TITLE PAGE

ASSESSMENT OF SENIOR SECONDARY SCHOOL STUDENTS' CONCEPTUAL UNDERSTANDING OF FORCE AND MOTION.

BY

UGWUANYI, CHRISTIAN SUNDAY
PG/M.Ed/09/51641.

DEPARTMENT OF SCIENCE EDUCATION
UNIVERSITY OF NIGERIA, NSUKKA

AUGUST, 2012

APPROVAL PAGE

This project has been approved for the department of Science Education, University of Nigeria, Nsukka.

Prof. B.G Nworgu
Supervisor

Dr. C.R Nwagbo
Head of Department

Prof. Joshua, Monday
External Examiner

Dr. B.C Madu
Internal Examiner

Prof. S.A Ezeudu
Dean, Faculty of Education

DEDICATION

This work is dedicated to God Almighty for his blessings and guidance throughout my Masters programme.

ACKNOWLEDGEMENT

I thank almighty God for his immeasurable blessings upon me at the course of my Masters programme. My profound gratitude goes to my project supervisor and Daddy as well Prof. B.G. Nworgu whose corrections and constructive criticisms of this work provided the basis for the success of this project. In a special way I thank my amiable Head of department and my Mummy, Dr. C.R Nwagbo for her supports during my programme. I equally remain ever grateful to all the Staff of the Department of Science Education, University of Nigeria for their individual and collective supports to me at the course of my programme.

Also, in this litany of appreciation, I wish to express my profound gratitude to my parents Mr. and Mrs. Ugwuanyi, Nicholas for without them I should not have known the only University of Nigeria. My elder Brother, Mr. Ugwuanyi, Kenneth and my Sister, Mrs. Ojobo Juliana are highly appreciated for their supports. My siblings Juliet, Sarah, Samson, Ogochukwu, Chinedu, and Chiemezie are highly commended for their patience and encouragement during my programme. All my fellow Postgraduate Students of 2009/2010 session in the department of science education are highly appreciated. In a special way I thank Mr. Orji Emmanuel who stood by me throughout the articulation of my project work. To the rest of my friends not mentioned, I thank you and love you all.

TABLE OF CONTENTS

CHAPTER FOUR: RESULTS

CHAPTER FIVE: DISCUSSION, CONCLUSION AND SUMMARY

LIST OF TABLES

ABSTRACT

This project work was based on the assessment of senior secondary school students' conceptual understanding of force and motion. Four researcher questions and two null hypotheses guided the study. Two research designs were adopted for the study namely: descriptive survey and Ex-post facto designs. Two hundred and twenty two senior secondary two (SS2) physics students in seven intact classes from six senior secondary schools that were purposively sampled from the area of study constituted the sample for the study. Those seven intact classes were drawn through simple random sampling within the SS2 physics classes in each of the six schools. The instrument for data collection was Force Motion Concept Evaluation (FMCE) developed by Thornton and Sokoloff, (1998). This instrument was adapted by the present researcher. Data collected was analyzed using WinBUGS computer program, frequency, percentage, mean, standard deviation and t-test of independent samples. While WinBUGS computer program, frequency, percentage, mean and standard deviation were used to answer the research questions, t-test of independent samples was used to test the null hypotheses. The results of the study showed that: (1) most of the Students responded to force and motion concepts using the impetus model (model 2) which is not exactly the correct model. (2) The nature of items determines the students' choice of models. (3) Gender has a significant influence on the students' conceptual understanding of force and motion in the direction of the female students. (4) School location has no significant influence on the students' conceptual understanding of force and motion. The implications of the above findings as well as recommendations were highlighted.

CHAPTER ONE
INTRODUCTION

Background of the Study

Physics is a natural science that involves the study of matter and its motion, along with related concepts such as energy and force. In this context, motion is the change in the position of an object as a result of an applied force (Anyakoha, 2007). Thus, force is an agent that brings about motion. It was Aristotle who first developed a systematic set of ideas about the physical world, which is often referred to as Aristotelian physics. Related to his concept of force is his classification of motion as natural, voluntary, and forced (Jammer, 1957). In the natural motion category, Aristotle believed that objects intrinsically either have a natural tendency to fall down to the earth, which he called gravity, or a natural tendency to rise into the sky, which he called levity. He thought that heavy bodies fall faster because the falling speed is in proportion to the weight of the objects. The earth and the sky are natural places objects would move to according to their internal natural tendencies. Voluntary motion refers to motion of living organisms such as animals and humans, who are agents able to exert force to make other inanimate things move. Nonliving objects are obstacles that stop or guide motion, but they do not exert forces.

In the forced motion category, an object moves because of the moving force applied to it by an agent. The object continues to move after the agent is no longer in contact with it because force is still transmitted to the object through a medium such as air. Motion in a vacuum is thus not possible. A force does not move an object unless it overcomes the object's inertia, an intrinsic resistance of the object. A constant force applied to an object produces a constant speed which is also inversely proportional to the inertia of the object. In the absence of force, an object would stop immediately. So, forced motion is made consistent to the other two types of motion by Aristotle through his notion of antiperistasis and his theory of entelechy or agency (Leclerc, 1972). In summary, for Aristotle, no motion is possible without force acting on the moving object, or in other words, motion and force are inseparable and a moving object is always an effect of some kind of entelechy, either visible or invisible.

According to Eryılmaz (2004), physics is the most basic of all sciences. It is at the root of every field of sciences and underlines all natural phenomena. In other words physics is the mother of all sciences. It is not a set of facts and rules to be memorized. Instead, memorization is a fruitless way to try to learn physics (Bueche, 1988). This observation by Bueche is in line with the assertion by Eryilmaz, (2004; 2) that;

> Physics is a difficult course to construct meaningful learning and so the achievement of students in physics is very low. There are some factors affecting students' achievement in physics. Some of them are related with students; like students' preconceptions, mathematics achievement, cognitive development level, attitudes towards physics, prior experience with the related fields, socio-economic level, age, and gender.

Among all of these factors, students' preconceptions play important role. Students' preconceptions in this context are all the ideas or knowledge they have about a given concept before the actual learning takes place. Researchers have been trying to diagnose students' preconceptions about physical concepts and rules. For instance, Clement (1982) and Eryılmaz (2004) in their separate studies on the influence of students' preconceptions on their achievement indicated that preconceptions held by students about physics concepts are significant. The above finding goes to support the fact that students' preconceptions contribute to their poor achievement in physics.

It is a popular idea in education that students come into a classroom with preconceptions about the material they are taught that can alter or interfere with their understanding of a concept (Tara, 2009). For example, when learning to read graphs of motion for the first time, students often interpret the graph literally, as they would interpret a picture or a map. As students learn, they move from naïve or intuitive state of understanding or conception to more acceptable conceptions.

Conceptual understanding according to Johnson (2005) refers to a person's representation of the major concepts in a system. Conceptual understanding is rich in relationships and understanding. It is a connected web of knowledge, a network in which the linking relationships are as prominent as the discrete bits of information. Conceptual understanding according to Johnson (2005) cannot be learned by rote. It must be learned by thoughtful, reflective learning. On the contrary, procedural understanding is the understanding of formal language or symbolic representations. It is the understanding of rules, algorithms, and procedures. Conceptual understanding is also known as the kind of knowledge that may be transferred between situations. The students' ability to develop conceptual understanding involves seeing the connections between concepts and procedures, and being able to apply physics principles in a variety of contexts.This is different from routine knowledge, which is knowledge that is applicable only to certain situations. For example, a student who decided to cram an aspect of a course for examination will quickly forget the crammed concepts after the examination. Conceptual understanding in physics develops when students "see the connections among concepts and procedures and can give arguments to explain why some facts are consequences of others" (National Research Council, 2001; 119). This implies that facts are no longer isolated but become organized in coherent structures based on relationships, generalizations and patterns.

Rittle-Johnson, Siegler, & Alibali, (2001) found that developing students' procedural knowledge had positive effects on their conceptual understanding, and conceptual understanding was a prerequisite for the students' ability to generate and select appropriate procedures. Thus, conceptual understanding is intertwined with procedural knowledge. This makes the isolated study of either difficult, requiring more than the determination of the correctness/incorrectness of a student's answer. It requires further investigation into the response, which can provide valuable insight into the thinking (Gould, 2005).The relationship between conceptual understanding and procedural understanding from their respective definitions above is that procedural understanding increases conceptual understanding. The difference between the two forms of understanding also is that while conceptual understanding leads to full adoption and transfer of the instructed procedure, procedural understanding leads to adoption but only limited transfer of the instructed procedure. This highlights the causal relations between conceptual and procedural understanding and suggests that conceptual understanding may have a greater influence on procedural understanding than the reverse.

According to Black and William (1998), it is increasingly appreciated that learning is tied to effective assessment by monitoring students' progress and feeding that information back to students. Assessment is an ongoing process of setting high expectations for student learning, measuring progress toward established learning outcomes, and providing feedback to improve academic progress (Black and William, 1998). There are many aspects of learning that can be assessed. However, if we seek to empower students to transfer the knowledge gained to new situations, then a deep understanding must be developed (National Research Council, 1999). In many physical science courses of which physics is part, deep understanding is usually associated with understanding of concepts. Lack of understanding of concepts by students may lead to their poor achievement. Thus, some efforts need to be devoted to identifying core concepts and then to devising means of gauging students understanding of those concepts. Thornton and Sokoloff (1998) designed Force and Motion Conceptual Evaluation (FMCE) to probe student's conceptual understanding of Newtonian Mechanics. Force and Motion Conceptual Evaluation was administered to more than 1000 students in the non-calculus and calculus-based general physics lecture courses and in the introductory physics laboratory at the University of Oregon and Tufts University. Results demonstrated that students high achievement in FCME indicate a high conceptual understanding of force and motion.

McDermott (1984), on reviewing research on conceptual understanding in mechanics (e.g., gravitational force, velocity and acceleration, and force and motion), pointed out some interesting and unexpected results from several studies. Studies about “passive” forces (e.g., the tension in a string) indicated that students, regardless of ages, have the same conceptual difficulty understanding those forces, yet most physics instructors proceed as if the concept of a passive force is easily understood (Minstrell, 1982). A study in velocity and acceleration revealed that students with greater facility with mathematics do not necessarily have a deeper conceptual understanding than those who have less training in mathematics (Whitaker, 1983). This shows that one’s knowledge of mathematics does not guarantee his/her conceptual understanding of physics concepts. It demands that one should posses an in-depth knowledge of the concepts not necessarily the mathematical knowledge only. Studies such as force and motion relation showed that many students have a well-integrated system of beliefs about the behavior of objects in motion. The students believe that an object will not stop moving unless the initial force acting on the subject is “used up”, the “Aristotelian” or “medieval” belief. This believe contradicts the Newtonian view that a body will continue in its state of uniform motion except intercepted by an external force (McCloskey, Caramazza, & Green, 1980).

Hake (2002) pointed out that most of the analyses of physics assessment tests have been done within the framework of "classical test theory" in which only the number of correct answers is considered in the scoring. Although CTT has served the measurement community for most of this century, IRT has witnessed an exponential growth in recent decades. The major advantages of CTT are its relatively weak theoretical assumptions, which make CTT easy to apply in many testing situations. At the item level, the CTT model is relatively simple. Classical test theory does not invoke a complex theoretical model to relate an examinee's ability to success on a particular item. Instead, CTT collectively considers a pool of examinees and empirically examines their success rate on an item (assuming it is dichotomously scored). The item response theory (IRT) was introduced to take care of the limitations of CTT. Item response theory (IRT), on the other hand, is more theory grounded and models the probabilistic distribution of examinees' success at the item level. As its name indicates, IRT primarily focuses on the item-level information in contrast to the CTT's primary focus on test-level information. For test items that are dichotomously scored, there are three IRT models, known as three-, two-, and one-parameter IRT models. In item response theory, assessments are made in an attempt to determine a student's place on the underlying ability scales. Andersen/Rasch (AR) multivariate IRT model is different from the ordinary IRT model in that it can deal with mixtures of strategies within individuals (Bao & Redish, 2004). In this model, students are not judged interms of getting correct options to questions posed to them but using any of the following models; Newtonian, Impetus or Aristotelian models. In this case, Newtonian model is the correct option; Impetus is an option very close to the correct option while Aristotelian model is a null option or totally incorrect option. This study adopted these models in order to characterize the students' conceptual understanding of force and motion interms of the three models discussed above.

Eryilmaz (2004) observed that gender contributes to poor achievement of students in physics. Gender according to Yang (2010) refers to the social attributes and opportunities associated with being male and female and the relationships between women and men; girls and boys, as well as the relations between women and those between men. These attributes, opportunities and relationships are socially constructed and are learned through socialization processes. Current studies show that female enrolment in physics and science subjects in general is very poor. This is in line with the study by Gonzuk and Chargok (2001) which revealed that the number of females who study physics in secondary and tertiary institutions is small compared to the number of boys. This difference in the number of females and males in the study of physics has created gender disparity in the academic achievement of students in physics and science subjects as a whole.

Gender difference was first investigated by sociologist of education. The focus was largely on female under achievement at every level of the educational system. Therefore there is need to promote the teaching and learning of physics in schools especially among female student. Ajejalami (1990) identified the following factors as contributing to under representation of females in science and technology Education in Africa;

- Lack of functional guidance and counseling services
- Relationship of sex to occupational prestige
- Influence of schooling

- Family background
- Interest among other factors
- Lack of adequate orientation programme
- Societal discrimination against females in education
- Occupational choice and adaptation of science and technology.

Fakorede (1999), in his own contribution posited that poor enrolment of girls in science subjects is due to:

- Inadequate opportunity for girls to study science,
- Inadequate achievement of girls in science,
- Inadequate interest of girls in science,
- Unfavourable attitude of girls to science learning and
- Inadequate knowledge of girls on the true nature of science.

The critical belief of biological theorists is that gender differences are natural and therefore unalterable. It would be right and proper to treat boys and girls in schools differently because their natural inclinations are different roles. Thus, theories were advanced that females excelled in language based subject because of their greater and reasoning abilities yet under performed in sciences because of their lower level of innate ability of shape and form factors.

Where a school is situated says a lot about the achievement of students (Ma & Wilkins, 2002). According to Ezeudu (2003), school location means urban-rural setting. The urban-rural influence is also expected in physics just like any other science subject because of the psychosocial influence it may have on the teachers and students resulting mainly from school location. This may even dictate their academic achievement in science of which physics forms a part. Therefore, the area in which a school is located can affect the educational achievement of a student. A school in the heart of the government reserved area (G.R.A) or housing estate cannot be compared with a school located in an unsuitable place like motor garage, main street, noisy environment, and nearness to a big market among others. Noisy environment is capable of hampering teaching and learning conditions. Long journey to school can be drudgery. These deplorable states inform the present journey.

From the foregoing, it can be seen that assessment of students' conceptual understanding of force and motion has not been extensively studied in order to draw clear picture of the influence of gender and school location on the achievement of students in physics. The above reason necessitated the current study.

Statement of Problem

The academic achievement of students in secondary schools has been a subject of concern to many people including parents, administrators, educators, psychologists and researchers. The poor achievement of students in science especially physics has continued to be a major concern to all and particularly those in the main stream of science education. This has also resulted into tension, depression, and social maladjustment among some secondary school students who were not able to attain the desired grade required for admission into higher institutions. Acquisition of physics knowledge is not a one way task but can effectively be achieved through gradual developmental stages. Thus students learn physics through three stages or modes. These stages are the naïve stage or the no knowledge stage from where they develop to common knowledge or what is called the impetus stage and finally the real conceptual understanding state. Assessments of students' achievement in physics in Nigeria have been done based on the conventional achievement test which considers only the overall scores of students on the concepts in question. This type of testing does not consider the developmental stages the students pass through in conceptualizing a concept.

Besides, it has not been objectively established whether gender has influence on the students' conceptual understanding from the review of some foreign works on the students' conceptual understanding of force and motion. The little that has been known, though not quite well is the influence of gender on the students' achievement in physics and science related subjects. In this respect, there seems to be controversy over the influence of gender vis-à-vis school location on students' achievement in science subjects particularly physics. These two variables need to be critically studied in order to objectively establish their influence on the students' achievement in physics. Hence, the problem of this study is: what is the assessment level of senior secondary school students' conceptual understanding of force and motion? What is the influence of gender on the students' conceptual understanding of force and motion? What is influence of location on students' conceptual understanding of force and motion?

Purpose of the Study

The purpose of the study was to assess senior secondary school students' conceptual understanding of force and motion. Specifically the study will determine;

1. The students' propensity to use Newtonian, Aristotelian or Impetus model in responding to force and motion concepts.
2. The propensity of the items to elicit a particular response model to force and motion concepts.
3. The influence of gender on students' conceptual understanding of force and motion.

4. The influence of location on the students' conceptual understanding of force and motion.

Significance of the Study

This study as a conceptual assessment work supports the constructivist or cognitive approach to learning, ie they are based on the belief that students construct their own understanding of concepts by expanding or modifying their existing views. The procedure also reinforces the value of cooperative learning and the individual students' active role in learning. Besides the outcome of this study has added to Arnderson/Rash multivariate theory. In that respect, the findings of the study revealed that the students are always in a mixed state of mind when presented with force and motion concepts. Those implied the theoretical significance of the study as it adds to the cognitive theory of learning.

The findings of this study will be of great benefit to the following; physics teachers, physics students, curriculum planners and physics stakeholders.

The findings will be useful to the physics teachers owing to the fact that physics knowledge is acquired by the students based on the three developmental models which represent the mixed state of students during teaching and learning. By being aware of this, the teacher will prepare his/her lessons in a way that will take care of the three developmental models for proper assessment of the students.

The physics students will benefit from the outcome of the current study. The findings of this study will reveal the fact that students are in mixed state during teaching and learning situation. That is some students responding to instructions in the right direction while others respond in the wrong direction following the three developmental models. Students being aware of this mixed state situation as well as the developmental models before an instruction will make frantic efforts to get the right instruction in order to develop the proper conceptual knowledge. Also they will understand that physics demands proper conceptualization rather than rote learning or cramming.

For the curriculum planners, the findings will bring to a limelight the developmental models which are the impetus or common knowledge model, the Aristotelian or null model and the real conceptual understanding. Those models will enable the planners to plan the curriculum in way that the pedagogical approach to the teaching of physics will incorporate the three developmental models the students pass through at the course of learning physics. This will serve as a guide to the teachers on how proper conceptual knowledge should be developed by the students.

Finally, the findings of this study will be of great importance to the physics stakeholders who are much concerned on the influence of gender and school location on the students' conceptual understanding of force and motion. This study will empirically through the use of conceptual assessment test study the influence of those variables.

Scope of the Study

This study was on the Assessment of Students' Conceptual Understanding of force and motion based on Conceptual Assessment Framework. The physics curriculum for senior secondary school class one (SS1) constitutes the learning content document for the current study. The aspects of Newtonian mechanics that are taught at senior secondary one include; kinematics, forces and motion, work, energy and momentum, circular motion and gravitation, and oscillations. In this study, the major focus will be on force and motion. The concept of force and motion will only be studied for the fact that all other topics are products of force and motion. This implies that understanding of the concepts of force and motion facilitates the understanding of the other topics. The study will make use of SS2 students who have SS1 as their penultimate year. This is to ensure that the students must have learnt the concepts of force and motion in their penultimate year.

Research Questions

The following research questions guided the study;

1. What is the students' propensity to use Newtonian, Aristotelian or Impetus model in responding to force and motion concepts?
2. What is the propensity of the items to elicit a particular response model to force and motion concepts?
3. What is the influence of gender on students' conceptual understanding of force and motion?
4. What is the influence of school location on students' conceptual understanding of force and motion?

Hypotheses

The following hypotheses were tested at 0.05 significance level.

HO_1: Gender is not a significant factor in the students' conceptual understanding of force and motion.

HO_2: Location is not a significant factor in the students' conceptual understanding of force and motion.

CHAPTER TWO
LITERATURE REVIEW

The review of literature was done with respect to the following sub themes;

CONCEPTUAL FRAMEWORK

- ❖ Concept of Newtonian Mechanics
- ❖ Principles of Cognitive Science of learning
- ❖ Markov Chain Monte Carlo (MCMC) Estimation
- ❖ Concept of Evidence Centered Design

THEORETICAL FRAMEWORK

- ❖ Cognitive Theory of learning
- ❖ The Andersen/Rasch (AR) Multivariate Theory.
- ❖ Comparisons among three Measurement Theory

RELATED EMPIRICAL STUDIES

- ❖ Students Learning in Physics
- ❖ Influence of Gender on Students' Achievement in Science
- ❖ Influence of Location on Students' Achievement in Sciences

Summary of Literature Review

Conceptual Framework
Concept of Newtonian Mechanics

Newtonian mechanics is the study of the causal relationship among force, mass, and motion. Natural philosopher and 17th century scientist Isaac Newton developed a set of universal principles, to help explain and predict the motion of objects in the natural world, and the degree to which these objects change their relative motion by interacting with external forces. According to Anyakoha (2007), Newton embodied the interrelationship between the physical concepts of force, mass, and acceleration into his three laws of motion. The ability of Newtonian mechanics to accurately describe natural phenomenon under observation is derived from the application of these distinct laws of motion. The first law of motion holds that a body in motion tends to remain in motion, and a body at rest tends to stay at rest. This principle explains the concept of inertia, namely, the application of force that is required to move a stationary object. Similarly, the deceleration of a body, otherwise moving at a constant speed, can only occur when an outside force acts on it. For example, a bullet fired from a rifle would continue its motion in a certain direction perpetually, were if not for the simultaneous forces of gravity and the resistance of the air in the atmosphere. These forces act together on the bullet to cause it to stop at a certain distance from the spot where it was initially fired.

Newton's second law of motion is a mathematical or quantitative formula that describes the inherent nature of force. Newton postulated that the amount of force exerted is directly proportional to the mass of a body, times its acceleration, ($F = ma$). If two distinct bodies are moving at a constant acceleration, the object with the larger mass will produce the greater force. This principle of Newtonian mechanics can be illustrated by the example of an automobile and a locomotive that are traveling towards each other at an equal rate of speed. When the two objects collide, the force exerted on the automobile will be vastly greater due to the much greater mass of the locomotive. The third law of motion can be summarized with the statement ‘for every action, there is an equal and opposite reaction’. In other words, the forces of two bodies acting upon each other are always equal and directly opposite. For instance, the force that a baseball exerts on a bat is equal and opposite to the force that the bat exerts on the baseball.

According to Serway and Beichner (2001), Newtonian mechanics provides analytical tools that an observer can use to accurately predict changes in a body’s motion resulting from external forces acting upon it. These principles are equally applicable to the movement of large celestial objects or the motion of a simple tennis ball. By using vectors to depict the magnitude and direction of a body in motion, as well as the manner in which external forces impact that body, an observer can accurately predict the net resultant velocity and direction of a body which is the sum of all the external forces acting on that body at any given point in time. This review is important for the fact the current study is tied to students’ conceptual understanding of the concept of force and motion.

Principles of Cognitive Theory of Learning

These principles are summarized below.

1. Learning takes place within communities of practice. Even when we are learning how to do things for ourselves or on our own, like learning how to develop a presentation or spreadsheet, we rely on a variety of tools and others to do our work. Learning is embedded within everyday work activities: we learn by observing others who are more skilled than we are or by

participating in more peripheral but nonetheless significant tasks as we are learning the different aspects of our work.

2. Novices learn to become experts through practice in solving a variety of problems in a domain. Developing expertise in one's job requires extensive experience with the kinds of problems one is likely to face on the job. Experienced telephone sales operators, for example, have made thousands of calls to potential customers and are aware of a variety of kinds of problems they are likely to encounter on the job. They are more likely than novice sales operators to close sales because they are able to use fewer outside resources to answer customer questions. Moreover, they ask potential customers a targeted set of questions focused on the most relevant cost-related items, unlike novices, who follow scripts provided and treat all kinds of questions as equal in importance.
3. Becoming an expert means applying learning to new contexts. Simulating situations commonly encountered on the job can provide learners with models for ways to solve new problems as they come up and better prepare learners than simply having them read in a book or training manual about problems they might face. On the other hand, learning that is too tightly coupled to a directly experienced situation is brittle. When confronted with a novel situation in which previously accumulated ways of approaching problems fail, we need something to fall back on. Because more general knowledge can be useful in coping with novelty, the most robust learning is likely to result from a combination of many experiences close to the target situation with some knowledge at a more general level.

4. Prior knowledge mediates learning. What we know from our prior experience can get in the way of our ability to solve new problems. For example, many of our everyday concepts of motion and physics get in the way of our understanding of the scientific concept of force. On the job, school based learning can interfere with the development of cognitively flexible solutions to complex problems, unless that learning can be reorganized to help solve the problems at hand.
5. Learning is enhanced when thinking is made visible by collaboration and reflection among learners. Learners can improve their comprehension and mastery of a particular domain through active monitoring of their own learning. Active monitoring is best achieved when learners have the opportunity to share their ideas with others and reflect on their practice.

These principles of cognitive science point to an approach to learning that is quite different from the approach of many workplace training departments. Learning, according to cognitive science, should no longer be viewed as a process of simply transmitting information from a teacher to a learner or from an expert to a novice. Rather, learning should be viewed as an active, constructive process, involving collaboration and reflection among people who learn through the course of their everyday activity.

Markov Chain Monte Carlo (MCMC) Estimation

Markov chain Monte Carlo (MCMC) is a computation that is carried out through Bayesian Analysis Using Gibbs Sampling. WinBUGS implements various MCMC algorithms to generate simulated observations from the posterior distribution of the unknown quantities (parameters or nodes) in the statistical model. The idea is that with sufficiently many simulated observations, it is possible to get an accurate picture of the distribution. Both item and person parameters under the AR model will be estimated by MCMC sampling-based methods using a Bayesian approach. Compared with other estimation procedures (e.g., joint maximum likelihood estimation, JMLE, and marginal maximum likelihood estimation, MMLE, which has become standard IRT methodology in practice), one advantage of using a Bayesian approach is that the estimation is direct given that the priors are specified in advance; therefore, it requires a much smaller sample size than maximum likelihood estimation (MLE) procedures to yield stable estimates. Also, no artificial constraints need to be imposed on the parameter space as with MLE, since outward drifts of the estimates are naturally and effectively controlled by the priors. Swaminathan and Gifford's simulation study (1986) on the 3-PL model showed that the Bayesian estimates stay in the parameter space. Furthermore, the Bayesian estimates show a closer relationship to the true values than JML estimates. Although from a frequentist perspective, Bayesian estimates are biased toward the mean of the prior distribution (i.e., exhibits shrinkage), the use of priors is also what keeps the parameter estimates (especially the a and c parameters in the 3-PL model) in the admissible parameter space. Another advantage of using Bayesian estimation over MLE (especially MMLE) is that uncertainty in item parameter estimates are easily incorporated into examinee inferences, and vice versa (e.g., Kim, 2001; Patz & Junker, 1999a; Tsutakawa & Johnson, 1990).

Finally, Bayesian estimation provides solutions for those examinees with perfect or zero scores. JMLE fails unless those examinees are removed prior to estimation. MCMC techniques have been recently applied to estimate parameters for latent variable measurement models, especially IRT models. For example, Albert (1992) applied a Gibbs sampling method to estimate item parameters under the two-parameter normal ogive model for a 33-item Mathematics Placement Test administered to 100 examinees. He then compared the estimates with those derived from MLE using EM algorithm based on a normal approximation. It was found that in terms of item difficulty parameters, these two estimation procedures yielded similar results; for discrimination parameters, the estimates based on Gibbs sampler tended to be larger than those resulting from MLE/EM, indicating that the marginal posterior distributions exhibited right skewness. By examining the standard error of estimates, he further suggested that the normal approximation to the posterior of the item parameters based on the mode and information matrix (used to compute the EM standard errors) might be a poor approximation to the exact posterior distribution.

Kim (2001) examined the accuracy of parameter estimates in the one-parameter logistic model using MCMC with Gibbs sampling. Four datasets were analyzed using Gibbs sampling method along with MLE methods, including conditional maximum likelihood estimation (CMLE), JMLE, and MMLE (expected a posterior method, EAP, was used to estimate θ parameters). He found that item parameter estimates from the four methods were almost identical, and θ estimates from Gibbs sampling were similar to those obtained from EAP.

Patz and Junker (1999a) applied MCMC using a Metropolis-Hastings sampling algorithm to estimate parameters for the two-parameter logistic IRT model. Later they extended this strategy to the data with multiple item formats (multiple-choice and partial credit items), missing data, and rated responses (Patz & Junker, 1999b). They demonstrated how MCMC approach is more straightforward and relatively easier to implement than MML/EM as IRT model complexity increases, since computationally MCMC does not involve exact numerical quadrature (for the E step) or pre-calculation of derivatives (for the M step). However, the cost of this ease of implementation is that the execution time is generally slower than EM due to the fact that MCMC is trying to estimate the entire joint posterior distribution function of all the parameters while EM only estimates one or two values for each parameter – the MML estimate and its standard error.

The examples reviewed above show some advantages of using MCMC to estimate model parameters, including its flexibility, ease of implementing for complex IRT models, and accuracy of parameter estimates. The major drawback is that it usually requires a considerable amount of computing time. Perhaps the rapid development of personal computers (PCs) will alleviate this problem. Relatively few applications (e.g., Hoijtink & Molenaar, 1997) using MCMC techniques for LC models were found in the literature.

The key idea of Markov chain simulation is to create a chain whose stationary distribution is a specified posterior distribution and run the simulation long enough (i.e., repeating the sampling process by starting with a possible value for each variable then drawing a sample from the updated distribution and continuing to do so) that the distribution of the current draws conditioning on the previous draws is a sufficiently close approximation to the stationary distribution. At this point, approximate distributions and summary statistics for each variable can be obtained based on these many draws.

There are two widely used Markov chain simulation methods – the Gibbs sampling and the Metropolis algorithm or the Metropolis-Hastings approximation method. We start with the Gibbs sampling, since to date most statistical applications of MCMC in psychometrics have used it. There are some important properties for MCMC. First of all, MCMC exhibits the known Markov property of “no memory”, meaning that draws in cycle t+1 only depend directly on values in cycle t, not on previous cycles. Second, an indirect dependence on previous values introduces autocorrelations across cycles. That is, although the sequence of draws of a given parameters does approximate the posterior of that parameter, the values are not independent draws from the distribution. Smaller autocorrelation coefficients are preferred; the value depends on the parameterization of the model, and the amount of information in the data for a given parameter. Third, under regularity conditions (e.g., sampling can "cover the space", or can choose any point in each parameter's range), dependence on starting values is "forgotten" after a sufficiently long run. Therefore, the "burn in" cycles – the first few hundreds or thousands of draws that are to be discarded because the sampled values in those cycles are dependent upon the starting values – will not be included in calculating the summary statistics for the variable one is interested in. One can run multiple chains with over-dispersed starting points to examine if they look like they are sampled from the same stationary distribution (see Gelman et al., 1995, on convergence diagnostics).

Several computer programs have been developed for specific purposes to carry out the model parameter estimations using MCMC, including a computer program written in MATLAB (The Mathworks, Inc., 1996) by Albert (1992), a FORTRAN program written by Baker (1998), and a specialized code in S-PLUS (MathSoft, Inc., 1995) by Patz and Junker (1999). Each of these programs is specific to the model the researchers were studying. For the current study, the WinBUGS computer program is used for estimating model parameters, assessing model convergence, and comparing the nested models under the AR model (these are discussed in the next chapter). It is an interactive Windows version of the BUGS program (**B**ayesian inference analysis **U**sing **G**ibbs **S**ampling, Spiegelhalter, Thomas, & Gilks, 1997) for Bayesian analysis of complex statistical models using

MCMC techniques (in particular Gibbs sampling, with Metropolis steps within Gibbs for full conditionals with hard-to-calculate forms). It is a widely used freeware program and has been used to a wide range of complex problems due to its flexibility. This is particularly useful in the current study because no other computer programs are readily available for estimating parameters for the AR model. Therefore, because of the program's flexibility along with those advantages and nice features of using MCMC over other estimation procedures as discusses earlier, BUGS where MCMC techniques are implemented would be the best choice to estimate parameters for the AR model in the current study.

Concept of Evidence Centered Design

Evidence Centered Design (ECD) of assessment is a design of educational assessments in terms of evidentiary arguments. This type of assessment design addresses the following questions: What do you want to make inferences about (e.g., language proficiency), what do you need to see (e.g., performance with what qualities), and what features in tasks evoke the evidence you need (Messick, 1994). In this study, the target inference is the mixtures of conceptions (some correct, some erroneous) that students bring to bear on mechanics problems, with evidence coming from tasks designed to reveal those conceptions. Evidence-centered assessment design (ECD) (Mislevy, 2003) is a formal framework for designing assessments from this evidentiary-reasoning perspective. The models of ECD and their relationships according to Messick (1994; 17), help to answer the following questions:

- "What complex of knowledge, skills, or other attribute should be assessed?" The proficiency model (sometimes called student model or competency model), describes characteristics of examinees upon which the inferences are to be based.
- "What behaviors or performances should reveal those constructs, and what tasks or situations should elicit those

behaviors?" The task model describes features of situations that will be used to elicit performance and how to structure the situations to obtain the kinds of evidence needed for the evidence model.

- "The rational development of construct-based scoring criteria and rubrics." The evidence model expresses how what is observed in a given task constitutes evidence about student model variables. It includes scoring criteria and rubrics and the statistical model of how to update the belief of student knowledge given the performance.

Evidence-Centered Design for Learning (ECDL) augments the main ECD models with a pedagogical model that represents how to foster growth and learning, given existing evidences and student proficiency level. ECDL also provides a way to consider a wide range of quality characteristics, including learner engagement, learning effectiveness and efficiency, validity, and accessibility, in the context of a more comprehensive product quality argument. ECDL is intended to be useful not only for the design of new products but also the redesign of existing products. It is worth pointing out the importance of assessments in learning-oriented products, in which the assessment may be used to: (a) detect learning, (b) to guide learning (such as by helping identify the next appropriate learning activity or providing task-specific feedback) (Black & William, 1998). Thus, success in the domain of learning-oriented products relies heavily upon high-quality assessments. ECDL addresses the issue of assessment validity in learning-oriented products by establishing arguments for both the design phase (deductive reasoning) and the operational use phase (inductive reasoning). Reasoning deductively, we say, "If the student's proficiency value is high (or low) then high (or low) scores should be obtained on given items." On the other hand, during operational use of the assessment, we emphasize inductive reasoning (i.e., inference), "If high (or low) scores are obtained on these problems, then the student's proficiency value is high (or low)." While inductive reasoning is what we ultimately care about in an assessment, it is the deductive reasoning that takes precedence, both logically and chronologically. In the deductive reasoning we needed to play “what if” which involves considering what would happen if we knew that the student had a certain level of proficiency and imagining that student interacting with the tasks and then determining what scores would be obtained.

There are four stages in the ECD: Domain Analysis, Domain Modeling, Conceptual Assessment Framework, and Operational Assessment. Domain Analysis concerns gathering information about how people acquire knowledge or skills, and how they use them. This information is essential in assessment since it will help the assessment designer to know, for instances, under what situations we can see people doing the kinds of things and using the kinds of knowledge related to assessment. This analysis can provide clues about important features of performance situations. This information then is organized in terms of design objects called paradigms in the second stage of the ECD, Domain Modeling. There are three paradigms: the proficiency paradigms are the structures that organize potential claims about aspects of proficiency for students; the evidence paradigms state the kinds of things student might say or do that would provide evidence about these proficiencies; and the task paradigms are the kinds of situations that might evoke the evidence we need to see. At this stage, by knowing the interrelationship among these three paradigms, one starts to rough out the structure of an assessment that will be needed for future operational assessment.

The next stage of the design is Conceptual Assessment Framework. The Conceptual Assessment Framework specifies the more technical elements of an operational assessment, including measurement models. This research is mainly based on this stage, where three models under the AR model will be compared. CAF consists of three major models that coordinate an assessment's substantive, statistical, and operational aspects, and provides the technical details required for implementing the assessment. The student model (SM) specifies the variable(s) in terms of which we wish to characterize students. It may contain a single variable, representing an overall proficiency, or multiple variables, characterizing several aspects of knowledge or competences. Technically, the SM model can be presented by a possibly vector-valued parameter (usually denoted by θ), and a joint probability distribution p (θ). A student model then can be viewed as a mathematical structure containing variables that can take a range of possible values, and a joint probability distribution function quantifying relationships among these variables. Reasoning from observable behavior in task situations with given features, we can characterize the students' knowledge, skills, or proficiency which we are interested in making inferences about, and use probability distributions over a SM variables to express our belief about their values. Values of SM variables correspond to claims that can be made about students, for example, as to their level of proficiency for getting correct answers in a domain of tasks, as in traditional testing, or as to the way they may be thinking about problems in the domain.

A task model (TM) in the CAF concerns substantive considerations about the features of tasks that are necessary to evoke evidence about SM variables. It embodies beliefs about the nature and structure of task situations, as they are important under the conception of knowledge that guides the assessment's design. With regard to work products (e.g., what the student says, does, or produces), the task model also specifies what student behaviors or productions will be observed as they provide clues about their knowledge, again as they are important under the conception of knowledge that guides the assessment's design. Therefore, for a particular task, the values of task model variables consist of information characterizing the situation with regard to its salient features and the kinds of performances that will be captured. In addition, the TM also describes features of tasks that are needed to inform the operational activities for particular assessment tasks. Although many tasks can be created given a task model, the collection is constrained only to suit the needs of the assessment project.

An evidence model (EM) in the CAF concerns reasoning from what we observe in a given task situation to update beliefs about SM variables. It contains two components, which connect students' work products to their knowledge and skill: the evaluation component and the measurement component. One can think of the evaluation component as "task scoring" since it describes rules for extracting evidence from individual performances, as values of observable variables. In other words, the evaluation component indicates how one identifies and evaluates (e.g., through rubrics) students' work or performance (what they say, do, or produce in a given task), and expresses salient aspects of them as values of observable variables (e.g., item or task scores). In comparison, one can view the measurement component as "test scoring," for it contains statistical models used to synthesize information or analyze data from observable variables across performances, in order to reflect belief about SM variables. Technically speaking, the measurement component specifies models used to construct likelihood functions for SM variables (as induced by the values of the observed variables) and to estimate model parameters to obtain estimators for SM variables.

Therefore, one can see that the measurement models in this context make connections between student models and task models. If tasks are well developed jointly with measurement model, one can take advantage of efficient statistical computing to use the complex model or estimation procedures (e.g., full Bayesian analysis) to support inferences in terms of preplanned and substantively important patterns in data. Bayesian estimation procedures (Markov Chain Monte Carlo techniques, MCMC) are used to estimate models' parameters and address their comparative fit to the data. Finally, the last stage of ECD is the Operational Assessment. It concerns the operation of the implemented assessment based on the design generated in the previous stages (in particular, the CAF). The design reviewed above is the appropriate design for any study that is based on the students, conceptual understanding based on the conceptual assessment framework. Thus the reason for the review is that the present study especially the instrument for data collection is based on the conceptual assessment framework.

Theoretical Framework

Cognitive Theory of Learning

Cognitive learning theory assumes that learners are active in their attempts to understand the world, new understanding depends on prior learning, learners construct understanding, and learning is a change in people's mental structures instead of changes in observable behavior. A major theme in the theoretical framework of Bruner is that learning is an active process in which learners construct new ideas or concepts based upon their current/past knowledge. The learner selects and transforms information, constructs hypotheses, and makes decisions, relying on a cognitive structure to do so. Cognitive structure (i.e., schema, mental models) provides meaning and organization to experiences and allows the individual to "go beyond the information given". As far as instruction is concerned, the instructor should try and encourage students to discover principles by themselves. The instructor and student should engage in an active dialog (i.e., Socratic learning). The task of the instructor is to translate information to be learned into a format appropriate to the learner's current state of understanding. Curriculum should be organized in a spiral manner so that the student continually builds upon what they have already learned.

Bruner (1966) states that a theory of instruction should address four major aspects: (1) predisposition towards learning, (2) the ways in which a body of knowledge can be structured so that it can be most readily grasped by the learner, (3) the most effective sequences in which to present material, and (4) the nature and pacing of rewards and punishments. Good methods for structuring knowledge should result in simplifying, generating new propositions, and increasing the manipulation of information.

In his more recent work, Bruner (1996) has expanded his theoretical framework to encompass the social and cultural aspects of learning as well as the practice of law. Bruner's constructivist theory is a general framework for instruction based upon the study of cognition. Much of the theory is linked to child development research (especially Piaget). The ideas outlined in Bruner (1996) originated from a conference focused on science and math learning. Bruner illustrated his theory in the context of mathematics and social science programs for young children.

This latter aspect of acquiring knowledge becomes more important in current assessments. An important purpose that one would like to be able to address with assessment is not only to examine how much individuals know (e.g., how many items examinees answer correctly or incorrectly) but also to assess how, when, and whether they can apply what they know. To be able to do this, however, requires more complex tasks (than traditional tests that target only on students' overall performance) to reveal information about how individuals respond to the questions through their cognitive processes, including, for example, reasoning strategies and evolving understanding over time.

In the current study, we take the cognitive approach to study how students solve physics tasks. In particular, we are interested in examining how individuals organize information or knowledge that is processed. Knowing this will help us understand how people answer questions or solve problems.

The Andersen/Rasch (AR) Multivariate Theory

The AR psychometric model is consistent with the above-reviewed literature concerning students' cognitive process in physics learning for the measurement paradigm instantiated in assessments such as the FCI and FMCE. Based on the literature, a student's mental model state could consist of distinct competing physics models, and which one would be used for problem-solving depends on the features of the item presented to him/her. The AR model is an appropriate one from psychometrics since it was developed to model students in terms of propensities toward characteristic types of response which in this case will be conceptions of the domain (it can be viewed as a "mixture-within-persons" approach, as further explained below), rather than in terms of expected correctness. The analysis based on the AR model addresses the same kinds of student/task interactions as the Model Analysis methodology that Bao and Redish (2004) developed in their research to study students' physics learning.

The AR model states that at a given point in time, a person is seen as having propensities to answer in accordance with any of the conceptions. Tasks are also parameterized in terms of their tendency to provoke different conceptions as well. Given physics tasks (e.g., a multiple-choice test with j items, each item having m choices that are each associated with a particular way of thinking about situations in the targeted domain) administered to N examinees, the idea of AR model can be presented as follows. In terms of items, each item can be characterized by a vector value containing m elements, with each value of the element corresponding to a location on a continuum for a certain property in physics in a sense that can be described as follows. For example, the first choice may represent the Newtonian approach (those who pick this choice have behaved in a way consistent with a Newtonian strategy for problem-solving), the second choice may represent the strategy using “impetus theory”, the third choice may reflect an Aristotelian belief, and so forth. In other words, the item parameter is a vector-valued parameter which contains m elements, and larger values for an element indicate a greater tendency for that item to elicit responses in line with the corresponding problem-solving approach. In particular, the choice with the highest value indicates that that way of thinking is more common on this item, all other things being equal. In line with the science-learning research noted above, particular features of given items can tend to evoke particular misconceptions. For persons, similar to items, each examinee is characterized in terms of a vector-valued parameter that also contains m elements, with each element representing the associated propensity level on the continuum. If, for instance, person A has the greatest propensity level on the Newtonian approach (the first choice on the above example), this indicates that person A tends to respond in accordance with Newtonian strategies for problem-solving. Furthermore, if the task is designed to examine whether or not students have mastered Newton’s third law, then we can make an inference about person saying that he/she probably understands how to apply Newton’s third law in this situation.

Let X_{ij}; i = 1… n, j = 1… k be independent random variables (i is the index for examinees while j is the index for items) and further assume that there are m discrete choices for each item (so Xij can be any integer between 1 and m). The m options are associated with kinds of response that are the same across all items, with respect to strategy, perspective, style, conception, or some other way of partitioning responses in the domain. In our example above, category 1 responses are consistent with a Newtonian approach, category 2 with an "impetus theory" approach, and category 3 with a naïve or Aristotelian approach, which might be called a null approach. The formal AR model can be written as:

$$P(X_{IJ} = p) = \frac{\exp(\theta_{ip} + \beta_{jp})}{\sum_{p=1}^{m} \exp(\theta_{ip} + \beta_{jp})},$$

Where,

p is an integer between 1 and m;

ip θ is the pth element in the person i's vector-valued parameter; and

jp β is the pth element in the item j's vector-valued parameter.

Again, note that there are m probabilities for each examinee on a given item, representing the probability of choosing any particular choice for that person on that item.

Comparisons among Three Measurement Theories

IRT models are the most familiar item-level test theory models that concern students' propensities to do well on tasks in a defined domain; e.g., right answers rather than wrong answers, high ratings rather than low ratings on essays. The 3-PL model, applied to the Bao-Redish data, would therefore characterize students simply as to their propensity to make correct (Newtonian) answers to the items, or an overall proficiency level in students' learning. This is the kind of inference one can make about a student model through the 3-PL model. As discussed earlier, the 3PL would also characterize items as to their operational properties such as difficulty, discrimination, and/or guessing parameters. The 3-PL model would be appropriate for the Bao-Redish data sets which were collected from the tests consisting of multiple-choice items if the objective of the analyst were simply to characterize students in terms of their propensity to make correct responses.

However, an analysis based on the 3-PL model would not be able to probe the details of the processes students may have used for problem-solving. The 3PL embodies a certain way of thinking about students and items – which characteristics are important, and how those characteristics are, in probability, reflected in students' performances. It is a student's tendency to produce right rather than wrong answers. This is the major difference between the AR model and the 3-PL model.

Other IRT models have been developed to incorporate knowledge from cognitive science of learning into psychometric analyses (e.g., Embretson, 1995; Samejima, 1995). They are not conformable with current research in that the cognitive process hypothesized and parameterized in those models is not consistent with Bao and Redish's model for how students learn physics, and those models are not coded for the data sets used in the current study (i.e., response categories represent various strategies possibly used by students for problem-solving, and the responses are not necessarily involved in a systematic relationship to cognitive approaches).

The student model that accords with a latent class (LC) model presumes that each student is a member of a class, associated with distinct probabilities for responses of different types. Applied to the Bao-Redish assessment, one might posit that each student is associated with a certain conception (or misconception) of Newtonian mechanics problems, and tends to respond in that way. These might be called “pure states” of understanding and expected. It is possible that some polytomous IRT models (e.g., Masters’ partial credit model or Muraki’s generalized partial credit model – Muraki, 1992) could be used to obtain students’ overall proficiency levels. These models are not widely used as the 3-PL model in the current educational measurement. The more important point is that recoding students’ responses based on their level of partial knowledge as opposed to their tendencies to use different approaches would not be coherent with the task design in this study. These models are not considered here. There are occasional inconsistencies in responses, however, so that answers corresponding to other conceptions would occur with probabilities to be estimated. In this sense, an LC model could be posited which would be similar to the AR model in that it can be used to model students in terms of conceptions of the domain rather than in terms of expected correctness (as in the 3-PL model).

However, the LC differs in that a given student may not be in an arbitrary “mixture state” as in the AR model, but is instead modeled as in a “pure state” – that is, a student uses a consistent theory or model to respond, although there are some probabilities of responses of other strategies. Therefore, the latent class analysis (LCA) provides a different perspective in understanding students’ learning – one that is farther than the AR model and even the 3-PL model in terms of investigating individual student’s learning. It should be noted that some mixture LC/IRT models exist in the literature (e.g., Mislevy & Verhelst, 1990; Rost, 1991); however, they too are not considered here because those models concern propensity toward correct answers under different strategy uses. Neither the modeling of strategies nor the form of the data required for those analyses are consistent with the framework of the Boa-Redish analyses. The present aim is not to explore a given data set with a compendium of models, but rather to fit a model (the AR model here) most aligned with the intention of the assessment and to draw out the cognitive basis of that model.

The AR model matches up with the literature in cognitive science of physics learning in the area of focus, and is fully compatible with the Bao-Redish study. This is the main point in the current study i.e., to show an example of using a psychometric model that is consistent with the studies in science of learning, and to make inferences about students by examining their responses to tasks designed under the same conception of learning.

Review of Related Empirical Studies

Students Learning in Physics

Bao and Redish (2004) presented their view about students' learning based on the evidence centered design approach. A total of 468 second year college physics students at the University of Maryland formed the sample for their study. The data for the study were analyzed using mean, person correlation coefficient, polysearal correlation coefficient and winsBUGS computer programme. Two methods were developed by Bao and Redish (2004) in their studies on students' physics learning: Concentration Analysis and Model Analysis. They first developed Concentration Analysis to measure how students' responses on multiple-choice questions are distributed. This information can be used to explore if the students have common correct/incorrect models or if the question is effective in detecting models of students' reasoning for problem-solving.

To look for the detail of those possible situations of student models for reasoning in physics learning, Bao and Redish (2004) developed the second method, Model Analysis, to extract the probability states of students' use of different models. The analysis was mainly based on the cognitive science of learning (e.g., context dependence, as discussed earlier) and the knowledge from qualitative research (i.e., interview students to find out possible contextual features in learning the Newtonian mechanics). It involves two important concepts which have been mentioned but not well-defined in the preceding sections: common models and student model states. The common models are those models commonly used by students. They often consist of one correct expert model and a few incorrect or partially correct student models. When students are presented with a set of questions for a new physics concept or situation, they may respond using a previously well-formed model or create a new model based on their past experience or knowledge (e.g., mapping of a reasoning primitive). As an example, with the concept of the force-motion relation in the Newtonian mechanics, there are three common models students may use as indicated on the Bao and Redish's paper:

Model 1: An object can move with or without a net force in the direction of motion. (an expert model using Newtonian way of thinking)

Model 2: There is always a force in the direction of motion. (an incorrect student model using "impetus theory" belief)

Model 3: Null model. (an unsystematic, inconsistent, or Aristotelian approach)

Methods developed by Bao and Redish as described above provide a better way than the traditional, total score, evaluation method in revealing how students learn new physics concepts in Newtonian mechanics. It helps to design a more valid instrument and to improve instruction in physics. However, there are some limitations in their methods.

First, it is not clear whether these methods can be applied to other fields of learning (e.g., mathematics). The cognitive learning process for mathematics may be different from learning physics since they are two different subjects (i.e., their substances are not the same). More generalized analyses would be preferred in educational testing. For example, psychometric models which are statistical models are substance-independent, yet when appropriately constructed and applied; reflect the key patterns in the substantive problem at hand. Thus analyses based on them can be used in various subjects of learning while remaining true to the learning theory of the domain of interest.

Second, the Boa-Redish analyses are not connected with the well developed psychometric machinery, where much has been learned over the past century about issues such as estimation, model criticism, and modeling approaches. Some measurement models (e.g., IRT models) have been widely used in educational testing. Although they may not be sufficient for Bao and Redish's interests in knowing how students learn new physics concepts, the little-known AR model as described above does exist, and it is consonant with the conception of student model states in the Bao-Redish approach. Their analysis is based on the belief that naïve students use different models, with probabilities that depend in part on the features of the test items. This could be described as a mixture-within-persons model as the AR model.

Third, their analyses are data dependent, in the same sense as those of the classical test theory. That is, all of statistics in their analyses will vary if different sets of questions are given to students. This limits the test use, and it does not allow comparing students' performance if they have taken quite different subsets of test items. In modern test theory, like the AR and other IRT models, once the assumptions of the model are satisfied (and they can be examined using statistical procedures), the item (or person) parameter estimates are independent of the particular sample of students (or test items) (Hambleton & Swaminathan, 1985).

Appreciating the patterns they seek to model yet recognizing the limitations of the methods, psychometric analyses based on the AR model are conducted in the current study.

Reiner, Slotta, Chi and Resnick (2000), studied the students' conceptual understanding of force, light, heat and electricity using evidence centered design approach. The sample of their study was 379 first year physics students. In their reviews on misconceptions of the concept of force, light, heat, and electricity, Reiner, Slotta, Chi and Resnick (2000), provided evidence that naïve conceptions often reflect an underlying commitment to preexisting knowledge of material objects or substances. Using force as an example, in physics, force is seen "as a process of interaction involving two or more material objects in which these objects are speed up, slowed down, or redirected". However, physics novices do not conceive of force as a process of interaction between two material objects. Rather, they think of force as either some act of the object itself or a property of a material object (e.g., novices tend to explain gravity by assuming an innate, inexhaustible internal property called weight, or even explain the fact that an object will fall if dropped because of the contact of air pressure). According to the researchers, understanding how prior knowledge affects students' learning would help us to explain, for example, why students tend to use "impetus theory" or the Aristotelian approach for solving Newtonian mechanics problems even after instruction.

Chun-Wei (2003), used a formal psychometric model (i.e., the Andersen-Rasch multivariate measurement model, AR; Andersen, 1973 & 1995) to study students' conceptual understanding in physics. The perspective was based on the "evidence-centered" design (ECD; Mislevy, Steinberg, & Almond, 2003) framework. The study builds on the Force Concept Inventory (FCI; Hestenes, Wells, &

Swackhammer, 1992) and the Force Motion Concept Evaluation (FCME; Thornton & Sokoloff, 1998) task design and on previous analyses of the cognitive processes of physics problem-solving. It thus focuses on the measurement component of evidence model (EM) in the ECD stage called the Conceptual Assessment Framework (CAF). The researcher used one hundred and ninety seven (197) second year students of University of Maryland. The findings show that the use of the AR model for tasks designed to reveal students' conceptions in physics is consistent with a cognitive perspective of learning, namely that students solve problems using approaches that can often be identified with conceptions or common misconceptions, and their propensity to use a certain approach (in this case Newtonian, "impetus theory", or Aristotelian) for problem-solving depends on the features of the item presented to him/her. Despite that the study was carried out on foreign students, the researcher used second year University Students but not secondary school students which the current study is interested in.

Influence of Gender on Students' Achievement in Science

Ezanya (2004) using a total of 200 students drawn through stratified random sampling carried out a study to investigate any sex related difference in students' performance in chemistry achievement test. The study revealed that males' performance in CAT was higher than their female counterpart.

Lynch and Paterson (2002) investigated the gender difference in respect of students' recognition of science concepts. The sample used for the study was a total of 1,635 students. This study was done in Tasmania high schools and students were chosen on account of location, urban or rural environment and social economic status of their parents. The hypothesis was tested using chi square test. The study revealed that boys' recognition of science concepts was higher than the girls.

Shaibu and Mari (1997) did a survey of gender related difference in the understanding of science process skills amongst junior secondary school students in some Nigerian schools. The main aim of the study was to determine the level of understanding of science process skills amongst male and female junior secondary school students in the population covered by the study. A total of eleven schools were used from where three hundred and thirty students were drawn. Mean, standard deviation and t-test were used for data analysis. The findings of the study revealed that the performance of females in inventory of science process skills was higher than the males.

From the review of the related literature on the influence of gender on students' achievement in science, there seem to be disagreement between the researchers with respect to the influence of gender. This goes to show that there still exists a gap in the study of the influence of gender on the students' achievement in science. Thus the present study will empirically carry out such study.

Influence of School Location on Students' Achievement in Science

Okonkwo (2000) investigated the nature of relationship between some school variables and students' achievement in mathematics. In a specific term, the study investigated the nature of relationship between location and school type, sex, qualification and experience of teacher on the students' achievement in mathematics. The study used a sample one thousand one hundred and forty eight junior secondary one students of eighteen randomly selected schools in Delta state. The findings show a significant effect of school location on students' achievement in mathematics with the rural students performing higher than the urban students.

Jegede (1984) did a survey of the ways in which characteristics of the educational environment of the home, school and classroom affect students' achievement in physics. The sample for the study consists of senior secondary two students from nine schools in Imo state. The results show that though the mean achievement scores of the urban schools were higher than that of the rural schools, the mean difference was not significant. The result however further implied that school location and adequacy of the number of teachers when taken together influenced students' achievement in physics. The researcher concluded that school location on its own has no significant effects on the students' achievement in physics.

Ayogu and Nworgu (1999) did a survey of the influence of gender and school location on students' achievement in physics. In the study, physics students of SS1, SS2 and SS3 were randomly drawn from twelve secondary schools in Enugu State to represent the population for the study. Revised physics achievement test was used as the instrument for data collection. The result shows that the influence of school location was significant for SS1 and SS2. The researchers based on their findings concluded that the urban students outperformed their rural counterparts across all the levels and that there was location influence in students' achievement in the revised PAT.

Erubami (2003) carried out a study on school location as a correlate of students' academic achievement at the senior secondary certificate examination in Ogun state. The result shows that there is no significant difference between the performance of students in science in urban and rural schools.

Effiong (2002) carried out a study on the effect of school location on students' achievement in physics. Four hundred students of SS2 formed the sample for the study. The results of the study show that the urban students achieved significantly higher than the rural students.

Ezeife (1990) examined sex and environment as factors in physics achievement. The study focused on finding out whether there is any difference attributable to school location in the performance of Nigerian secondary schools student in physics. A sample of one hundred and twenty SS2 students from four schools in Akwa Ibom state was used for the study. Analysis of variance was used for data analysis. The results show that students in the urban schools performed better than those in the rural schools.

Unachukwu (1999) using a sample of six thousand nine hundred and twenty one primary school pupils in Enugu state investigated sex differentials and school location in mathematics achievement of primary school pupils in the state. The sample for the study was drawn randomly from class 2, 3 and 4 in forty primary schools in the state. The result revealed that pupils in urban primary schools achieved higher than their counterparts.

The review of the related literature on the influence of school location on the achievement of students in science revealed controversial views of the researcher on such study. Equally there is a need to study in the detail the influence of school location on students' achievement in science particularly physics.

Summary of Literature Review

As briefly discussed in the very beginning, this study is meant to provide an example of analyses integrating ideas from several areas of current research. First, it is based on the ECD framework developed. Within this framework, this study focuses on the measurement component of the evidence model in the CAF. Second, as the primary goal, it is desired to compare and explore the utility of analyses that draw upon the area of psychometrics to make inferences about the student variable(s) using data from physics as an example. From the review of literature, the analysis developed by Bao and Redish to study college students' learning in physics (especially in Newtonian mechanics) was shown to be effective but with the limitations discussed above.

Next, this line of research also integrates findings from the psychology of science learning with psychometric methods. One of major findings that Bao and Redish drew upon is that naïve students' responses on questions are affected by their pre-existing knowledge or experience (i.e., context dependence). Therefore, task questions designed to measure students' understanding of physics concepts are suggested to embody the contextual features associated with each item targeting on the specific concept.

The related literature reviewed in relation to the influence of gender and school location on the achievement of students in physics and science related subjects indicated that various researchers in their respective studies had diverse views in respect to the influence of such factors on students' achievement. Besides, the researchers used ordinary achievement test instruments for their studies. This type of test instrument does not consider the mixed state of students in response to physics concepts which the present study will study.

CHAPTER THREE
RESEARCH METHOD

This chapter was discussed in line with the following sub-headings.

- Research Design
- Area of Study
- Population of the Study
- Sample and Sampling Techniques
- Instrument for Data Collection
- Validation of Instrument
- Reliability of the Instrument
- Method of Data Collection
- Method of Data Analysis

Design of the Study

Two separate research designs were adopted for the study. These are survey and Ex-post facto research designs. According to Nworgu (2006), survey research design is a design in which a group of people or items is studied by collecting and analyzing data from only a few people or items considered to be representative of the entire group. Specifically, the study will adopt the descriptive survey which according to Nworgu (2006) involves collecting data on and describing in a systematic manner, the characteristics, features or facts about a given population. This particular design was adopted for the study because the researcher will be interested in collecting and describing some relevant data with respect to certain variables in relation to the population.

Besides, Ex-post facto design was adopted since the influence of two variables that cannot be manipulated were investigated. According to Nworgu

(2006), Ex-post facto design is a research design which is used to study to influence of variables that cannot be manipulated such as gender, location etc.

Area of the Study

This study was carried out in Nsukka Education Zone of Enugu State. Nsukka Education zone is made up of three local Governments viz: Nsukka local Government, Uzo-Uwani and Igbo Etiti. This education zone was chosen by the researcher for the fact that the schools within the zone have the necessary attributes such as urban location, rural location and mixed schools as well as single sex schools that are needed for the study. Besides, the schools within the zone are at the reach of the researcher.

Population of the Study

The population of the study was the entire senior secondary two (SS2) physics students in all the government owed senior secondary schools in Nsukka Education zone. There are forty eight (48) senior secondary schools in Nsukka education zone. The population of SS2 physics students in the entire senior secondary schools is two thousand four hundred (2,100). The distribution of the school population follows thus; Nsukka local government has twenty three (23) senior secondary schools with population of SS2 physics students of one thousand one hundred and fifty (1,150); Igbo Etiti has thirteen (13) senior secondary schools with population of SS2 physics students of six hundred and fifty (450); while Uzo Uwani has twelve (12) senior secondary schools with population of SS2 physics students of six hundred (500) (Post Primary School Management Board, 2011).

Sample and Sampling Technique

The sample size for the study was two hundred and twenty five (225) senior secondary two (SS2) physics students in seven (7) intact classes from six senior secondary schools sample from the area of study. The six schools were drawn from the forty eight (48) government owned senior secondary schools in the Nsukka Education zone through purposive sampling. Purposive sampling was used, since the influence of gender and school location will be studied.

Secondly, from the sampled schools, simple random sampling technique was used to generate the seven intact classes based on the number of science classes in each of the schools.

Instrument for Data Collection

Conceptual assessment test instrument was used for data collection. This instrument is Force Motion Concept Evaluation (FMCE) which is a thirteen (13) item questionnaire as can be seen in appendix B. The instrument was developed by Thornton and Sokoloff, (1998) and was adapted by the researcher. The adaption was in terms of wording of the items of the instrument. Synonymous terms familiar to students in Nigeria was used in some cases rather than the main terms.

The instrument was developed in the year 1998 for the fact that the ordinary achievement test instrument cannot be used to study students' conceptual understanding of force and motion. The instrument measures the students' ability to respond to questions on force and motion using either of the following models: Newtonian model, Impetus model or Aristotelian model. For instance, a question like, if a large box is being pushed across the floor at a constant speed of 4.0 m/s. What can you conclude about the forces acting on the box? The students might answer the question using any of the following options; A). The amount of force applied to move the box at a constant speed must be equal to the amount of the frictional forces that resist its motion. B) The amount of force applied to move the box at a constant speed must be more than the amount of the frictional force that resist its motion, C) The amount of force applied to move the box at a constant speed must be more than its weight. Option “A” represents the Newtonian model; option “B” represents the Impetus model while option “C” represents the Aristotelian model.

The response options of this instrument were framed systematically as to reflect the three models stated above. One or more of the options will be collapsed into any of the three models. Items number one (1) to number (5) have five response options while those from number six (6) to number (13) have seven response options. During data analysis, these options will be collapse into the three models used by the students in responding to the items. During the scoring process, any student that answered using the Newtonian model will get a score of one (1), Impetus model will be scored two (2) while Aristotelian model will be scored three (3). In this case one is greater than two while two is greater than three.

Validation of the Instrument

The instrument was validated by three persons from the department of Science Education, University of Nigeria, Nsukka. The validates were asked to validate the instrument based on the wording of the items of the instrument, appropriateness of the items with respect to the purpose of the study. Two of the Validates are purely measurement and evaluation specialists while the third Validate is from Science Education.

Reliability of the Instrument

The instrument was subjected to trial testing using twenty (20) SS2 physics students from a government secondary school in Ogrute, Enugu Ezike. An estimate of reliability coefficient known as Kuder-Richardson 20 (K-R20) was used to estimate the reliability of the instrument. The reliability coefficient for the instrument was found to be 0.87. This value implies that the instrument is reliable for the study.

Method of Data Collection

The data for this study were collected through the use of force motion concept evaluation which is an instrument that was developed based on the evidence centered design of assessment. The researcher visited the sampled schools to collect the data for the study. At the course of the visit, the copies of the instrument were administered to the students through the assistance of the physics teachers in the respective sampled schools. Those teachers were intimated on the major purpose of the study. The administration of the instruments was done once and retrieval of the questionnaire was at the spot.

Method of Data Analysis

The data were analyzed based on Anderson/Rasch (AR) model. The AR model states that at a given point in time, a person is seen as having propensities to answer in accordance with any of the conceptions (Newtonian, Impetus or Aristotelian). In AR model, tasks are parameterized in terms of their tendency to provoke different conceptions as well. Given physics tasks (e.g., a multiple-choice test with j items, each item having m choices that are each associated with a particular way of thinking about situations in the targeted domain) administered to N examinees, the idea of AR model can be presented as follows. In terms of items, each item can be characterized by a vector value containing m elements, with each value of the element corresponding to a location on a continuum for a certain property in physics in a sense that can be described as follows. For example, the first choice may represent the Newtonian approach, which is the correct approach, the second choice may represent the strategy using “impetus theory”, the third choice may reflect an Aristotelian belief, and so forth. In other words, the item parameter is a vector-valued parameter which contains m elements, and larger values for an element indicate a greater tendency for that item to elicit responses in line with the corresponding model approach. In particular, the choice with the highest value indicates that that way of thinking is more common on this item, all other things being equal.

For the responses to be adequate to be analyzed for the AR model, the data set was coded based on the three response categories. After the data was recoded, standard descriptive item analyses will be conducted for the data set, including a frequency distribution table, mean and percentage to study the influence of gender and location on students' achievement. Even though the categories 1-3 are, strictly speaking, nominal, they can be ordered from 1 representing a higher level of understanding than 2, which is again higher than 3. The data at this point was analyzed under the AR model, using the computer program WinBUGS. The result of the analysis from the WinBUGS was interpreted thus; for the item parameter b (j, k) where j represents the jth items, k represents the response categories, the b that has the highest value of the vector parameter has the highest tendency to elicit a particular response model followed by the next highest value and so on. Same is applicable to the person parameter ϴ (j, k) where j represents the jth persons, k represents the response categories. The ϴ for the three response categories that has the highest vector value indicates the most likely model that will be used by a person in responding to the FMCE items.

Finally, t-test for independent samples was used to test the two null hypotheses.

CHAPTER FOUR
RESULTS

This chapter was presented according to the research questions and the hypotheses that guided the study.

Research Question One: What is the students' propensity to use Newtonian, Aristotelian or Impetus model in responding to force and motion concepts? In order to answer this question, student's parameter (Ө) estimate was done using WinBUGS computer program.

Table 1: <u>Summary of the Number of Students whose Parameter Estimates Correspond to a Particular response Model</u>

	Newtonian	Impetus	Aristotelian	Total
Number of Students	34	138	53	225
Percentage (%)	15	61	24	100

Note: The data presented above is a summary of Appendix C

Table 1 above shows that out of the two hundred and twenty five students (225) used for the study, 34 (15%) of the students responded to force and motion concepts using the Newtonian model; that is, the model 1 which is the correct model, 138 (61%) of students responded using Impetus model; that is, model 2 which is nearer to the correct model while 53 (24%) of the students responded using Aristotelian model that is model 3. Considering the percentage values of the models, it is obvious that the students' propensity to use Impetus model (model 2) is higher since it has the highest percentage value. This is evident for the fact that the parameter estimate of appendix C for the model 2 has the greatest value for majority of the students. The implication of the students' propensity to use any of the models to scientific understanding of the concepts can be explained thus; students that use Newtonian model are expected to have acquired the required conceptual understanding of force and motion concepts. Those are the ones that can respond to force and motion concepts correctly irrespective of how twisted the questions may be. To the students, an arrow (with very small frictional forces) continued to fly through the air because of the law of inertia, while a block of wood on a table stopped sliding once the applied force was removed because of frictional forces. That is purely a Newtonian view of force and motion. Students that use Impetus model are the ones that have inadequate conceptual understanding of force and motion. To this class of students, When a mover sets a body in motion he implants into it a certain force called impetus, that is, a certain force enabling a body to move in the direction in which the mover starts it, be it upwards, downwards, sidewards, or in a circle. The implanted impetus increases in the same ratio as the velocity. It is because of this impetus that a stone moves on after the thrower has ceased moving it. Students that use Aristotelian model are the ones that do have little or no conceptual understanding of force and motion. To those students, a constant force applied to an object produces a constant speed which is also inversely proportional to the inertia of the object. In the absence of force, an object would stop immediately. The students believe that an object will not stop moving unless the initial force acting on the subject is “used up”, the “Aristotelian” or “medieval” belief. This believe contradicts the Newtonian view that a body will continue in its state of uniform motion except intercepted by an external force.

Research Question Two: What is the propensity of the items to elicit a particular response model to force and motion concepts? Question two was answered through the estimation of item parameter for the whole items of the instrument.

Table 2: Item Parameter Estimates for the FMCE Items

Item	Newtonian Model Mean(SD)	Impetus Model Mean(SD)	Aristotelian Model Mean
1	-0.40 (0.75)	0.13 (0.74)	0.27*
2	-0.87(0.75)	0.09(0.74)	0.78*
3	-0.73 (0.75)	0.06 (0.74)	0.67*
4	2.12 (0.81) *	1.02 (0.81)	-3.14
5	-0.49 (0.79)	1.93 (0.77) *	-1.44
6	-0.55 (0.82)	1.81 (0.71) *	-1.27
7	0.64 (0.80) *	0.03 (0.71)	-0.68
8	-1.15 (0.82)	0.71(0.71) *	0.44
9	-0.83 (0.82)	1.78 (0.71) *	-0.95
10	0.26 (0.71)	0.78 (0.74) *	-1.04
11	0.92 (0.71) *	-0.12 (0.74)	-0.81
12	0.23 (0.70) *	-0.30 (0.74)	0.07
13	0.43 (0.71)	0.46 (0.74) *	-0.90

*Note: * represents the model elicited by a particular item. The figures in parenthesis ( ) represent the standard deviations of items.*
Since the parameters with response category 3 are not estimated here, they were obtained by computing the sum of parameter estimates with the first two response categories and then reversing the sign of the sum, the associated SDs are not available.

Table 2 shows that items 4,7,11and 12 have the tendency of eliciting the use of model 1; items 5,6,8,9,10, and 13 have the tendency of eliciting the use of model 2 while items 1, 2, and 3 have the tendency of eliciting the use of model 3 in responding to force and motion concepts by the students respectively. That was arrived at by considering the fact that the parameter value for model 1 was greater than the other models for items 4, 7, 9, and 1. That was same for model 2 and model 3 for the other items. This implies that the use of the three models depends on the nature of the items. By implication, students that do not have adequate conceptual understanding of force and motion might at random use any of the three models in response to force and motion questions that are not clearly framed.

Research Question Three: What is the influence of gender on students' conceptual understanding of force and motion?

Table 3: Percentage and Mean Analysis of the Influence of Students' Gender on their Achievement on FMCE.

Students	Frequency	Percent	Mean Achievement	Standard Deviation
Male	97	43.1	4.11	1.74
Female	128	56.9	4.63	1.40
Total	225	100.0		

Table 3 shows that the mean achievement and the standard deviation of male students on FMCE test are 4.11 and 1.74 respectively while those of the female students are 4.63 and 1.40 respectively. Comparatively, female students achieve better than their male counterparts in FMCE test. This can be seen from the gap between the two mean achievements of which mean achievement for the females is higher than that of the males.

Research Question Four: What is the influence of school location on students' conceptual understanding of force and motion?

Table 4: Percentage and Mean Analysis of the Influence of Students' Location on their Achievement on FMCE.

Location	Frequency	Percent	Mean Achievement	Standard Deviation
Urban	105	46.7	4.34	1.58
Rural	120	53.3	4.49	1.56
Total	225	100.0		

The analysis in Table 4 implies that students from schools in urban location have their mean achievement and the standard deviation on the FMCE as 4.34 and 1.58 respectively while those for the students from schools in rural location are 4.49 and 1.56 respectively. Also it is clear from the analysis that there is small difference in the mean achievements for the students from urban schools and those from the rural schools. The mean achievement of those from the urban schools is relatively low in comparison with that of the students from rural schools. That suggests that students from rural schools might achieve higher in FMCE test than their counterparts from the urban schools.

Hypotheses

HO_1: Gender is not a significant factor in the students' conceptual understanding of force and motion. This null hypothesis was tested using t-test of proportion. This statistic is normally used for two independent samples.

Table 5: t-test Analysis of Influence of Students' Gender on their Achievement on FMCE test at 0.05 level of Significance.

Gender	N	Mean	Std. Deviation	Df	T	Sig. (2-tailed)
Male	97	4.11	1.74			
Female	128	4.63	1.40	223	-2.485	.014

The analysis in table 5 shows that the mean achievement and the standard deviation for the male students are 4.11 and 1.74 respectively while the female students' mean achievement and standard deviation are 4.63 and 1.40 respectively. Also the probability associated with t-test value of -2.485 is 0.014. Since this probability value is less than the 0.05 level of significance, the null hypothesis was rejected. Hence, the independent t-test analysis above shows that there is a significant difference in the mean achievements of male and female students on FMCE test at 0.05

significance levels in favour of female students. This indicates that the above difference observed in table 3 cannot be attributed to a chance factor.

HO_2: Location is not a significant factor in the students' conceptual understanding of force and motion.

Table 6: t-test Analysis of the Influence of Students' School Location on their Achievement on FMCE test at 0.05 level of Significance

Location	**N**	**Mean**	**Std. Deviation**	**Df**	**t**	**Sig. (2-tailed)**
Urban	105	4.34	1.58			
Rural	120	4.49	1.56	223	-.686	.494

Table 6 shows that the mean achievement and standard deviation for the students in urban schools are 4.34 and 1.58 respectively while those of the students in rural schools are 4.49 and 1.56 respectively. The probability associated with a t-test value of -.686 is .494. Since this value (.494) is greater than the 0.05 level of significance, the null hypothesis was accepted. The analysis in table 6 above reveals that there is no statistical significant difference in mean achievements between students in urban schools and those in rural schools at 0.05 significance levels. Thus, the mean achievement difference observed in table 4 was due to a chance factor.

Summary of Findings

- Most of the Students responded to force and motion concepts using the impetus model (model 2) which is not exactly the correct model.
- The level of conceptual understanding of the students is low. Only 15% have adequate conceptual understanding, 61% have inadequate or partial conceptual understanding while 24% have little or no conceptual understanding.
- The nature of items determines the students' choice of models.

- Gender has a significant influence on the students' conceptual understanding of force and motion in the direction of the female students.
- School location has no significant influence on the students' conceptual understanding of force and motion.

CHAPTER FIVE
DISCUSSION, CONCLUSION AND SUMMARY

This chapter was presented based on the discussion of the findings obtained through the analysis of data in chapter four. Based on the discussion of the major findings of the study, conclusions were drawn as well as summary of the major findings. Also presented in this chapter were: Educational implications of the study, recommendations, limitations of the study and suggestion for further study.

Discussion of the Findings

The discussion was done based on the analysis of data to answer the research questions and test the null hypotheses.

Students' Propensity to use Different Models in Response to Force and Motion Concepts

The analysis of data in Table 1 of chapter four reveals that out of the three models which students can adopt when presented with force and motion concepts, the model 2 has the highest value of the student's parameters when compared to the other two models. This indicates that majority of the students responded to force and motion concepts using the impetus model which is nearer to the correct model. This finding is in agreement with the findings of Chun-Wei (2003), Bao and Redish (2004), in their respective studies which are similar to the present study. Chun-Wei (2003) further observed that the students' conceptual understanding does change after instruction. This was evident when pretest and post testing of students were done in a single study by Chun-Wei (2003). In this case the students were using either model 1 or model 2 after instruction. This indicates that students are in a mixed model state (i.e., in a transition toward understanding Newtonian mechanics). This also implies that they still have difficulties in understanding some concepts related to force-motion (Chun-Wei, 2003). This could not be observed in the current study for the fact that the study was limited to single administration of the instrument.

Nature of Items and the Students' Choice of Models

The analysis of data in Table 2 indicates that the students' choice of models depends on the nature of the questions asked. That is, the item's level of difficulty determines the choice of response model that the student will make. This finding is in concordance with the findings of Chun-Wei (2003) who observed that the particular features of tasks evoke different response categories to equivalent items, indicating that the students are not in a "pure state" of Newtonian view of concepts.

Gender and Students' Achievement

The analysis in Table 3 reveals that mean achievement of female students is slightly higher than those of the male students. This mean achievement difference was strengthened by the t-test analysis in table 5 which showed that gender is a significant factor for students' conceptual understanding in favour of female students. This finding is in agreement with the findings of Shaibu and Mari (1997) who observed a gender difference in achievement in science process skills in favour of the female students. On the contrary, Ezanya (2004); Lynch and Paterson (2002) studies showed that gender is significant in chemistry achievement and science process skills respectively in favour of male students. This current study has been able to show that gender is a significant factor with respect to students' conceptual understanding of force and motion.

School Location and Students' Achievement

The analysis of data in Table 4 indicates that the mean achievement of students in rural schools is higher than their counterparts from the urban schools. This was proven to be due to a chance factor when t-test analysis of table 6 reveals that school location is not a significant factor in students' conceptual understanding of force and motion. This finding is in agreement with the studies by Jegede (1984) and Erubami (2003). Those studies revealed that school location is not significant in students' achievement in physics. This finding contradicts the findings of Isiugo-Abanihe and Labo-popoola (2004) that students in urban areas performed significantly better than those in rural areas and also this is in support of school location playing a significant role in academic performance of students.

Conclusion

Based on the discussion of the findings, the following conclusions were drawn;

1. Majority of the physics students used impetus model (model 2) when responding to force and motion concepts.
2. The level of conceptual understanding of the students is low.
3. The nature of items determines the students' choice of models when presented with questions on force and motion.
4. Gender has a significant influence in the students' conceptual understanding of force and motion in favour of the female students.
5. School location has no significant influence in the students' conceptual understanding of force and motion than.

Educational Implications of the Study

It is evident from the study that students use different models in response to force and motion concepts. In particular most of the students use model 2 which is nearer to the correct model. This by implication shows that the physics teachers need to improve on their teaching of physics in particular force and motion concepts in order to impact the right Newtonian view of the concepts to the students. It was observed that students consistently use model 2 due to the misconceptions they have on the concepts. If this misconception is not checked, the students achievement in those concepts will continue to be low.

It was also observed that the nature of questions does influence the students' choice of model. This implies that the nature of items determines the students' choice of models. To this effect, the teachers should be meticulous in setting questions based on force and motion concepts in order to elicit the right model from the students otherwise the students' chance of giving the correct response will be marred.

The influence of gender was found to be significant. This implies that the female students' achievement in conceptual understanding is higher than the male students. It is then the duties of the physics teachers to provide a classroom environment where the male students will be better predisposed to improve on their conceptual understanding in order to meet up with their counterpart.

School location was found not to be a significant factor in students' conceptual understanding of force and motion. This shows that irrespective of one's school location, the achievement of conceptual understanding of force and motion is the same. The implication of this is that the physics teachers should give every student equal opportunity to learn not minding the school location of the students in order to get the best out of the students.

Recommendations of the Study

Based on the above mentioned implications, the following recommendations were made;

1. Physics teachers should be properly assessed to understand their level of conceptual understanding of force and motion. Doing this will help to improve on the students' conceptual understanding of force and motion.
2. Wording of questions on force and motion concepts by the physics teachers should be done with ultimate caution bearing in mind that the nature of items influences the students' choice of models.
3. More efforts should be made by the physics teachers in the schools to enable the male students in those schools to compete favourably with their female counterparts.
4. Every physics student should be given equal opportunity to prove him/herself in the learning of force and motion concepts irrespective of one's school location.

Limitations of the Study

When making inferences from this study, one should be aware that the sample size for each data set is relatively small, and the content scope was drawn from physics. The students used may not be a good representation of other student populations. Therefore, the results of this study need to be interpreted with caution. Due to the relatively small sample size (by the standards of psychometric analyses such as latent class and IRT modeling), not much information is available to estimate parameters for individual items and students. The study was also limited to only single administration of the instrument and so could not come up with the conclusion that students' conceptual understanding could change after instruction.

Suggestions for Further Studies

There are four perspectives to study the nature of human mind as recognized by the study as: the differential, behaviorist, cognitive, and situative perspectives.

The last two approaches have not been explored in great detail, especially in the field of educational measurement. The current study explored students' learning in terms of cognitive perspective. However, the study only focused on a small piece of cognitive process (i.e., the mixture-within-persons strategy for problem-solving). For this aspect of learning, modeling students' problem-solving in terms of common misconceptions has proven useful.

Other fields of learning may not be the same as physics in this regard, so they may focus on different aspects of cognitive process. This study focused on the assessment of students' conceptual understanding based on the conceptual assessment framework. There is need to continue to follow this line of research to explore, for example, how students solve a class of math tasks and how it is different from physics learning. More studies are needed to explore how social or cultural factors affect students' learning. For example, one may be interested in investigating whether the different test formats (multiple-choice test vs. open-ended questions) affect students' conceptual understanding of force and motion and how it occurs.

Summary of the Study

The major purpose of this study was to assess the SS2 physics students' conceptual understanding of force and motion based on the tenets of conceptual assessment framework. To achieve the major purpose of the study, four research questions were posed and as two null hypotheses were formulated. From the review of related literature, two research designs were used for the study. Two hundred and twenty five senior secondary two physics students from six government's owned secondary schools in Nsukka Education Zone were used for the study. The six schools were purposively sampled, three from urban schools and three from rural schools. Seven intact classes were used for the study.

The data obtained for the study were analyzed using winBUGS computer program, percentage, mean and t-test statistics. The findings of the study are:

1. Most of the Students respond to force and motion concepts using the impetus model (model 2) which is not exactly the correct model.
2. The difficulty level of items does influence the students' choice of models.
3. Gender has a significant influence on the students' conceptual understanding of force and motion in the direction of the female students.
4. School location has no significant influence on the students' conceptual understanding of force and motion.

From the findings of the study, conclusions were drawn as well as the implications of the study. Also from the educational implications, recommendations were made. Limitations as well as suggestions for further studies were highlighted.

References

Adedayo, A.O (1997). The Influence of parental Socio-Economic factors on Students' Achievement in Mathematics. *Nigerian Journal of Research in Education. 1(1)*, *175-179.*

Adedayo, O.A. (1997). Gender, environment and co-education as factors of performance in the Ravens Standard Progressive Matrices. *Gombe Technical Education Journal. 1(1), 16-20.*

Ajejalami, D. (1990). Science and Technology Education in Africa focus on seven sub-Saharan Countries. Lagos: University of Lagos Press, Lagos.

Akinbode, M (2005): *Effect of school location and gender differences on students academic achievement in Economic*: A case study of Abeokuta North Local Government Area of Ogun State. *(An unpublished Ph.D thesis)*

Albert, J.H. (1992). Bayesian estimation of normal ogive item response curves using Gibbs sampling. *Journal of Educational Statistics, 17,* 251-269.

Andersen, E. B. (1995). Polytomous Rasch models and their estimation. In G. H. Fischer & I. W. Molen (Eds.), *Rasch models: Foundations, recent, and applications* (pp. 271-291). New York: Springer-Verlag.

Andersen, E.B. (1973). *Conditional inference and models for measuring.* Copenhagen: Danish Institute for Mental Health.

Anyakoha, M.W. (2007). *New school physics for senior secondary schools.* Africana Feb Publishers limited.

Ariyo, A. O. (2006). School and student factors as determinants of students' achievement in physics at the senior secondary school level in Oyo State, Nigeria. *An unpublished Ph.D thesis, University of Ibadan, Ibadan.*

Ates, J.B. (2008). Research on gender issues in the classroom. *Handbook of research on science teaching and learning*. New York: Macmillan.

Ayogu, Z.U; & Nworgu, B.G (1999). Influence of gender and school location on students' achievement in Physics. *40th Annual Conference Proceedings of Science Teachers' Association of Nigeria.217-221*

Baker, F. B. (1998). An investigation of the item parameter recovery characteristics of a Gibbs sampling procedure. *Applied Psychological Measurement, 22(2)*, 153-169.

Bao, L. & Redish, E. F. (2004). Educational assessment and underlying models of cognition., *The scholarship of teaching and learning in higher education: Contributions of research universities* (chap. 11). Indiana University Press.

Bao, L., & Redish, E. F. (2001). Concentration analysis: A quantitative assessment of student states. *Physics Education Research Section of American Journal of Physics, 69 (7)*, 45-53.

Beguin, A. A. & Glas, C. A. W. (2001). MCMC estimation and some model-fit analysis of multidimensional IRT models. *Psychometrika, 66(4)*, 541-561.

Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education. 5,* 7-74.

Bruner, J. (1999). *The Process of Education*. Cambridge, MA: Harvard University Press.

Bruner, J. (1966). *Toward a Theory of Instruction*. Cambridge, MA: Harvar University Press.

Bueche, F.J. (1988). *Principles of Physics*. (5th Edition). Singapore: Mc. Graw-Hill Inc.

Chun-Wei, H. (2003). Psychometric analyses based on Evidence-centered design and Cognitive Science of learning to explore students' Problem-solving in physics. *Dissertation submitted to the Faculty of the Graduate School of the University of Maryland, College Park.* Retrieved from http://drum.lib.umd.edu/bitstream/1903/33/1/dissertation.pdf.

Clement, J. (1982). Students' Preconceptions in Introductory Mechanics. *AmericanJournal of Physics 50*, 66-71.

Effiong,U*.U* (2002). Science classroom environment factors, cognitive preferences and achievement in physics among senior secondary school students in Cross River State. *Unpublished Ph.D thesis, University of Calabar, Calabar.*

Embretson, S. E. (1995). A measurement model for linking individual learning to processes and knowledge: Application to mathematical reasoning. *Journal of Educational Measurement, 32(3)*, 277-294.

Emina, F.I. (1989). Effect of learning environment on pupils' achievement in science in Bendel State. *Journal of Research in science Education. 1(2),269-275.*

Erubami, D.A (2003). Gender and socio-economic factors as correlates of students' academic performance in sciences at the senior secondary Certificate Examination in Ogun State. *Unpublished M.Ed thesis, University of Benin, Benin City.*

Eryilmaz, H (2004). The effect of peer instruction on high school students' achievement and attitudes toward physics. *Ph.D thesis of Middle East Technical University.*

Ezanya L.4 (2004). Development and preliminary validation of a chemistry achievement test. *unpublished M.Ed thesis. University of Nigeria, Nsukka.*

Ezeife, A.N (1990). Sex and Environment as factors in physics Achievement. *International Journal of Empirical studies in psychology and Education.1(2),52-59.*

Ezeudu S.A (2003). Classroom environment as correlate of students' cognitive achievement in senior secondary school geography. Ythe journal of WCCI Nigerian chapter. Vol.4(2) oct, 2003, 65-73

Fakorede A.D. (1999). An investigation into gender difference and students achievement in secondary school biology: A case study of Oyo State. *An unpublished M.Ed project*.

Fox, J .P. & Glas, C. A. W. (2001). Bayesian estimations of a multilevel IRT model using Gibbs sampling. *Psychometrika, 66(2)*, 271-288

Gelman, A., Carlin, J.B., Stern, H.S., & Rubin, D.B. (1995). *Bayesian data analysis.* London: Chapman & Hall.

Glas, C. A. W., & Falcón, J. C. S. (2003). A comparison of item-fit statistics. *Applied Psychological Measurement, 27(2)*, 87-106.

Glas, C. A. W., & Verhelst, N. D. (1995). Testing the Rasch model. *Rasch models: foundations, recent developments, and applications* (pp. 69-96). New York: Springer-Verlag.

Gonzuk , S. & Chargok, H. (2001) , *Gender differences in science: Parallels in interest, experience, and performance*. *International Journal of Science Education 9, 467-481 (1987).*

Gould, A.(2005). What students' learning of representations tells us about constructivism. *The Journal of Mathematical Behavior 19*, 481-502.

Greeno, J. G., Collins, A. M., & Resnick, L. B. (1996). Cognition and learning. *Handbook of educational psychology* (pp. 15-46). New York: Macmillan.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Kluwer: Boston.

Hestenes, D., Wells, M., & Swackhammer, G. (1992). Force Concept Inventory. *The Physics Teacher, 30*, 141-153.

Hoijtink, H., & Molenaar, I. W. (1997). A multidimensional item response model: Constrained latent class analysis using the Gibbs sampler and posterior predictive checks. *Psychometrika, 62*, 171-189.

Isiugo-Abanihe I. & Labo-Popoola O. S. (2004). School type and location as environmental factors in learning English as a second language. *West African Journal of Education, 23*(1), 55-71.

Jammer,M. (1957). Concepts of force: a study of dynamics. Cambridge Mass: Harvard University Press: Retrieved from, http://www.thecatalyst.org/physics/chapter-two.html

Jegede, B.A (1984). Non-Cognitive Correlates of Senior Secondary Students' Achievement in Physics. *Journal of Science Teachers' Association of Nigeria. 22(2), 78-88.*

Johnson S.K (2005). Conceptual and procedural knowledge in mathematics: An introductory analysis. *Conceptual and procedural knowledge: The case of mathematics* (pp. 1-27). Hillsdale, NJ: Erlbaum. Retrieved from, http://wik.ed.uiuc.edu/index.php/Conceptual_knowledge

Kim, S.H. (2001). An evaluation of a Markov chain Monte Carlo method for the Rasch model. *Applied Psychological Measurement, 25(2)*, 163-176.

Leclerc, I. (1972). The nature of physical existence. New York: Humanities press. Retrieved from, http://www.thecatalyst.org/physics/chapter-two.html

Lynch, D. A; & Paterson, M. (2002). Gender inequality in science recognition in Africa: The causes, consequences and solution. *Unpublished Seminar paper*.

Ma, X and Wilkins, J.L.M (2002). The development of science achievement in middle and high school. individual difference and school effects. *Evaluation Review, 26(4), 395-417.*

MathSoft, Inc. (1995). S-PLUS (Version 3.3 for Windows) [Computer program]. Seattle WA: Author.

McCloskey, J.,Caramazza, A. & Green, B. (1980). Naïve Beliefs in "Sophisticated" Subjects: *Misconceptions about Trajectories of Subjects Cognition*, 9, 117-123.

McDermott, L.C. (1984). Research on Conceptual Understanding in Mechanics: Physics Today, July, 24-32.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Education Researcher, 23*(2), 13-23.

Minstrell, J. (1982). Explaining the "at rest" condition of an object. *The Physics Teacher*, *20(1)*, 10-14.

Mislevy, R. J., & Verhelst, R. D. (1990). BILOG 3: *Item analysis and test scoring with binary logistic models [Computer program*]. Mooresville IN: Scientific Software.

Mislevy, R.J., Steinberg, A & Almond, D.C. (2003). On the structure of educational assessments. *Measurement, 1,*3-67.

Mpaoro, I.E (2003). The relationship among test anxiety, academic achievement and interest of students in senior secondary school mathematics in River state.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16(2)*, 159-176.

National Research Council (1999). *How people learn: Brain, mind, experience, and school.* Committe on Developments in the Science of Learning. Washington, DC: National Academy Press.

National Research Council (2001a). *Knowing what students know: The science and design of educational assessment.* Committee on the Foundations of Assessment. Washington, DC: National Academy Press.

National Research Council. (2001b). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Nworgu, B.G. (1994). The influence of environment and gender on the level of science students: Implications for Rural Development. *Journal of Research in Education and Humanities,* 2(1); 9-19.

Nworgu, B.G. (2006). *Educational research: basic issues and methodology (revised and enlarged edition).* Nsukka: University Trust Publishers.

Okonkwo, S.C (2000). Relationship between some school and teacher variables and students achievement in mathematics. *Journal of Science Teachers' Association of Nigeria. 35(1&2),43-49.*

Olubunmi A, (2001). The impact of school location and gender difference on students' achievement in agricultural science: A case study of Ogun State.

Patz, R. J., & Junker, B. W. (1999a). A straightforward approach to Markov chain Monte Carlo methods for item response models. *Journal of Educational and Behavioral Statistics, 24(2)*, 146-178.

Patz, R. J., & Junker, B. W. (1999b). Applications and extensions of MCMC in IRT: Multiple item types, missing data, and rated responses. *Journal of Educational and Behavioral Statistics, 24(4)*, 342-366.

Rasch, G. (1961). On general laws and meaning of measurement in psychology. *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability* (Vol. 4). Berkeley: University of California Press.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*, 346-362.

Rost, J. (1991). A logistic mixture distribution model for polychotomous item responses. *British Journal of Mathematical and Statistical Psychology, 44*, 75-92.

Samejima, F. (1995). Acceleration model in the heterogeneous case of the general graded response model. *Psychometrika, 60(4)*, 549-572.

Serway, R.A., & Beichner, R.J (2001). *Physics for scientists and engineers with Modern Physics.* Saunders College publishing, a division of Harcourt College Publishers.

Shaibu, A.M and Mari, J.K (1997). Gender related differences in the understanding of science process skills amongst junior secondary school students in some Nigerian Schools. *Journal of Scienc Teachers' Association of Nigeria. 32(1&2),21-27*

Spiegelhalter, D.J., Thomas, A., & Gilks, W.R. (1997). BUGS: *Bayesian inference using Gibbs sampling (Version 0.60) [Computer program]*. Cambridge, UK: University of Cambridge, Institute of Public Health, Medical Research Council Biostatistics Unit.

Swaminathan, H., and Gifford, J. A. (1986). Bayesian estimation in the three-parameter logistic model. *Psychometrika, 51(4)*, 589-601.

Tara, S.A (2009). Preconception as a factor influencing students' understanding of classical mechanics. *American Journal of Physics 48*, 201-279.

The MathWorks, Inc. (1996). MATLAB: *The language of technical computing [Computer program]*. Natick MA: Author.

Thornton, P. K., & Sokoloff, D. R. (1998). Assessing student learning of Newton's laws: the force and motion conceptual evaluation. *American Journal of Physics, 66(4)*, 338-351.

Tsutakawa, R. K., & Johnson, J. C. (1990). The effect of uncertainty of item parameter estimation on ability estimates. *Psychometrika, 55*, 371-390.

Unachukwu, G.C (1999). Sex differentials in mathematics achievement of primary school pupils in Enugu State of Nigeria, *Nigeria Journal of Empirical Studies in Psychology and Education. 1(1), 85-90*

Van den Wollenberg, A. L. (1982). Two new test statistics for the Rasch model. *Psychometrika, 47*, 123-139.

Whitaker, E.A. (1983). A cross-age study of student understanding of the concept of velocity, *Journal of Research in Science Teaching, 28(8)*, 649-660.

Yang, D. H. (2010). Gender and classroom learning. *Psychology in the Schools, 22,* 08–223.

Young, S. L. (1994) Self-efficacy beliefs, motivation, race, and gender in middle school science. *Journal of Women and Minorities in Science and Engineering, 7*, 271–285.

APPENDIX A

S/N	Name of school	Location	No of students
1	Model Secondary School, Nsukka	Urban	25
2	St. Teresa’s College Nsukka	Urban	45
3	Queen of the Rosary Sec School, Nsukka	Rural	40
4	Urban Girl’s Secondary School, Nsukka	Urban	35
5	Community Secondary School, Edeoballa	Rural	55
6	Community Secondary School, Nru	Rural	25
		Total	225

APPENDIX B
INSTRUMENT FOR DATA COLLECTION

SECTION A: Personal Data of SS2 Students

Name: ---

Class: -----------------------

Sex: Male ☐

Female ☐

School Location: Urban ☐

Rural ☐

SECTION B: FORCE MOTION CONCEPT EVALUATION

Instruction: Choose the correct option(s) that best fit(s) each of the items of the instrument.

1. A boy throws a steel ball straight up. Discarding any effects of air resistance, the force(s) acting on the ball until it returns to the ground is (are):

A) Its weight vertically downward along with a steadily decreasing upward force.

B) A steadily decreasing upward force from the moment it leaves the hand until it reaches its highest point beyond which there is a steadily increasing downward force of gravity as the object gets closer to the earth.

C) A constant downward force of gravity along with an upward force that steadily decreases until the ball reaches its highest point, after which there is only the constant downward force of gravity.

D) A constant downward force of gravity only.

E) None of the above, the ball falls back down to the earth simply because that is its natural action.

2. The main forces acting, after the “kick”, on the ball along the path you have chosen are:

A) The downward force due to gravity and the effect of air pressure.

B) The downward force of gravity and the horizontal force of momentum in the direction of motion.

C) The downward force of gravity, the upward force exerted by you, and horizontal force acting on the ball in the direction of motion.

D) The downward force of gravity and an upward force exerted on the ball by you.

E) Gravity does not exert a force on the ball, it falls because of intrinsic tendency of the object to fall to its natural place.

3. An elevator is being lifted up an elevator shaft by a steel cable. When the elevator is moving up the shaft at a constant velocity:

A) The upward force on the elevator by the cable is greater than the downward force of gravity.

B) The amount of upward force on the elevator by the cables equals that of the downward force of gravity.

C) The upward force on the elevator by the cable is less than the downward force of gravity.

D) It goes up because the cable is being shortened, not because of the force being exerted on the elevator by the cable.

E) The upward force on the elevator by the cable is greater than the downward force due to the combined effects of air pressure and the force of gravity.

4. A golf ball driven down a fairway is observed to travel through the air with a trajectory (flight path). Which following force(s) is (are) acting on the golf ball during its entire flight?

1. The force of gravity

2. The force of the "hit"

3. The force of air resistance

A) 1 only

B) 1 and 2

C) 1, 2, and 3

D) 1 and 3

E) 2 and 3

5. A large box is being pushed across the floor at a constant speed of 4.0 m/s. What can you conclude about the forces acting on the box?

A) If the force applied to the box is doubled, the constant speed of the box will increase to 8.0 m/s.

B) The amount of force applied to move the box at a constant speed must be more than its weight.

C) The amount of force applied to move the box at a constant speed must be equal to the amount of the frictional forces that resist its motion.

D) The amount of force applied to move the box at a constant speed must be more than the amount of the frictional force that resist its motion,

E) There is a force being applied to the box to make it move but the external forces such as friction are not "real" forces they just resist motion.

A sled on ice moves in the ways described in questions 6-7 below. Friction is so small that it can be ignored. A person wearing spiked shoes standing on the ice can apply a force to the sled and push it along the ice. Choose the one force (**A** through **G**) which would **keep the sled moving** as described in each statement below.

A. The force is toward the right and is increasing in strength (magnitude).

B. The force is toward the right and is of constant strength (magnitude).

C. The force is toward the right and is decreasing in strength (magnitude).

D. No applied force is needed.

E. The force is toward the left and is decreasing in strength (magnitude).

F. The force is toward the left and is of constant strength (magnitude).

G. The force is toward the left and is increasing in strength (magnitude).

6. Which force would keep the sled moving toward the right at a steady rate (constant) velocity?

7. The sled was started from rest and pushed until it reached a steady (constant) velocity toward the right. Which force would keep the sled moving at this velocity?

Questions 8-10 refer to a coin which is tossed straight up into the air. After it is released it moves upward, reaches its highest point and falls back down again. Use one of the following choices (A through G) to indicate the force acting on the coin for each of cases described below. Ignore any effects of air resistance.

A. The force is down and constant.

B. The force is down and increasing.

C. The force is down and decreasing.

D. The force is zero.

E. The force is up and constant.

F. The force is up and increasing.

G. The force is up and decreasing.

8. The coin is moving upward after it is released.

9. The coin is at its highest point.

10. The coin is moving downward.

Questions 11-13 refer to a toy car which is given a quick push so that it rolls up an inclined plane. After it is released, it rolls up, reaches its highest point and rolls back down again.

Friction is so small that it can be ignored. Use one of the following choices (**A** through **G**) to indicate the net force acting on the car for each of the cases described below.

A. Net **constant** force **down** plane.

B. Net **increasing** force **down** plane.

C. Net **decreasing** force **down** plane.

D. Net force zero.

E. Net **constant** force **up** plane.

F. Net **increasing** force **up** plane.

G. Net **decreasing** force **up** plane.

11. The car is moving up the plane after it is released.

12. The car is at its highest point.

13. The car is moving down the plane.

Model 1= Newtonian model

Model 2 = Impetus model

Model 3 = Aristotelian model

Questions	Model 1	Model 2	Model 3
1)	d	a, b,c	e
2)	a, d	b, c	e
3)	b	a, e	c, d
4)	a, d	b, c, e	N/A

5)	c	a, d, e	b
6)	d	b	others
7)	d	b	others
8)	a	g	others
9)	a	d	others
10)	a	g	others
11)	a	d	others
12)	a	b	others
13)	a	b	others

APPENDIX C
Students' Parameter Estimates for the FCME Items

Students	Newtonian Model Mean (SD)	Impetus Model Mean (SD)	Aristotelian Model Mean
1	-0.39 (1.02)	0.10 (0.94)	0.29
2	0.15 (0.98)	-0.35 (0.95)	0.20
3	-1.01 (1.05)	0.55 (0.93)	0.45
4	1.15 (1.01)	0.56 (0.99)	-1.71
5	0.24 (1.03)	0.29 (0.98)	-0.53
6	0.71 (1.01)	0.93 (1.01)	-1.65
7	0.71 (1.02)	0.93 (1.01)	-1.64
8	0.28 (0.96)	0.08 (0.96)	-0.36
9	1.09 (1.01)	0.54 (1.01)	-1.63
10	0.71 (1.01)	0.93 (1.07)	-1.64
11	0.18 (0.77)	0.88 (0.76)	-1.06
12	-0.62 (0.76)	-0.12 (0.75)	0.74
13	0.63 (0.76)	-1.09 (0.76)	0.46
14	-0.39 (1.02)	0.10 (0.94)	0.29
15	0.15 (0.98)	-0.35 (0.95)	0.20
16	-1.01 (1.05)	0.55 (0.93)	0.45
17	1.15 (1.01)	0.56 (0.99)	-1.71
18	0.24 (1.03)	0.29 (0.98)	-0.53
19	0.61 (1.01)	0.88 (1.01)	1.65
20	0.71 (1.02)	0.93 (1.01)	-1.64
21	0.28 (0.96)	0.08 (0.96)	-0.36
22	1.09 (1.01)	0.54 (1.01)	-1.63
23	0.71 (1.01)	0.93 (1.07)	-1.64
24	0.18 (0.77)	0.88 (0.76)	-1.06
25	-0.62 (0.76)	-0.12 (0.75)	0.74
26	0.63 (0.76)	-1.09 (0.76)	0.46
27	-0.39 (1.02)	0.10 (0.94)	0.29
28	0.15 (0.98)	-0.35 (0.95)	0.45
29	1.15 (1.01)	0.56 (0.99)	-1.71
30	0.24 (1.03)	1.15 (1.01)	0.45
31	0.56 (0.99)	0.29 (0.98)	-0.53
32	0.71 (1.01)	0.93 (1.01)	-1.65
33	0.28 (0.96)	0.08 (0.96)	-0.36
34	1.09 (1.01)	0.54 (1.01)	-1.63
35	0.71 (1.01)	0.93 (1.07)	-1.64
36	0.18 (0.77)	0.88 (0.76)	-1.06
37	-0.62 (0.76)	-0.12 (0.75)	0.74
38	0.63 (0.76)	-1.09 (0.76)	0.46
39	-0.39 (1.02)	0.10 (0.94)	-0.39
40	0.10 (0.94)	-1.09 (0.76	0.29
41	0.15 (0.98)	0.35 (0.95)	0.20
42	1.15 (1.01)	0.56 (0.99)	-1.71
43	0.24 (1.03)	0.29 (0.98)	-0.53
44	0.71 (1.01)	0.93 (1.01)	-1.65
45	0.28 (0.96)	0.08 (0.96)	-0.36
46	1.09 (1.01)	0.54 (1.01)	-1.63

47	0.71 (1.01)	0.93 (1.07)	-1.64
48	0.18 (0.77)	0.88 (0.76)	-1.06
49	-0.39 (1.02)	0.10 (0.94)	0.20
50	0.15 (0.98)	-0.35 (0.95)	0.45
51	-1.01 (1.05)	0.55 (0.93)	-1.71
52	1.15 (1.01)	0.56 (0.99)	-0.53
53	0.24 (1.03)	0.29 (0.98)	-1.65
54	0.71 (1.01)	0.93 (1.01)	-1.64
55	0.71 (1.02)	0.93 (1.01)	-0.36
56	0.28 (0.96)	0.08 (0.96)	-1.63
57	1.09 (1.01)	0.54 (1.01)	-1.64
58	0.71 (1.01)	0.93 (1.07)	-1.06
59	0.18 (0.77)	0.88 (0.76)	0.74
60	-0.62 (0.76)	-0.12 (0.75)	0.46
61	-0.39 (1.02)	0.10 (0.94)	0.20
62	0.15 (0.98)	-0.35 (0.95)	0.45
63	-1.01 (1.05)	0.55 (0.93)	-1.71
64	1.15 (1.01)	0.56 (0.99)	-0.53
65	0.24 (1.03)	0.29 (0.98)	-1.65
66	0.71 (1.01)	0.93 (1.01)	-1.64
67	0.71 (1.02)	0.93 (1.01)	-0.36
68	0.28 (0.96)	0.08 (0.96)	-1.63
69	1.09 (1.01)	0.54 (1.01)	-1.64
70	0.71 (1.01)	0.93 (1.07)	-1.06
71	0.18 (0.77)	0.88 (0.76)	0.74
72	-0.62 (0.76)	-0.12 (0.75)	0.46
73	0.63 (0.76)	-1.09 (0.76)	0.29
74	-0.39 (1.02)	0.10 (0.94)	0.20
75	0.15 (0.98)	-0.35 (0.95)	0.45
76	-1.01 (1.05)	0.55 (0.93)	-1.71
77	1.15 (1.01)	0.56 (0.99)	-0.53
78	0.24 (1.03)	0.29 (0.98)	1.65
79	0.61 (1.01)	0.88 (1.01)	-1.64
80	0.71 (1.02)	0.93 (1.01)	-0.36
81	0.28 (0.96)	0.08 (0.96)	-1.63
82	1.09 (1.01)	0.54 (1.01)	-1.64
83	0.71 (1.01)	0.93 (1.07)	-1.06
84	0.18 (0.77)	0.88 (0.76)	0.74
85	-0.62 (0.76)	-0.12 (0.75)	0.46
86	0.63 (0.76)	-1.09 (0.76)	0.29
87	-0.39 (1.02)	0.10 (0.94)	0.45
88	0.15 (0.98)	-0.35 (0.95)	-1.71
89	1.15 (1.01)	0.56 (0.99)	0.45
90	0.24 (1.03)	1.15 (1.01)	-0.53
91	0.56 (0.99)	0.29 (0.98)	-1.65
92	0.71 (1.01)	0.93 (1.01)	-0.36
93	0.28 (0.96)	0.08 (0.96)	-1.63
93	1.09 (1.01)	0.54 (1.01)	-1.64
95	0.71 (1.01)	0.93 (1.07)	-1.06
96	0.18 (0.77)	0.88 (0.76)	0.74

97	-0.62 (0.76)	-0.12 (0.75)	0.46
98	0.63 (0.76)	-1.09 (0.76)	-0.39
99	-0.39 (1.02)	0.10 (0.94)	0.29
100	0.10 (0.94)	-1.09 (0.76	0.20
101	0.15 (0.98)	0.35 (0.95)	-1.71
102	1.15 (1.01)	0.56 (0.99)	-0.53
103	0.24 (1.03)	0.29 (0.98)	-1.65
104	0.71 (1.01)	0.93 (1.01)	-0.36
105	0.28 (0.96)	0.08 (0.96)	-1.63
106	1.09 (1.01)	0.54 (1.01)	-1.64
107	0.71 (1.01)	0.93 (1.07)	-1.06
108	0.18 (0.77)	0.88 (0.76)	0.20
109	-0.39 (1.02)	0.10 (0.94)	0.45
110	0.15 (0.98)	-0.35 (0.95)	-1.71
111	-1.01 (1.05)	0.55 (0.93)	-0.53
112	1.15 (1.01)	0.56 (0.99)	-1.65
113	0.24 (1.03)	0.29 (0.98)	-1.64
114	0.71 (1.01)	0.93 (1.01)	-0.36
115	0.71 (1.02)	0.93 (1.01)	-1.63
116	0.28 (0.96)	0.08 (0.96)	-1.64
117	1.09 (1.01)	0.54 (1.01)	-1.06
118	0.71 (-0.39	0.93 (1.07)	0.74
119	0.15 (0.98)	0.88 (0.76)	0.46
120	-1.01 (1.05)	-0.12 (0.75)	0.29
121	1.15 (1.01)	0.10 (0.94)	0.20
122	0.24 (1.03)	-0.35 (0.95)	0.45
123	0.71 (1.01)	0.55 (0.93)	-1.71
124	0.71 (1.02)	0.56 (0.99)	-0.53
125	0.28 (0.96)	0.29 (0.98)	-1.65
126	1.09 (1.01)	0.93 (1.01)	-1.64
127	0.71 (1.01)	0.93 (1.01)	-0.36
128	0.18 (0.77)	0.08 (0.96)	-1.63
129	-0.62 (0.76)	0.54 (1.01)	-1.64
130	0.63 (0.76)	0.93 (1.07)	-1.06
131	-0.39 (1.02)	0.88 (0.76)	0.74
132	0.15 (0.98)	-0.12 (0.75)	0.46
133	-1.01 (1.05)	-1.09 (0.76)	0.29
134	1.15 (1.01)	0.10 (0.94)	0.20
135	0.24 (1.03)	-0.35 (0.95)	0.45
136	0.61 (1.01)	0.55 (0.93)	-1.71
137	0.71 (1.02)	0.56 (0.99)	-0.53
138	0.28 (0.96)	0.29 (0.98)	1.65
139	1.09 (1.01)	0.88 (1.01)	-1.64
140	0.71 (1.01)	0.93 (1.01)	-0.36
141	0.18 (0.77)	0.08 (0.96)	-1.63
142	-0.62 (0.76)	0.54 (1.01)	-1.64
143	0.63 (0.76)	0.93 (1.07)	-1.06
144	-0.39 (1.02)	0.88 (0.76)	0.74
145	0.15 (0.98)	-0.12 (0.75)	0.46
146	1.15 (1.01)	-1.09 (0.76)	0.29

147	0.24 (1.03)	0.10 (0.94)	0.45
148	0.56 (0.99)	-0.35 (0.95)	-1.71
149	0.71 (1.01)	0.56 (0.99)	0.45
150	0.28 (0.96)	1.15 (1.01)	-0.53
151	1.09 (1.01)	0.29 (0.98)	-1.65
152	0.71 (1.01)	0.93 (1.01)	-0.36
153	0.18 (0.77)	0.08 (0.96)	-1.63
154	-0.62 (0.76)	0.54 (1.01)	-1.64
156	0.63 (0.76)	0.93 (1.07)	-1.06
157	-0.39 (1.02)	0.88 (0.76)	0.74
158	0.10 (0.94)	-0.12 (0.75)	0.46
159	0.15 (0.98)	-1.09 (0.76)	-0.39
160	1.15 (1.01)	0.10 (0.94)	0.29
161	0.24 (1.03)	-1.09 (0.76	0.20
162	0.71 (1.01)	0.35 (0.95)	-1.71
163	0.28 (0.96)	0.56 (0.99)	-0.53
164	1.09 (1.01)	0.29 (0.98)	-1.65
165	0.71 (1.01)	0.93 (1.01)	-0.36
166	0.18 (0.77)	0.08 (0.96)	-1.63
167	-0.39 (1.02)	0.54 (1.01)	-1.64
168	0.15 (0.98)	0.93 (1.07)	-1.06
169	-1.01 (1.05)	0.88 (0.76)	0.20
170	1.15 (1.01)	0.10 (0.94)	0.45
171	0.24 (1.03)	-0.35 (0.95)	-1.71
172	0.71 (1.01)	0.55 (0.93)	-0.53
173	0.71 (1.02)	0.56 (0.99)	-1.65
174	0.28 (0.96)	0.29 (0.98)	-1.64
175	1.09 (1.01)	0.93 (1.01)	-0.36
176	0.71 (1.01)	0.93 (1.01)	-1.63
177	0.18 (0.77)	0.08 (0.96)	-1.64
178	-0.62 (0.76)	0.54 (1.01)	-1.06
179	0.18 (0.77)	0.93 (1.07)	0.74
180	-0.62 (0.76)	0.88 (0.76)	0.46
181	-0.39 (1.02)	-0.12 (0.75)	0.51
182	0.15 (0.98)	0.10 (0.94)	-1.13
183	-1.01 (1.05)	-0.35 (0.95)	0.29
184	1.15 (1.01)	0.55 (0.93)	0.20
185	0.24 (1.03)	0.56 (0.99)	0.45
186	0.71 (1.01)	0.29 (0.98)	-1.71
187	0.71 (1.02)	0.93 (1.01)	-0.53
188	0.28 (0.96)	0.93 (1.01)	-1.65
189	1.09 (1.01)	0.08 (0.96)	-1.64
190	0.71 (1.01)	0.54 (1.01)	-0.36
191	0.18 (0.77)	0.93 (1.07)	-1.63
192	-0.62 (0.76)	0.88 (0.76)	-1.64
193	0.63 (0.76)	-0.12 (0.75)	-1.06
194	-0.39 (1.02)	-1.09 (0.76)	0.74
195	0.15 (0.98)	0.10 (0.94)	0.46
196	-1.01 (1.05)	-0.35 (0.95)	0.29
197	1.15 (1.01)	0.55 (0.93)	0.20

198	0.24 (1.03)	0.56 (0.99)	0.45
199	0.61 (1.01)	0.29 (0.98)	-1.71
200	0.71 (1.02)	0.88 (1.01)	-0.53
201	0.28 (0.96)	0.93 (1.01)	1.65
202	1.09 (1.01)	0.08 (0.96)	-1.64
203	0.71 (1.01)	0.54 (1.01)	-0.36
204	0.18 (0.77)	0.93 (1.07)	-1.63
205	-0.62 (0.76)	0.88 (0.76)	-1.64
206	0.63 (0.76)	-0.12 (0.75)	-1.06
207	-0.39 (1.02)	-1.09 (0.76)	0.74
208	0.15 (0.98)	0.10 (0.94)	0.46
209	1.15 (1.01)	-0.35 (0.95)	0.29
210	0.24 (1.03)	0.56 (0.99)	0.45
211	0.56 (0.99)	1.15 (1.01)	-1.71
212	0.71 (1.01)	0.29 (0.98)	0.45
213	0.28 (0.96)	0.93 (1.01)	-0.53
214	1.09 (1.01)	0.08 (0.96)	-1.65
215	0.71 (1.01)	0.54 (1.01)	-0.36
216	0.18 (0.77)	0.93 (1.07)	-1.63
217	-0.62 (0.76)	0.88 (0.76)	-1.64
218	1.09 (1.01)	0.08 (0.96)	-1.64
219	0.28 (0.96)	0.93 (1.01)	1.65
220	1.09 (1.01)	0.08 (0.96)	-1.64
221	0.71 (1.01)	0.54 (1.01)	-0.36
222	0.71 (1.02)	0.88 (1.01)	-0.53
223	0.28 (0.96)	0.93 (1.01)	1.65
224	0.18 (0.77)	0.93 (1.07)	-1.63
225	-0.62 (0.76)	0.88 (0.76)	-1.64

Note: Since the parameters with response category 3 are not estimated here, they were obtained by computing the sum of parameter estimates with the first two response categories and then reversing the sign of the sum, the associated SDs are not available.