conception "force compromise determines motion" (Clement, 1982; Halloun & Hestenes, 1985a). Table 5.3(t) below indicates the percentage of students, in each of the eight groups, choosing an option and their average confidence levels.

Table 5.3(t) Performance and Confidence levels of all students for item 25

Student Group	Number of Students	Options (%)					Average Confidence Level		
		A	(B)	C	D	E	All	Correct Option	Incorrect Options
UPteach	31	35.5	51.6	9.7	0.0	3.2	1.6	1.7	1.5
UPadp	68	33.8	57.4	4.4	2.9	1.5	1.8	1.8	1.7
UPsc	483	25.6	66.0	7.5	0.6	0.2	1.8	1.9	1.6
UPmaj	33	36.4	57.6	3.0	3.0	0.0	2.1	2.4	1.8
CTadp	143	30.3	53.5	14.8	0.0	1.4	1.7	1.8	1.6
ULfy	102	17.6	49.0	26.5	3.9	2.9	1.8	2.0	1.6
ULsc	79	12.7	29.1	27.8	10.1	20.3	1.6	1.5	1.6
ULmaj	43	20.9	34.9	27.9	11.6	4.7	1.6	1.7	1.6

More than 50% of the students from UPteach, UPadp, UPsc, UPmaj and CTadp, and less than 50% of the students from ULfy, ULsc and ULmaj, have chosen the correct option, option B. Option A has received some attention from students, the percentage of students choosing this option ranges from 12.7% to 36.4%. The conceptual content of this item is essentially the same as that of item 24, but students performed up to 23% better

than in item 24 as judged by the percentage correct answers per group. The only exception is ULsc.

Students from the eight groups have shown that they were confident about their chosen options, their average confidence levels range from 1.6 to 2.1. Table 5.3(t) also indicates the confidence levels of students who have chosen the correct option and those choosing the incorrect options. Students from all eight groups have shown that they were confident about their correct option. Their confidence levels range from 1.5 to 2.4. Students, from the eight groups, who have chosen the incorrect options, have indicated that they were also confident about their incorrect options, and the confidence levels of the students in all the eight groups range from 1.5 to 1.8.

5.6. Summary

The difference in the confidence levels associated with correct and incorrect answers for each item will be analyzed in this section. The difference between the average confidence levels associated with a correct answer and with the combination of incorrect answers is indicative of the quality of judgment about the correctness of the answer provided that cohorts are capable of. The summary of the analysis of the multiple-choice component of the results is based on the performance by the UPmaj cohort. The UPmaj cohort was chosen as a benchmark, because this cohort was the best performing group of all. Table 5.4 below indicates the ability of students to make accurate judgment in terms of the difference in the confidence levels. Item difficulty in Table 5.4 represents the percentage of UPmaj students who have chosen the correct answer.

Table 5.4 Summary of differences between average confidence levels for correct and incorrect responses, for all cohorts and item difficulty for the UPmaj cohort.

Item	Conceptual Dimension	Item Difficulty	Difference in the average confidence levels associated with correct and incorrect answers							
			UPteach	UPadp	UPsc	UPmaj	CTadp	ULfy	ULsc	ULmaj
6	Gravitation	Easy (78.8%)	0.1	0.2	0.5	0.7	0.4	0.3	0.1	-0.1
7	Newton's third law	Moderate (57.6%)	0.0	0.0	0.2	0.5	0.1	0.1	0.1	0.2
8	Gravitation	Difficult (36.4%)	0.1	0.1	0.3	0.5	0.3	0.1	0.0	0.1
9	Kinematics and Gravitation	Moderate (57.6%)	0.0	0.4	0.4	0.4	0.2	0.5	0.4	-0.1
10	Superposition principle and Newton's first law	Moderate (63.6%)	0.4	0.5	0.6	0.4	0.1	0.3	0.1	0.3
11	Newton's first and third laws and superposition principle	Easy (75.5%)	0.4	0.6	0.9	0.7	0.6	0.4	0.0	-0.1
12	Newton's second law	Difficult (36.4%)	0.4	0.1	0.2	0.1	-0.1	0.5	0.1	0.5
13	Newton's first law and superposition principle	Moderate (63.6%)	0.1	0.1	0.4	0.5	0.2	-0.3	0.2	0.6
14	Superposition principle	Moderate (57.6%)	0.4	0.0	0.1	0.1	0.3	0.4	0.0	-0.1
15	Kinematics	Moderate (57.6%)	0.4	0.5	0.4	1.0	-0.1	0.3	-0.3	0.5
16	Kinematics	Moderate (66.7%)	0.2	0.4	0.8	1.5	0.7	0.3	0.2	0.2
17	Newton's second law	Easy (87.9%)	0.7	0.2	0.4	0.4	0.5	0.4	0.7	0.0
18	Newton's first law and superposition principle	Moderate (54.5%)	-0.1	0.3	0.6	0.4	0.3	-0.1	-0.1	0.6
19	Kinematics	Moderate (63.7%)	-0.2	0.1	0.3	0.6	0.5	0.0	0.7	0.2
20	Kinematics and Newton's first law	Moderate (69.7%)	-0.3	0.7	0.3	0.4	0.4	0.4	0.5	0.4

Item	Conceptual Dimension	Item Difficulty	Difference in the average confidence levels associated with correct and incorrect answers							
			UPteach	UPadp	UPsc	UPmaj	CTadp	ULfy	ULsc	ULmaj
21	Newton's second law	Moderate (63.7%)	1.4	0.4	0.4	0.0	0.3	0.9	0.1	1.2
22	Superposition principle	Difficult (6.1%)	0.4	0.0	0.2	0.4	-0.7	0.2	0.2	0.1
23	Kinematics and Newton's second law	Difficult (30.3%)	0.0	0.2	0.2	0.5	0.2	-0.4	0.3	-0.6
24	Newton's first law	Moderate (42.2%)	0.3	0.2	0.2	0.5	0.1	0.2	0.2	-0.2
25	Newton's first law	Moderate (57.6%)	0.2	0.1	0.3	0.6	0.2	0.4	-0.1	0.1

The performance of the UPmaj cohort will be used in determining the difficulty of an item. The UPmaj cohort is used in this analysis because this cohort has shown to be the best overall performance amoung the eight cohorts chosen in the study. For the purpose of the study, an item is regarded as being easy if 70% and above of the students were able to answer it correctly; an item is regarded as being moderate if between 40% and 69% of the students were able to answer it correctly; and an item is said to be difficult if between 0% and 39% of the students answered it correctly.

The analysis of Table 5.4 above indicates that there are three groups of items. The first group is made up of items that are classified as being easy, and those are items 6, 11 and 17. The second group consists of items that have been classified as being moderately difficult, and these are items 7, 9, 10, 13, 14, 15, 16, 18, 19, 20, 21, 24 and 25. The third group is made up of items that are regarded as being difficult, and these are items 8, 12, 22 and 23.

The difference in the average confidence levels associated with correct and incorrect answers, for each cohort, is calculated by subtracting the average confidence

levels associated with incorrect answers from the average confidence levels associated with correct answers. A positive difference between the average confidence levels of students choosing the correct answer and those choosing the incorrect answers can be interpreted to mean that the majority of students in that cohort were making a correct judgment about their knowledge. However the degree to which this is true would depend on how big the difference between average confidence levels is. In Table 5.4 above, negative differences can be observed. This indicates the poor accuracy of judgment among the students, meaning that those that are getting answers wrong are more confident about their incorrect answers than those who answered correctly.

From Table 5.4, two sets of unique items can be identified. Items 12, 14 and 21 for UPmaj are flagged by their small values for differences in average confidence, despite the difference in their difficulty levels. The other uniqueness is observed in items 15 and 16 which have exceptionally large differences between average confidences while having comparable difficulties. At this stage it is not clear why these sets of items are unique, however an attempt will be made in order to find out what this means in the next chapter where an analysis of the written responses will be made.