COMPARING THE EFFECTIVENESS OF RESEARCH-BASED CURRICULA FOR TEACHING INTRODUCTORY MECHANICS

By
Trevor I. Smith
B.S. University of Maine, 2005

A THESIS
Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Teaching

The Graduate School
The University of Maine
May, 2007

Advisory Committee:
Michael C. Wittmann, Assistant Professor of Physics, Advisor
John R. Thompson, Assistant Professor of Physics
Donald B. Mountcastle, Associate Professor of Physics
LIBRARY RIGHTS STATEMENT

In presenting this thesis in partial fulfillment of the requirements for an advanced degree at The University of Maine, I agree that the Library shall make it freely available for inspection. I further agree that permission for “fair use” copying of this thesis for scholarly purposes may be granted by the Librarian. It is understood that any copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Signature:  
Date: April 25, 2007
This study examines the effects of implementing various research-based forms of curricula for teaching introductory physics course at the College of DuPage in Glen Ellyn, IL and the University of Maine in Orono, ME. Introductory courses at each of these institutions were modified at each of these institutions over the course of several years. For each course, baseline data were collected during a year in which at least one form of research-based curricula was used. In subsequent years, other research-based curricula were implemented in addition to the baseline treatment at the University of Maine and in place of the baseline treatment at the College of DuPage. Data collected using the Force and Motion Conceptual Evaluation and the Force Concept Inventory were analyzed using the Model Analysis method developed by Bao and Redish[1] to determine the manner(s) in which the curricular modifications affected students’ conceptual development throughout a course. The results of
this analysis show that the majority of curricular changes had little overall effect on students’ conceptual understanding of physics. Several exceptions to this generality are discussed.
ACKNOWLEDGEMENTS

This work was supported in part by the National Science Foundation under CCLI Grant DUE-0510614. I wish to thank Michael Wittmann for his hours of consultation and comment throughout this project. I would also like to thank John Thompson and Don Mountcastle for their helpful comments and support. And lastly, I am indebted to Tom Carter from the College of DuPage for providing data and course information from that institution as well as valuable insight into the corresponding results.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. iii

LIST OF TABLES ........................................................................ viii

LIST OF FIGURES ......................................................................... x

CHAPTER

1  INTRODUCTION ................................................................. 1

2  LITERATURE REVIEW ....................................................... 3
   2.1 Relevant Findings in Educational Psychology ....................... 3
      2.1.1 Prior Knowledge Base .............................................. 4
      2.1.2 Social Learning ..................................................... 7
      2.1.3 Concepts vs. Primitives .......................................... 8
      2.1.4 Epistemologies .................................................... 14
3.2 Categorization of Responses: FMCE

3.2.1 Force Sled Cluster

3.2.2 Vertical Motion Cluster

3.2.3 Force Graphs Cluster

3.2.4 Acceleration Graphs Cluster

3.2.5 Newton III Cluster

3.2.6 Velocity Graphs Cluster

3.2.7 Energy Cluster

3.2.8 Old Clusters vs. New Clusters

4 DATA AND ANALYSIS

4.1 College of DuPage

4.1.1 Algebra-based Course

4.1.2 Calculus-based Course

4.2 University of Maine

4.3 Less Common Models

5 DISCUSSION OF RESULTS

5.1 College of DuPage Algebra-based Course

5.2 College of DuPage Calculus-based Course

5.3 University of Maine

5.4 Comparing Results Between Institutions

6 CONCLUSIONS
<p>| Table 3.1 | Misconceptions and related questions on FCI | 44 |
| Table 3.2 | Bao’s FCI cluster | 45 |
| Table 3.3 | FCI: Question clusters | 48 |
| Table 3.4 | FCI: Newton III cluster | 48 |
| Table 3.5 | FCI: Circular Motion cluster | 50 |
| Table 3.6 | FCI: After Impetus cluster | 52 |
| Table 3.7 | FCI: Constant Force–1D cluster | 53 |
| Table 3.8 | FCI: Constant Force–2D cluster | 55 |
| Table 3.9 | FCI: “Strobe” Motion cluster | 56 |
| Table 3.10 | FCI: Free-fall cluster | 57 |
| Table 3.11 | FMCE: Question Clusters | 59 |
| Table 3.12 | FMCE: Force Sled cluster | 60 |
| Table 3.13 | FMCE: Vertical Motion cluster | 63 |
| Table 3.14 | FMCE: Force Graphs cluster | 65 |
| Table 3.15 | FMCE: Acceleration Graphs cluster | 67 |
| Table 3.16 | FMCE: Newton III cluster | 69 |</p>
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.17</td>
<td>FMCE: Velocity Graphs cluster</td>
<td>72</td>
</tr>
<tr>
<td>3.18</td>
<td>FMCE: Energy cluster</td>
<td>74</td>
</tr>
<tr>
<td>3.19</td>
<td>FMCE: Old Question Clusters</td>
<td>75</td>
</tr>
<tr>
<td>4.1</td>
<td>COD: Changes to algebra-based course</td>
<td>80</td>
</tr>
<tr>
<td>4.2</td>
<td>COD: Changes to calculus-based course</td>
<td>84</td>
</tr>
<tr>
<td>4.3</td>
<td>UMaine: Changes to algebra-based course</td>
<td>89</td>
</tr>
<tr>
<td>4.4</td>
<td>Less Common Models</td>
<td>95</td>
</tr>
<tr>
<td>A.1</td>
<td>FCI—Original vs. Revised</td>
<td>129</td>
</tr>
<tr>
<td>C.1</td>
<td>Sample data from the Force Sled cluster of the FMCE</td>
<td>138</td>
</tr>
<tr>
<td>C.2</td>
<td>Response coding for sample data</td>
<td>139</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 2.1  Example of a model plot  ........................................  40

Figure C.1  Sample Data Model Plot  ........................................  142

Figure D.1  Data: COD, Algebra-based course, Force Sled cluster  ....  143
Figure D.2  Data: COD, Algebra-based course, Vertical Motion cluster  .  144
Figure D.3  Data: COD, Algebra-based course, Force Graphs cluster  .  144
Figure D.4  Data: COD, Algebra-based course, Acceleration Graphs cluster  145
Figure D.5  Data: COD, Algebra-based course, Velocity Graphs cluster  .  145
Figure D.6  Data: COD, Algebra-based course, Newton III cluster (Mass Dependence)  ......................................................  146
Figure D.7  Data: COD, Algebra-based course, Newton III cluster (Action Dependence)  ......................................................  146
Figure D.8  Data: COD, Algebra-based course, Energy cluster  ............  147
Figure E.1  Data: COD, Calculus-based course, After Impetus cluster  ...  148
Figure E.2  Data: COD, Calculus-based course, Newton III cluster (Mass Dependence model)  .............................................  149
Figure E.3 Data: COD, Calculus-based course, Newton III cluster (Action Dependence model)........... 149

Figure E.4 Data: COD, Calculus-based course, Circular Motion cluster (Force in Direction of Motion model)........... 150

Figure E.5 Data: COD, Calculus-based course, Circular Motion cluster (Force Outward from Center model)........... 150

Figure E.6 Data: COD, Calculus-based course, Constant Force–1D cluster 151

Figure E.7 Data: COD, Calculus-based course, Strobe Kinematics cluster 151

Figure E.8 Data: COD, Calculus-based course, Constant Force–2D cluster (Overcoming Impetus model)........... 152

Figure E.9 Data: COD, Calculus-based course, Constant Force–2D cluster (Straight Path model)........... 152

Figure E.10 Data: COD, Calculus-based course, Two Objects in Free-fall cluster ..................... 153

Figure F.1 Data: UMaine, PHY 111, Force Sled cluster ............. 154

Figure F.2 Data: UMaine, PHY 111, Vertical Motion cluster ............. 155

Figure F.3 Data: UMaine, PHY 111, Force Graphs cluster ............. 155

Figure F.4 Data: UMaine, PHY 111, Acceleration Graphs cluster ............. 156

Figure F.5 Data: UMaine, PHY 111, Velocity Graphs cluster ............. 156

Figure F.6 Data: UMaine, PHY 111, Newton III cluster (Mass Dependence) ............. 157
Figure F.7  Data: UMaine, PHY 111, Newton III cluster (Action Dependence) .................................................. 157

Figure F.8  Data: UMaine, PHY 111, Energy cluster ....................... 158
CHAPTER 1

INTRODUCTION

The art of teaching has been a well studied topic throughout the last century. Numerous articles have been published by skilled teachers who have discovered ways of reaching their students, connecting with them in such a way as to allow the students to learn to their hearts’ content. School children from all grade levels have praised their favorite teachers; those individuals who make learning fun and interesting. Teaching has been considered an art best learned in practice. Many student-teachers, after their first day “on the job,” lament that nothing could have prepared them for actually being in a classroom. There is no substitute for being in front of a class.

Yet, as with all great arts, there is a science which forms the foundations of teaching. Researchers in cognitive science have sought to determine how people think. Piaget, Vygotsky, and many others have studied the ways in which people learn and acquire new knowledge and understanding. Over the past three decades, researchers and teachers of physics have begun to apply these findings to studies that aim to determine the ways in which students learn physics. The field of physics education research has grown to include thousands of researchers worldwide who strive to answer two overarching questions: 1) How do students learn physics? and 2) How can we, as instructors, help them to learn physics better?
The results of physics education research, and the growing set of data toward answering these questions, have led to the development of several different types of curricula. Research-based curricula have been developed in many different institutions. The implementation of these curricula has also become a subject of study. Of particular interest is how well these curricula can be implemented at institutions other than those that spawned their creation. This study seeks to investigate how the use of research-based curricula at non-primary implementation sites affect student learning of introductory physics. Furthermore, this study examines the question, how do the implementations of different research-based curricula compare in terms of students’ gains in conceptual understanding of physics?

Data provided from the University of Maine and the College of DuPage are used to investigate this query. Each of these institutions has modified at least one of its introductory physics courses over the span of several years. These modifications include the use of research-based curricula to supplement the lecture, recitation, and/or laboratory portions of a course, as well as the simultaneous implementation of curricula developed by researchers at several different institutions. The results of these modifications are analyzed using pre- and post-instruction multiple-choice surveys designed to elicit students’ mental models concerning physics. The analysis of these survey responses is discussed with an eye toward highlighting the differences caused by each of the curricula.
CHAPTER 2

LITERATURE REVIEW

This chapter provides a brief summary of the research that has been conducted within the field of physics education research. This is not meant to be an exhaustive review by any means, merely a quick overview of the material that is relevant to my research.

2.1 Relevant Findings in Educational Psychology

The human mind has been the subject of study for several hundreds of years. What makes us different from other animals? What does it mean to have the “ability to reason” from a biological standpoint? How do we think and feel? How can we imagine whole other worlds that have properties contradictory to everything that we have seen in our own? These are all questions that psychologists seek to answer through studies of the human mind and the human brain. Educational psychologists focus on the questions dealing particularly with thought, knowledge, and learning, e.g. how do we acquire new knowledge and information and learn how to think about things in new ways?
2.1.1 Prior Knowledge Base

According to Byrnes, Thorndike modeled knowledge as a series of connections between situations and people’s corresponding responses.[2] The acquisition of these correlations is explained by Thorndike’s three laws: the law of exercise, the law of effect, and the law of readiness. The law of exercise states that the frequency with which two events occur together determines the strength of the connection between them, i.e. greater frequency affects a stronger connection and vice versa. The law of effect dictates that the desirability of the result of two subsequent events directly correlates with the strength of the connection between the events. That is, if two subsequent events are followed by a negative result, the connection between those events will weaken, and vice versa. Thorndike’s third law, the law of readiness, directly involves the biological make up of the brain and pathways of neurons which he dubs a “conduction unit.”[2] The law of readiness states that when a conduction unit is active, or ready to conduct, it will be the natural tendency of the conduction unit to, in fact, conduct. If, on the other hand, the unit is not active, its natural tendency will be to refrain from conducting. It is by these conductions that Thorndike explains the biological processes by which the connections between events become stronger or weaker.

Thorndike does not assume students to have a prior knowledge base that will affect their learning. By the law of exercise, teachers should be able to teach their students all they need know by reciting facts or questions and answers. The more often the students hear and speak the correct relationships, the more often they will
be able to answer questions correctly. Furthermore, by the law of effect, affirmations of correct responses to questions and negations of incorrect responses will strengthen the connections between the questions and the corresponding correct responses.

Unfortunately, students are not as easily molded as Thorndike originally believed. Studies throughout the past half century have shown that students begin their schooling with a vast quantity of observational knowledge, i.e. knowledge acquired from their surroundings. Byrnes also reports on Piaget’s proposed two-tier system for the acquisition of knowledge and information: assimilation and accommodation.[2] Assimilation is the process by which humans receive input and interpret it in terms of the ideas they already have about their world. A young child, for example, upon seeing a friend playing with a small blue object that bounces on the sidewalk, may recall his own toy that displays a similar behavior: a red rubber ball. This child will see that the behavioral similarities between the two objects outweigh the cosmetic differences (i.e. color) and determine that his friend is also playing with a rubber ball. This child has assimilated the sensory input provided by the new toy as part of his mental construct of a rubber ball. Accommodation, on the other hand, occurs when new input cannot be easily interpreted in terms of existing ideas, and the person must alter his ways of thinking about the world in order to accommodate it. Let’s imagine that the same child encounters a second friend playing with a different blue toy. This toy does not bounce when you throw it on the sidewalk. It does not squish if you push on it. It isn’t even round. It does have four round pieces on the bottom, and it is curved on the top. When you push it across the sidewalk, it will travel for a long
time. In this case, the child has to take all of these pieces of information and sensory input, and bring them together to create a mental representation of this new kind of toy, a mental construct that his friend will later tell him represents a toy car. This process of the child requiring his mind to alter its own state to accommodate new information depicts Piaget’s second form of learning.

By all accounts, accommodation is more difficult than assimilation. Recognizing and categorizing an object or event is far simpler than creating an entirely new way for your mind to think about its surroundings. Unfortunately for teachers, a significant portion of their work entails guiding students to accommodate information that is not available to them through direct observations of their surroundings. Byrnes correlates teaching with helping students form a wall of knowledge and understanding.[2] Teachers pass their students bricks of knowledge that the students are then expected to place in the wall. A student’s often would not appear to be the work of a master mason. The bricks are an assortment of varying sizes and shapes that are placed together in an attempt toward creating a work of beauty; at best an impenetrable fortress. Assimilation in this wall analogy is represented by simply placing a new brick in the wall. Accommodation, however, occurs when the student is passed a brick that doesn’t fit in the wall. She must rearrange the bricks she’s already laid in order to find a home for this new brick among its brethren.
2.1.2 Social Learning

Vygotsky, like Piaget, recognized that people enter new situations with a wealth of knowledge from their past experiences. Unlike Piaget, Vygotsky believed that people do not acquire new knowledge and information on their own. On the contrary, he emphasized the importance of social interactions for learning.[2] In particular, Vygotsky saw language as the vessel by which knowledge is transmitted. He claims that every step in a child’s development appears twice: first between the child and another person (e.g. a parent or teacher) and second, within the child herself. Vygotsky also theorized the tendency of people to speak with themselves either verbally (egocentric speech) or mentally (inner speech) as the process by which people can think and learn on their own. Egocentric and inner speech, however, are not believed to be as powerful as Piaget’s methods for learning, and most learning with Vygotsky’s methods still requires at least two people.

Another important aspect of Vygotsky’s theory is what is known as the “zone of proximal development.”[2] Vygotsky states that, as a child acquires a new skill, such as learning to read, he will develop from a complete novice requiring extensive help from his teacher or parents, to an expert who requires little or no assistance. At some point within this development, he will be able to perform at an expert level if given a slight boost from his mentors. The gap between his performance with and without this boost is the “zone of proximal development.” With this aspect of his theory, Vygotsky advocates students being taught at a level just above their current expertise. He argues that students who are taught at a level far exceeding
their understanding will not be able to achieve conceptual growth without extensive assistance, but if students are taught at or below their level of achievement, they will not be sufficiently challenged and will never grow intellectually. Revisiting Byrnes’ brick wall metaphor, the teacher’s role in Piaget’s theory of tossing the bricks to the students and allowing them to build their own wall would not satisfy Vygotsky in the least. Vygotsky would have the teacher acting as a scaffolding for the students to be able to build higher walls. The teacher would also be expected to show the student the appropriate region of the wall in which the new brick should be placed but not its precise destination. The teacher can never actually place the brick for the student as Thorndike’s theories suggest; the last step must still be taken by the student himself.

2.1.3 Concepts vs. Primitives

When discussing the ways in which students learn and the ways they think about their world, two well-known schools of thought repeatedly show up: concepts (conceptions[3], conceptual frameworks[4]) and primitives (pieces of knowledge, p-prims[5], facets[6], resources[7, 8]). Concepts are defined as coherent modes of thinking that people use to reason and think about problems and their surroundings.[3] Concepts are the bricks with which student build their walls of understanding. Misconceptions, or concepts that don’t agree with the scientifically accepted manner of thinking about a given topic, are typically defined by the following four properties:

1. They are strongly held, stable cognitive structures.

2. They differ from expert conceptions.
3. They affect, in a fundamental sense, how people understand natural phenomena and scientific explanations.

4. They must be overcome, avoided, or eliminated for people to achieve expert understanding.[9]

Within this list I wish to include:

5. They are content specific.

Item 5 is included for two reasons: 1) context dependence is an important property of misconceptions as described by the majority of the literature, and 2) it is an important distinguisher between misconceptions and primitives, which will be discussed in subsequent paragraphs. Misconceptions come from students observing their surroundings and interpreting their observations to form a coherent view of how the world works. Students use their view of the world to arrange their wall of understanding to incorporate new material. Misconceptions, however, are not the types of bricks that teachers wish their students to use while building their walls. The misconception types of bricks are off-color, and don’t fit as well as they should. Unfortunately, students are often blind to these details and are perfectly satisfied with their choice of bricks and their walls in general. It becomes the task of teachers to make their students realize the imperfections inherent in the misconception bricks. The teachers must also provide new bricks that are very similar to the students’ older versions, but fit the desired niches more precisely.
Making students see that their misconceptions are faulty is not typically an easy task. The stability of misconceptions (item 1 above) and their pervasive nature (item 3) make them very difficult to eradicate. Misconceptions come from students’ own observations and interpretations of the world around them. Observations and interpretations that have served them virtually without fail for many years. Often, getting students to realize that these ideas are incorrect is about as easy as trying to convince them that gravity pulls upward. The mere fact that many of students’ common misconceptions agree with historically held beliefs about the laws of the physical world are a testament to just how deeply ingrained some of these beliefs can be.\(^1\)

Demastes, Good, and Peebles offer several processes by which students may undergo a conceptual change to rid themselves of their misconceptions and adopt the correct, scientifically held view.[10] Among these processes for conceptual change are: incremental, cascade, wholesale, and dual construction. Incremental conceptual change describes the process by which a student gains understanding of a particular concept a little bit at a time. Incremental conceptual change agrees with Vygotsky’s theory pertaining to a young child learning to read with the help of a mentor. The child does not learn how to read “all of a sudden,” he progresses little-by-little until he has mastered the task. The cascade process for conceptual change occurs when a student makes one realization that in turn causes a domino-style effect in which

\(^1\)A more detailed discussion of the similarities between students’ misconceptions and historical views of physics can be found in section 2.2.
several other realizations are made within quick succession. Incremental and cascade change are similar in that each describes conceptual change as a step-by-step process, but unlike incremental change, cascade change is fairly rapid, often occurring within a single class period or study session. Moreover, each step of the process of incremental conceptual change does not necessarily trigger the next step, as is the case with cascade change.[10] Wholesale conceptual change is different from either incremental or cascade change in that students who undergo a wholesale change do not go through intermediate steps between their misconceptions views and the scientifically accurate views. The fourth type of conceptual change described by Demastes et al., dual construction, is again unlike any of the previously described processes. With dual construction, students have adopted the scientifically correct view for some aspects of the situation in which a particular misconception is held, but maintain their misguided views for other aspects. Students who hold dual constructions have effectively gone through a wholesale change for some portions of a given context but still maintain their previous views for the remaining portions.

The processes for conceptual change described by Demastes et al. depict how students can in essence move from a misconception of the world around them to the scientifically accurate belief. The definitions proposed by Demastes et al., furthermore, assume large-grain, context dependent concepts. The examination of smaller pieces of knowledge requires a different approach. Primitives are quite different from misconceptions. As the name suggests, primitives are derived from first (or prime) principles and broad observations of a person’s surroundings. The most prevalent
distinguisher between primitives and misconceptions, as mentioned previously, is the fact that primitives are not context dependent. Primitives are the results of broad observations of the world and cannot be tied to any specific context. As such, primitives are considered neither correct nor incorrect in and of themselves. The validity of a primitive is determined by its implementation within a given context. diSessa defined numerous phenomenological primitives (p-prims) throughout his investigations of students’ understanding of introductory physics.[5] One very prevalent p-prim is the notion that “closer means stronger.” This is a perfect example of inconclusive validity. The statement of this p-prim could be correct or incorrect depending on the context in which it is implemented. Let’s say, for example, that someone was discussing standing near a wood stove or in front of an electronic speaker. In each of these cases, the closer a person gets, the stronger he will feel the effect (heat from the wood stove or volume from the speaker). If, however, this same person was discussing the cause of the earth’s seasons, it would be inappropriate to use the “closer means stronger” p-prim. In fact, it is a documented misconception that many people believe that the earth’s seasons are due to the fact that the earth is closer to the sun in the summer time.\footnote{This misconception could also be considered a dual construction due to the fact that the people who hold the earth-is-closer-in-summer misconception also know perfectly well that summer in the northern hemisphere is winter in the southern hemisphere and vice versa.}[11] In terms of p-prims, however, these people are merely applying the “closer means stronger” p-prim in an inappropriate manner. Moreover, the notion of “closer means stronger” does not require justification within a person’s mind, another important property of p-prims. If a child was asked, “Why do you believe that
closer sources have stronger effects?” This child would most likely assume a puzzled expression and respond, “Because that’s the way things work.”

Facets and resources, like p-prims, are also closely related to primitives. Facets examine how students use various p-prims within specific contexts.[6] For example, applying the p-prim “dying away”—the idea that an effect lessens as time passes until it disappears or dies out—in the context of echoes in a canyon getting softer would be a different facet than applying the same p-prim in terms of a box that had been pushed across a floor slowing to a stop. Minstrell and his collaborators have created a list of facets within the context of physics and organized them based on the subtopic of physics to which they pertain and the difficulty with which they are corrected.[6]

Resources, on the other hand, are a generic form of a piece of knowledge a student uses, whether it be as small and diverse as a p-prim or as large and specific as a concept.[12] Resources, as the name suggests, are merely cognitive structures that students have available to use while thinking about a scenario. Wittmann discusses how students use networks of resources to reason about physics.[8] He goes on to show how these networks of resources can be shown to account for the various types of conceptual change defined by Demastes et al.: incremental, cascade, wholesale, and dual construction. Wittmann also presents evidence on how the use of a particular resource may cause the implementation of a second resource and so on. The links between these resources can be weak or strong, and their strength can vary based on the outcome of a particular pairing of resources. The ideas behind the strength of the links between resources are almost reminiscent of Thorndike’s theories of learn-
ing described in section 2.1.1. But unlike Thorndike, the resources view of student thinking relies heavily on students’ prior knowledge and understanding about their world. Hammer, Elby, Scherr, and Redish elaborate on the strength of links between resources.[7] As the links between a set of resources gets stronger, a student is more and more likely to use all of the resources in a set as soon as one of them is triggered. When the links become sufficiently strong, the set of resources can be used by the student as a single cognitive unit that can be linked to other resources. In this manner, concepts can be viewed as strongly linked sets of resources that incorporate observationally-based primitives and scenario-specific cues to create a single context-dependent cognitive unit.[7]

2.1.4 Epistemologies

In addition to the ways people learn and reason about their surroundings, research has also been conducted to investigate students’ epistemological stances on learning, i.e. how they think about learning.[13–15] Researchers in this area argue that it is insufficient to examine students’ beliefs of their surroundings without considering how the students feel about their responsibilities as learners. Hammer discusses how some students believe that knowledge of physics is a collection of semi-related facts while others hold the “scientist’s” view that physics is a complex, interconnected, and coherent net of concepts.[16] Students also vary in their approach to problem-solving. Some students recognize the need for a strong conceptual background while others believe it is sufficient to memorize formulas and definitions. The Maryland
Physics Expectations Survey[17] (MPEX) and the Epistemological Beliefs Assessment for Physical Science[18] (EBAPS) are two surveys that have been developed at the University of Maryland to measure students’ epistemologies and attitudes toward physics.[14] Both of these surveys present scenarios or viewpoints and ask students to use a Likert scale to indicate how well they agree or disagree with the statement. An example from the MPEX is that students would be presented with the following statement:

“Problem solving” in physics basically means matching problems with facts or equations and then substituting values to get a number.[17]

A student’s response is categorized as favorable if it aligns with a physicist’s opinion and unfavorable if the student disagrees with a physicist’s point of view. In this case, a student who disagrees with the above statement would be considered to have given a favorable response. Research into students’ epistemologies has yielded particularly interesting results when students are asked to complete the survey twice: once the way they truly believe and once the way they think their professors would like them to respond.[19]

Although epistemological beliefs are typically strongly held, Elby[14], Lising and Elby[15], and Hammer and Elby[13] describe how teachers can help their students change their beliefs to accept a scientist’s view of physics. Hammer and Elby[13] advocate a method of getting students to recognize their intuitions for what they are: interpretations of the students’ observations of their world. If a student’s intuitions are incorrect for a particular scenario, the teacher should encourage that student
to examine the observations that led to the misconception in a new light. This method of refining intuitions has sparked the development of the \textit{Maryland Open-source Physics Tutorials} which employ the method of guiding students to reinterpret their observations. A more detailed description of these tutorials can be found in section 2.3.2.

\section{Students’ Mental Models}

Throughout section 2.1, I discussed why educators and curriculum developers cannot consider students to be blank slates. It is imperative that students’ preconceptions about physics be analyzed in detail. With this goal, researchers in physics education have undertaken innumerable studies to examine and classify students’ naive conceptions of physics. Many of these studies have uncovered modes of thinking that are not only coherent within a student’s mind but also applicable to a majority of students who have not engaged in formal studies of physics.\footnote{20} Throughout this thesis these coherent modes of thinking will be referred to as mental models.

Students’ mental models contain several important aspects that should be clarified. First of all, mental models are not necessarily incorrect. Students may use the appropriate Newtonian model for thinking about and considering various aspects of physics. Furthermore, students may create dual constructions\footnote{10} using both correct and incorrect mental models; that is, a student’s use of Newtonian thinking in one scenario does not necessarily indicate that he will use the same correct model for a different aspect of physics. Students may have two or more—possibly conflicting—
mental models for a given situation. Furthermore, a student’s use of one mental model for a given situation does not guarantee its use for a seemingly similar situation. Students often view situations as being unrelated that a physicist would consider identical. Mental models, as they are used in this text, align closely with the literature describing student’s concepts (or conceptions).

Students’ mental models are largely based on their perceptions of their environment. As such, many of these models agree with the ideas of Aristotle and Galileo, but are often less coherent.[21] Students have many years’ worth of experience in a world where friction plays a key role in almost all kinematic processes. The frameworks that students have built regarding how the world works have served them well, and they have no a priori reason for questioning them. The wide applicability and complex structure of these mental models cause students to be very reluctant to give them up. In the following sections I will describe various non-Newtonian mental models that many students use when reasoning about kinematics, dynamics, and energy.

2.2.1 Kinematics

Motion is a deceptively difficult concept to master.[22] Our eyes are designed to be able to readily determine whether an object is moving or still. Furthermore, we can fairly easily recognize whether or not two objects are moving at the same speed. But what about changing speed? While riding in a car, it is rarely difficult to tell whether you are speeding up or slowing down. But what about the car next to you? What about the cars in the opposite lane? It is very difficult to accurately determine
an object’s acceleration without taking appropriate measurements. These difficulties give rise to misunderstandings about the nature of moving bodies with regard to Newtonian physics.

Velocity is among the easiest concepts for introductory physics students. Yet still, many problems arise when students are asked to reason about objects in a physics setting. Trowbridge and McDermott[22] adapted several of Piaget’s motion tasks[23] to elicit students’ mental models regarding velocity. In each of the tasks, two balls were set into motion on two separate tracks. Depending on the task, the balls could be either speeding up, moving at constant speed, or slowing down, and the balls did not move in the same fashion as each other. Students were interviewed individually and asked to answer questions regarding the motion of the balls. When asked to determine if there existed a time during the first demonstration in which both balls were moving at the same speed (or velocity), the majority of students indicated that such an event occurred when the two balls were side-by-side. During a subsequent demonstration in which the balls were never side-by-side, students indicated that the balls were never moving at the same velocity (even though at one instant they were, in fact, moving at the same velocity). Trowbridge and McDermott classify their results as indicating a confusion between velocity and position. This is not to say that students don’t know the difference between where an object is and how fast it’s going. This merely denotes the fact that students may have a tendency to misinterpret their observations. The students can easily see a sense of “sameness” (with regard to position) when the two
bolls are side-by-side. They then project this sameness onto the property of velocity as well.

The concept of acceleration is far more intricate and difficult to comprehend than that of velocity. Without equipment of some sort, it is nearly impossible for an average person to gain much information about an object’s acceleration. With the goal of studying students’ naive conceptions about acceleration, Trowbridge and McDermott[24] once again conducted individual demonstration interviews using modified versions of Piaget’s motion tasks. The results of this study were far more varied than those of their investigation of students’ conceptions of velocity. Once again, students observed two balls moving along two different tracks. The majority of the results from this study come from a demonstration in which the two balls are released at different times, from different locations, have different (constant) accelerations, but end at the same time, at the same position, moving at the same velocity. The students in these interviews are asked to determine whether or not the two balls have the same acceleration (the correct response is “not”). The first result that Trowbridge and McDermott reported is that many students have difficulty distinguishing between velocity and change in velocity.\(^3\) Another common reasoning used by students when reasoning about acceleration is the notion that the ball covering a greater distance in the same amount of time must have a greater acceleration. This idea indicates a confusion between velocity and acceleration similar to that previously found between

\(^3\)This result was originally discovered using the motion task from the studies on velocity in which the two balls pass each other twice. It was observed again in a second task in which the two balls have different accelerations and don’t end in the same position.
position and velocity. Some students believed that the balls had the same acceleration based on the observation that they ended at the same position. Others compared the accelerations based on their finals speeds without regard for the time taken for each to accelerate. Another interesting result reported is the sense that the object that is catching up to (or gaining on) the other must have a greater acceleration. That is, a faster object would have a greater acceleration than a slower one. Students again demonstrate a confusion between velocity and acceleration. This confusion manifests itself in the notion that acceleration is determined by an object’s velocity, highlighting once again students’ difficulties in distinguishing between velocity and change in velocity. These results, as well as those found regarding students’ understanding of velocity, agree with those found during other studies that report students’ poor differentiation between position, velocity, and acceleration.[21]

In addition to studies conducted by Trowbridge and McDermott using demonstration interviews and natural language dialogue, Beichner has investigated the ways in which students interpret kinematics graphs.[25] Beichner’s studies on students’ understanding of kinematics graphs provide many insights into the different forms in which physics concepts may be presented, and the ways in which they manifest themselves through student reasoning. According to Beichner, the two most common “errors” that students make are 1) envisioning the graph as a picture of the situation (rather than a representation of the manner in which position, velocity, or acceleration vary as functions of time) and 2) confusing the slope of the graph at a point with the height of the graph at that point. The latter of these corresponds nicely
with the position-velocity and velocity-acceleration confusion found by Trowbridge and McDermott.\cite{22, 24} Beichner’s results highlight the prevalence of the mental models represented by these confusions. Studies at the University of Colorado at Boulder have also shown that students’ reasoning about and success with physics are intricately related to the manner in which material is presented.\cite{26, 27}

### 2.2.2 Dynamics

What causes motion? What causes changes in motion? When are forces applied? Which objects can exert forces? These are all questions that the typical introductory student wrestles with upon entering the physics classroom. The answers to these are far from obvious for most students. To them, Newton’s laws of motion are a foreign concept that seems to only apply within the contexts of the classroom and the textbook. As with kinematics, students’ mental models of forces and their impact on objects have been widely studied.\cite{20} Halloun and Hestenes\cite{21} detail students’ common sense ideas about physics and their relation to various historical schools of thought (Aristotelian physics and Impetus physics). Unlike these historical models of nature, students’ mental models tend to be fraught with contradictions and unexplainable phenomena. Even so, students share many ideas with these icons of the past. This similarity shouldn’t be surprising given that all of their notions about the world come from direct observation. Within all of these ideologies, forces are divided into two categories: 1) those exerted by direct contact with an active agent, and 2) those that are intrinsic to the object itself (similar to momentum). Using this scheme
for forces, inanimate objects exert no forces. They are merely obstacles that get in the way of an object’s motion. Minstrell discusses students’ ideas of obstacles vs. active agents as they apply to objects at rest.[28] When presented with the situation of a book sitting on a table, many students assert that the table does not exert a force; it is merely “in the way” of the book’s motion. Brown and Clement also examine the “at rest” situation and suggest analogical reasoning as an instructional strategy.[29] The goal of their analogies is to get the students to realize that if a person’s hand exerts a force on a book while supporting it at rest, a table must as well. In addition to the absence of passive forces, long-range forces do not exist in impetus physics without a connection (such as a rope) between the active agent and the object on which the force is being applied. Gravity is seen as the natural tendency of all things to rest on the earth (an intrinsic force), and magnetism is not mentioned. All of these properties are contained in what Halloun and Hestenes deem “the causal property of motion: every motion has a cause” which they consider analogous to Newton’s second law in terms of common sense reasoning.

The idea of impetus is closely related to this notion of causality (and refutes Newton’s first law). Halloun and Hestenes[21] report students’ beliefs that active agents impart an impetus (force) to an object which keeps the object moving once the active agent is removed, along with a very eloquent quote from impetus physicist Jean Buridean:

A mover, while moving a body, *impresses* on it a certain *impetus*, a certain power capable of moving this body in the direction in which the mover set
it going, whether upwards, downwards, sideways, or in a circle. By the same amount that the mover moves the same body swiftly, by that amount is the impetus that is impressed on it powerful. It is by this impetus that the stone is moved after the thrower ceases to move it; but because of the resistance of the air and the gravity of the stone, which inclines it to move in a direction opposite to that towards which the impetus tends to move it, this impetus is continually weakened. Therefore the movement of the stone will become continually slower, and at length, the impetus is so diminished or destroyed that the gravity of the stone prevails over it and moves the stone down toward its natural place.[30]

The statement regarding the “gravity of the stone” is of particular interest as it clearly implies that gravity is not a force that is exerted on the stone but rather an intrinsic property of the stone.

Halloun and Hestenes also report on students’ ideas of the explicit relation between force, velocity, and acceleration. Many students believe that a constant applied force results in a constant velocity. This notion is not surprising living in a world in which friction almost always applies a force opposing the motion of an object. It follows directly that acceleration is the result of applying an increasing force.\(^4\) Also mentioned are the concepts of dominance and compromise for two competing forces. Dominance is the tendency for objects to move in the direction of the greater force.

\(^4\)I prefer the wording that a change in velocity (or increasing velocity) results from applying an increasing force given students’ confusions between acceleration, velocity, and change in velocity mentioned previously.
and ignore the weaker. Compromise (not surprisingly) is the tendency for objects to move in a direction somewhere in between the two forces. At first, the idea of compromise seems very close to the correct idea of superposition. Students’ notions of compromise, however, are often intricate and involve one force wearing out while the other takes over.

Thornton also reports various student views for analyzing physical situations.[31] Thornton used results from administering the Force and Motion Conceptual Evaluation[32] as well as individual interviews to identify these student views and the frequency with which they occur. The first case reported is that of an object moving with constant, nonzero velocity.\(^5\) The most common student view for this situation is that a constant velocity requires a constant, nonzero force in the direction of motion. A secondary student view is that an ever decreasing force will be present in the direction of motion. The next case involves an object speeding up at a steady rate. The most common student view (and the only non-Newtonian view reported) for this situation is that a force is in the direction of motion that is increasing at a constant rate. Both of these most common student views agree with the aforementioned concept that the force on an object is determined by its velocity (or force follows velocity). The third case, in which an object is slowing down at a steady rate, gives rise to several different student views. The most common again agrees with the force-follows-velocity concept and states that a force exists in the direction of motion that decreases at a

\(^5\)The first case is actually an object at rest, but this case is trivial and > 96% of students responded correctly.
constant rate. This view is seen less frequently than either of the other most common views. Other student views include: a decreasing force opposed to the motion, and increasing force opposed to the motion and no net force. This last student view is particularly interesting as it corresponds with Aristotle’s ideas of all things wanting to go to their natural state. If an object is left to its own devices it will eventually slow down to rest on the earth. One possible explanation of this view is that students are applying their knowledge of friction (even though they are explicitly told that the object is in a frictionless environment), but it is still troubling that these students do not consider friction as contributing to the net force on an object. This again leads to Aristotelian and impetus physics with the idea of intrinsic forces of bodies.

Newton’s third law is particularly difficult for students in that it has no common sense analog as Halloun and Hestenes would describe it.[21] Studies have shown that students use a variety of strategies when reasoning about the forces of two interacting bodies.[33–35] Bao, Hogg, and Zollman identified four “contextual features” that students use when responding to questions regarding Newton III: velocity, mass, pushing and acceleration.[33] The velocity feature indicates that an object with a larger initial velocity will exert more force than an object with a smaller initial velocity (during a collision). The mass feature is expressed in the idea that more massive objects exert more force than less massive ones. The velocity and mass features work well together to illuminate students’ implicit confusion between momentum and veloc-

---

635% of responses as opposed to 74% and 76% for the constant velocity and speeding up cases, respectively.
ity. Based on their ideas about kinematics, students often have a desire to represent force as \( F = mv \);[21] furthermore, students may use the terms momentum and force interchangeably.[33] The pushing feature is contained in the notion that, when one object pushes another, the object that is pushing must exert more force than the object that is being pushed. This idea is typically accompanied by the reasoning that if both objects exerted the same force on each other, neither would move. The pushing feature ties in with the concept of “overcoming” an object’s inertia and the idea of intrinsic forces mentioned previously. The acceleration feature is very similar to the velocity feature in that students display a belief that objects with greater acceleration exert more force than those with less (or zero) acceleration (during a collision). I am wary about making this distinction between the velocity and acceleration contextual features. It seems as though the students who use the acceleration feature are combining their desire to use the velocity feature with an attempt to apply Newton’s second law (\( \sum F = ma \)). It appears that these students have learned in class that force is related to acceleration, not velocity, but they still have the idea that more acceleration corresponds to more exerted force. These students also may be failing to make the distinction between forces exerted by two objects on each other and net force on a given object.

Smith and Wittmann also describe two mental models that students use when reasoning about Newton III: mass dependence and action dependence.[34] These mental models correspond with Minstrell’s facets of knowledge,[6] in particular facets 62 (The moving object or a faster moving object exerts a greater force), 63 (The more active or
energetic object exerts more force), and 64 (The bigger or heavier object exerts more force). The mass dependence facet has a direct correlation with the mass contextual feature described by Bao, et al.[33] The action dependence facet combines the velocity and pushing contextual features described above to create a mental model that is applicable to both pushing and collision situations. Smith and Wittmann found that students applied the action dependence facet more frequently than they did the mass dependence facet indicating students’ preference toward reasoning based on an object’s motion. The identities of these facets as well as their relative prevalence are consistent with Maloney’s findings on students’ use of a dominance principle to reason about the interaction between two bodies.[35] Thornton also reports students using ideas about an object’s “ability to push” or “do damage” using its mass or speed to determine the force it will exert on another object.[31] Interestingly, Thornton (like Bao, et al.[33]) explicitly separates his analysis of questions that consider collisions from those that consider pushing situations; however, this separation does not give rise to additional student views. Instead, Thornton uses the ideas of dominance due to initial speed or mass to make connections between the two types of situations while highlighting their differences and the range of the applicability of the students’ views.[31]

Epistemological research conducted by Hammer and Elby has indicated the high probability that students’ use of the aforementioned facets and contextual features are based on their observations that smaller or slower objects “react” more during

---

7Definitions of these facets can be found in Refs.[6, 34].
collisions than bigger faster ones.[13] These observations come from students’ prior knowledge and experience and are directly related to acceleration rather than force. Elby has created instructional materials to help students make this distinction and appropriately interpret their observation. A more detailed description of these materials will follow in section 2.3.2.

2.2.3 Energy

Lawson and McDermott report on student reasoning regarding the work-energy and impulse-momentum theorems.[36] Students were presented with a situation in which two objects of differing mass are placed on a frictionless surface and subject to the same constant force while traversing a given distance. The students were asked to first compare the momentum of each of the objects at the end of its trip and secondly to compare the kinetic energy of the objects at the end. The results of this experiment report that most students are comfortable using the definitions of kinetic energy and momentum learned in class ($KE = \frac{1}{2}mv^2$ and $p = mv$, respectively), but few use these definitions in conjunction with either the work-energy theorem ($\Delta KE = F\Delta x$) or the impulse-momentum theorem ($\Delta p = F\Delta t$) even though, when questioned explicitly, all students were familiar with these theorems.[36] Instead, most students maintained an almost exclusive dependence on the final velocity of each of the objects to make their comparisons. The constant force used throughout the study makes the aforementioned theorems imperative to students’ success in comparing the objects’ momenta and kinetic energy. O’Brien Pride, Vokos, and McDermott found
nearly identical results when studying students’ use of the work-energy and impulse-momentum theorems for two-dimensional motion.[37]

The results of the studies conducted by Lawson and McDermott, and O’Brien Pride, Vokos, and McDermott can be applied to the ways in which students reason about the velocity and kinetic energy of an object sliding down a frictionless hill. Of particular interest is which factor students will consider most important: the hill’s height or its steepness. Using the work-energy theorem, it is easy to determine that a hill’s height is the deciding factor in determining an object’s change in velocity or kinetic energy from the top of a hill to the bottom. Students, however, have a wealth of experience in which they have personally observed the fact that steeper hills allow objects to go faster. Moreover, objects take less time to go down steep hills than they do to go down shallow hills. The assertion that steeper hills cause a greater change in velocity, however, represents a confusion between the vertical component of an object’s velocity and the scalar magnitude of its velocity. This confusion is similar to that reported by O’Brien Pride, Vokos, and McDermott in which students only consider one component of an object’s change in momentum rather than its scalar magnitude. Another possible interpretation is that students do not consider the entire time during which the sled is accelerating. The students may be recognizing the fact that the acceleration of a sled increases as the steepness of the hill increases. This increased acceleration corresponds with an increased change in velocity per unit time, but not an increased overall change in velocity as is indicated by students’ responses.
2.3 Instructional Strategies

Research regarding students mental models is incredibly important to conscientious teachers, but these results are only the first step. In order to affect students’ understanding of introductory physics, results from these studies must be used to develop instruction strategies that will target students’ non-Newtonian mental models and help the students gain a better understanding of their surroundings. Luckily, many physics education researchers agree with this statement, as evidenced by the volume of reports on curriculum development.[13, 20, 38–41] These physicists have endeavored to combine their knowledge of students’ mental models with the findings of psychological research concerning the ways in which people learn, remember, and understand to create instructional materials and strategies that enhance physics instruction and allow students a more profound understanding of their world.

2.3.1 Interactive Lecture Demonstrations

Traditional physics instruction at the university level consists of three parts: lecture, recitation, and laboratory. One of the first steps toward improving this physics instruction is to modify the lectures to create a more engaging environment. Traditional lectures consist of a professor using a variety of presentation strategies including overhead transparencies, chalkboard formulas and computerized presentations to deliver information about physics. Demonstrations are often used to peak students’ interest for a given topic, but are the students really engaged in the physics they are witnessing? Thornton and Sokoloff argue that they are not.[40] As part of the Tools for
Scientific Teaching project, Thornton and Sokoloff have developed a series of Interactive Lecture Demonstrations (ILDs) that engage the students on a much deeper level than that of a traditional lecture. The basic format of an ILD is that the students are first presented with the apparatus with which the demonstration will be conducted and asked to make predictions regarding the outcome. Using a large-group response system, students are asked to input their predictions by responding to a multiple-choice question. These response systems allow a histogram of the class’s predictions to be created and displayed instantly so that the students can see trends and compare their responses with those of their classmates. Students discuss their predictions with their neighbors to compare the reasons for their decisions. The demonstration is then conducted and the students once again converse with their neighbors about the result of the demonstration in an attempt to reconcile their observations with their initial predictions. The idea behind ILDs is that students will be more mentally engaged in the presentation than they would be during a more traditional demonstration in which predictions are not required.

Studies have shown that requiring students to think about their predictions before seeing a demonstration has a profound effect on the students’ understanding of the physical situation in question. Furthermore, research has also shown that students who have participated in ILDs for kinematics also perform better on assessments regarding their knowledge of dynamics than students who have not received any ILD instruction. Thornton and Sokoloff feel that this carry over is a result of ILD students acquiring the tools to think logically and in depth about phys-
ical phenomena regardless of the specific knowledge acquired by participating in the ILDs. The claims have been made on students’ conceptual understanding of physics by comparing students who participated in ILDs with students who had received all traditional instruction.

2.3.2 Tutorials

A second step toward improving physics instruction at the university level would be to modify the recitation periods to promote the goal of cognitive development rather than improving mathematical problem-solving skills. This step has in fact become a leap toward improving students’ conceptual understanding through the implementing of tutorials, first developed at the University of Washington (UW).[38, 39, 43] Tutorials are worksheets designed to be completed in small groups (3–4 students) that implement guided inquiry methods to encourage students to discover various physical principles for themselves. Instructors in tutorial sections act as facilitators, asking questions of the students to probe their understanding and reveal misconceptions that students possess. McDermott and Shaffer, creators of UW’s Tutorials in Introductory Physics (TIP)[43] subscribe to the conceptual framework view of physics knowledge and address students’ misconceptions through a three step process dubbed “elicit, confront, and resolve.”[39] The first step in this process calls for students misconceptions to be brought to the front of their minds by presenting a situation in which many students will implement a particular misconception. Once these misconceptions have been elicited, the tutorial encourages the students to confront these misconcep-
tions by asking questions designed to cause cognitive conflict. These questions are typically presented in the middle of each of the tutorials and may be elaborated on by instructors if sufficient cognitive conflict does not arise. Once the students begin to question the validity of their ideas about the situation at hand, the tutorial provides them the opportunity to resolve their conflict and, in doing so, develop a Newtonian understanding of the physical situation.

The TIP materials are based on another curriculum from UW: *Physics by Inquiry* (PBI).[44] The PBI materials compose an entire curriculum that uses laboratory work and explorations as its key method for instruction. The PBI curriculum is intended to stand on its own to provide a course that focuses on developing students’ conceptual knowledge of physics with minimal focus on mathematical problem solving. Instruction using PBI is very student-centered and progresses only as quickly as the students can properly work through the activities and explorations. As such, PBI is ill-suited for typical university instruction that demands fairly strict topical guidelines for each semester. The TIP materials were developed for just this purpose. The tutorials act as a substitute for traditional recitations that will accomplish many of the same cognitive goals as PBI without requiring additional time or course restructuring.

Researchers at the University of Maryland College Park (UMD) have also developed many tutorials for introductory physics.[45, 46] These tutorials, however, employ very different instructional methods than those encountered in the UW varieties. The *Activity-Based Tutorials* (ABT)[45] use hands-on experimentation and modern, computer-based data taking techniques to allow the students to demonstrate
the principles of physics for themselves. Many of these activities are based on the *Tools for Scientific Thinking* materials created by Thornton and Sokoloff discussed in section 2.3.1. Students working on ABT materials typically use force and motion sensors to take data while encountering various physical phenomena.

The *Maryland Open-source Physics Tutorials* (OPT) materials employ yet another instructional strategy. Based on the work of Hammer and Elby, these tutorials use an epistemological basis of confronting students’ intuitions and beliefs about their surroundings. The OPT materials treat students’ intuitions about their world as valuable observations that may have been misinterpreted by the students. This attitude toward students’ observations is very similar to diSessa’s notion of p-prims as pieces of atomistic, undeniable knowledge. An example from the OPT tutorial for Newton’s third law begins by asking students about a situation in which a heavy truck collides with a small car. The majority of students immediately respond that the truck exerts more force because the car “reacts” more. The OPT tutorial guides the students through the process of recognizing that their intuitions and practical observations apply to the difference in acceleration of each of the vehicles, not necessarily the force. The students use Newton’s second law \((F = ma)\) to determine that the mass difference will offset the difference in acceleration and that the two objects will exert the same amount of force on each other. The OPT materials are by far the least well known of those mentioned, but they are growing in popularity. Moreover, research suggests that the OPT tutorials could be more effective than either the ABT or the TIP versions.
2.3.3 Modeling Method for Laboratory Work

Research-based reform has also touched on the third element of the triad of typical instruction: laboratory sessions. Wells, Hestenes, and Swackhamer have developed modeling methods to be implemented in the laboratory portion of the typical physics course.[41] Traditional laboratory work consists of students verifying physical principles through the use of step-by-step labs that tell them every procedure to complete and every measurement to take. The modeling method, on the other hand, more closely resembles the work done by scientists when examining unexplained phenomena. Students are first presented with an apparatus (e.g. a simple pendulum of varying mass and length) to be examined and used in their experiments. Working in small groups (again 3–4 students) the students decide which parts of the system they will explore and which variables will be important to consider. The students design and execute their own experiments and present their results to the rest of the class. Each group of students must defend their results to the satisfaction of their instructor as well as their classmates. Once all results have been presented, the students work to create a model for their physical system that completely describes the observed phenomena and is generalizable as an accepted physical principle. This procedure affords students the opportunity to gain experience in experimental design and presentation as well as obtain a deeper understanding of the physics behind each situation. Wells, Hestenes, and Swackhamer[41] and Halloun and Hestenes[47] provide data that these modeling laboratory methods improve student understanding of physics over courses
taught in a more traditional fashion, even when the instructor deliberately tries to emphasize conceptual understanding over problem solving.[47]

2.4 Methods for Analyzing Data

2.4.1 Interviews vs. Multiple-Choice Tests

Many methods have been developed for collecting and analyzing data with regard to students’ understanding of physics. diSessa’s methods for conducting an individual “clinical interview” provides the opportunity to study the depth of a student’s knowledge and determine how much they understand compared with how much they can merely regurgitate for the purposes of examinations.[48] Section 2.2.1 describes other interview-based studies conducted at the University of Washington to determine students’ ideas concerning velocity and acceleration. This qualitative method is excellent for discovering the ways in which individual students think and reason about physics and the world around them. Unfortunately, diSessa’s interviewing methods become cumbersome and vastly time consuming when working to determine if a given curriculum or instructional strategy works well for a broad range of students. With this goal in mind many researchers rely on quantitative methods for determining students’ understanding.\footnote{There are too many instances of this practice to list them all here. Some examples can be found in Refs. [34, 42, 49].} Furthermore, the results of interview data have been used to create multiple choice assessments that focus on eliciting students’ correct and incorrect mental models for various situations. The Force Concept Inventory (FCI)[50] and the

\[ \text{Force Concept Inventory (FCI)} \]
Force and Motion Conceptual Evaluation (FMCE)[32] are two such assessments that have become very popular throughout the physics education research community as relatively quick methods for determining students’ difficulties in mechanics. The FCI tests students’ understanding of one- and two-dimensional linear and rotational dynamics, while the FMCE focuses more intently on students’ understanding of linear kinematics and dynamics.⁹

2.4.2 Normalized Gain

In his 6000-student study, Hake developed a method for analyzing students’ FCI scores that he dubbed the normalized gain.[49] The normalized gain (or \(< g >\)) is a measure of a student’s improvement throughout a semester as a fraction of their capacity for improvement measured at the beginning of the semester. As an example, suppose there are five questions in a particular section of the FCI. Suppose also that a student answered one of them correctly at the beginning of the semester and three at the end. This student has improved by two questions out of a possible four and her normalized gain would be given by:

\[
< g > = \frac{4 - 2}{4} = 0.50
\]

The maximum normalized gain a student can obtain is 1.00. Hake’s method of normalized gain is very useful for comparing students at varying initial levels of understanding. If researchers were to consider only the raw increase in score, a student who went from responding correctly on one question to responding correctly on two would

⁹More detailed analyses of these tests can be found in Chapter 3.
be scored identically to a student who went from four to five (out of five) correct responses; however, the normalized gain scores for these students would be $<g> = 0.25$ and $<g> = 1.00$, respectively. The normalized gain in this case serves as a scale of difficulty. It is more difficult for students to gain the top 5% of understanding in a class than it is to achieve a similar gain in score at the lower range of understanding.

Normalized gain also provides far more information than looking at test scores alone. Let’s consider, for a moment, two more students. Student A received a 90% on his end-of-semester test, and Student B only scored a 70% on hers. At first glance it would appear that Student A performed far better in the class than Student B. If, however, we go back to their pre-semester assessments, we will see that Student A scored an 85% and Student B scored a 40%. In this case, Student B would have a normalized gain of $<g> = 0.50$ and her accomplishments above Student A’s $<g> = 0.33$ could be recognized.

### 2.4.3 Model Analysis

One flaw in Hake’s analysis is that he ignores much of the rich information that these well-crafted assessments contain regarding students’ reasoning and understanding. His normalized gain only considers whether or not students’ responses correspond with proper Newtonian thinking. This binary analysis disregards any information that conveys which (if any) mental model a student might be using. The method of model analysis for examining multiple choice assessment data, developed by Bao and Redish, provides a more in depth look at students’ understanding than this
typical binary analysis.[1] Using model analysis, each available response for a multiple-choice question is categorized in one of three ways: 1) correct Newtonian thinking, 2) corresponding with a well known and documented incorrect student mental model, or 3) not corresponding with any previously observed mental model, “other.” Mental models used for categorizing the available responses are specific to the particular branch of physics that the question aims to assess (i.e. linear kinematics, rotational dynamics, Newton’s third law, etc.). Each student’s responses are classified using this categorization scheme to determine the frequency of each well-defined model throughout a class of students. Data are also kept regarding the consistency with which each student uses each of the mental models to determine whether he is in a “pure model state” (employs only one mental model for answering all questions within a given domain) or a “mixed model state” (uses two or more models while answering questions in a given domain).[1] Once each student’s responses are categorized, the results for the entire class are combined into a density matrix that represents the knowledge state of the class as a whole. The eigenvalues and eigenvectors of this density matrix are analyzed to create a model plot which visually depicts the class’s improvement over the course of a semester. An example of a model plot can be seen in Figure 2.1.

The model 1 (or model 2) region indicates a high probability that students consistently use model 1 (model 2) when answering questions related to a particular aspect of physics. The mixed model region indicates that the class as a whole is in a “mixed model state.” This generally occurs when some students mostly employ
Figure 2.1: This figure depicts the various regions of the model plot and the information that can be ascertained therein for a set of questions that provide responses that can be classified as either correct (model 1) or corresponding to a single incorrect mental model (model 2).[1]

model 1 and some students mostly employ model 2. A “mixed model class” could also be populated by many students who are in mixed model states themselves. The distinction between these population can be observed by the proximity of the class’s data to the “probability = 1” barrier line displayed in Figure 2.1. Data close to the barrier indicate a class populated by students who individually possess pure model states. A class of students who primarily posses mixed model states, however, would be indicated by a data point that is close to the origin.

Classes typically begin the semester within the model 2 region, and ideally they move toward the model 1 region (visually up and to the left) as the semester progresses. These model plots provide a concise visual representation of the simultaneous process of students using an incorrect mental model less frequently and beginning to
use the correct model more frequently. Bao and Redish posit that students who are in a “mixed model state” are more likely to move to a pure Newtonian model state than those students that display properties of being in an incorrect “pure model state.”[1] Thornton’s research into student views also supports the assertion that students progress gradually from one mental model to the next, going through mixed states before consistently applying Newtonian reasoning.[31] Though not yet very popular, Bao and Redish’s method of model analysis provides a much needed step in the direction of analyzing multiple-choice assessments based on the implications of all given responses on students’ reasoning rather than focusing solely on whether or not each response corresponds with the appropriate form of Newtonian thinking.
CHAPTER 3

EXPERIMENTAL METHODOLOGY AND DESIGN

This thesis analyzes data from the Force Concept Inventory and the Force and Motion Conceptual Evaluation using model analysis developed by Bao and Redish. To implement model analysis it is necessary to separate each of these assessments into clusters of questions by the specific type of physics knowledge assessed by each question and identifying the mental models that students may use to choose their responses. The following sections define the clusters of questions that will be analyzed and detail the mental models contained within each.

3.1 Categorization of Responses: FCI

A detailed record of the development of the original version of the FCI can be found in ref. [50]. A table showing the correlation between the questions in the original (1992) version of the FCI and those in the revised (1995) version can be found in Appendix A. For the remainder of this discussion I will refer to the questions as they occur in the revised version. One of the beautiful things about the FCI is the breadth of physical understanding of mechanics that is necessary to obtain a high score. Each item on the test asks a question that pertains to a different bit of conceptual understanding than any of the other items. Furthermore, a deep understanding of the concepts is
required to correctly answer any one of the questions. This poses a problem, however, when one endeavors to identify clusters of questions that pertain to similar aspects of physics. Table 3.1 shows the vast number of misconceptions the original authors of the FCI intended it to identify as well as their corresponding questions.

Many of the misconceptions listed in Table 3.1 correspond with the mental models described in section 2.2. For example, misconceptions “AR1” and “AR2” correspond directly with the mass dependence and action dependence mental models identified by Smith and Wittmann.[34] It can be seen from Table 3.1 that questions on the FCI do not separate nicely into concise clusters. As such, it is up to researchers using Bao’s model analysis to define their own clusters. Bao and Redish chose several questions from the FCI as an example of their model analysis methods. In particular they chose a set of questions in which the alternative model was the notion that there always exists a force in the direction of motion.[1] Table 3.2 shows the question cluster Bao used for their example.

Unfortunately I do not agree with the cluster Bao and Redish use as their example. The first reason for this is that the two models they define are not mutually exclusive. In question 17 for example, there is in fact a force in the direction of motion, and response “b” indicates this fact just as clearly as responses “a” and “d.” The difference between these responses lies not in their direction but in their relative magnitudes. Response “b” indicates (correctly) that the net force on the elevator must be zero for it to move at a constant velocity, while responses “a” and “d” implicitly convey the notion of a nonzero net force. This distinction has nothing to do with the presence
Table 3.1: Misconceptions and related questions on the FCI: recreated from ref. [50].

Question numbers have been modified to correspond with the revised FCI.

<table>
<thead>
<tr>
<th>Misconception</th>
<th>Questions Number(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>0. Kinematics</strong></td>
<td></td>
</tr>
<tr>
<td>K1. position-velocity undiscriminated</td>
<td>19</td>
</tr>
<tr>
<td>K2. velocity-acceleration undiscriminated</td>
<td>19, 20</td>
</tr>
<tr>
<td>K3. nonvectorial velocity composition</td>
<td>9</td>
</tr>
<tr>
<td><strong>1. Impetus</strong></td>
<td></td>
</tr>
<tr>
<td>I1. impetus supplied by “hit”</td>
<td>11, 27, 30</td>
</tr>
<tr>
<td>I2. loss/recovery of original impetus</td>
<td>7, 8, 21, 23</td>
</tr>
<tr>
<td>I3. impetus dissipation</td>
<td>10, 12, 13, 14, 15, 24, 27</td>
</tr>
<tr>
<td>I4. gradual/delayed impetus build-up</td>
<td>8, 10, 21, 27</td>
</tr>
<tr>
<td>I5. circular impetus</td>
<td>6, 7</td>
</tr>
<tr>
<td><strong>2. Active Force</strong></td>
<td></td>
</tr>
<tr>
<td>AF1. only active agents exert forces</td>
<td>15, 16, 17, 28, 29, 30</td>
</tr>
<tr>
<td>AF2. motion implies active force</td>
<td>27</td>
</tr>
<tr>
<td>AF3. no motion implies no active force</td>
<td>29</td>
</tr>
<tr>
<td>AF4. velocity proportional to applied force</td>
<td>22, 25, 26</td>
</tr>
<tr>
<td>AF5. acceleration implies increasing force</td>
<td>3</td>
</tr>
<tr>
<td>AF6. force causes acceleration to terminal velocity</td>
<td>3, 22</td>
</tr>
<tr>
<td>AF7. active force wears out</td>
<td>22</td>
</tr>
<tr>
<td><strong>3. Action/Reaction Pairs</strong></td>
<td></td>
</tr>
<tr>
<td>AR1. greater mass implies greater force</td>
<td>4, 15, 16, 28</td>
</tr>
<tr>
<td>AR2. most active agent produces greatest force</td>
<td>15, 16, 28</td>
</tr>
<tr>
<td><strong>4. Concatenation of Influences</strong></td>
<td></td>
</tr>
<tr>
<td>CI1. largest force determines motion</td>
<td>17</td>
</tr>
<tr>
<td>CI2. force compromise determines motion</td>
<td>6, 7, 12, 14, 21</td>
</tr>
<tr>
<td>CI3. last force to act determines motion</td>
<td>8, 9, 21, 23</td>
</tr>
<tr>
<td><strong>5. Other Influences on Motion</strong></td>
<td></td>
</tr>
<tr>
<td>CF. Centrifugal force</td>
<td>6, 7</td>
</tr>
<tr>
<td>Ob. Obstacles exert no force</td>
<td>4, 11, 15, 16, 29</td>
</tr>
<tr>
<td><strong>Resistance</strong></td>
<td></td>
</tr>
<tr>
<td>R1. mass makes things stop</td>
<td>14, 27</td>
</tr>
<tr>
<td>R2. motion when force overcomes resistance</td>
<td>25, 26</td>
</tr>
<tr>
<td>R3. resistance opposes force/impetus</td>
<td>25, 26</td>
</tr>
<tr>
<td><strong>Gravitation</strong></td>
<td></td>
</tr>
<tr>
<td>G1. air pressure-assisted gravity</td>
<td>3, 11, 17, 29</td>
</tr>
<tr>
<td>G2. gravity intrinsic to mass</td>
<td>3, 11, 13</td>
</tr>
<tr>
<td>G3. heavier objects fall faster</td>
<td>1, 2</td>
</tr>
<tr>
<td>G4. gravity increases as objects fall</td>
<td>3, 13</td>
</tr>
<tr>
<td>G5. gravity acts after impetus wears down</td>
<td>12, 13, 14</td>
</tr>
</tbody>
</table>
Table 3.2: Bao and Redish’s example cluster. Question numbers have been modified to correspond with the revised FCI. Bao’s definition does not completely fit question 26 as original question 28 divides into revised questions 25 and 26.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Force Proportional to Acceleration</th>
<th>Force in Direction of Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>a, d</td>
<td>b, c</td>
</tr>
<tr>
<td>13</td>
<td>d</td>
<td>a, b, c</td>
</tr>
<tr>
<td>17</td>
<td>b</td>
<td>a, d</td>
</tr>
<tr>
<td>25</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>26</td>
<td>*</td>
<td>a</td>
</tr>
<tr>
<td>30</td>
<td>a, c</td>
<td>b, d, e</td>
</tr>
</tbody>
</table>

or absence of a force in the direction of motion. Questions 25 and 26 also share this disparity. In each of these, there is a force in the direction of motion. Question 25 asks the students to compare the magnitude of the force with which a woman pushes a box with that of the weight of the box and the forces resisting the box’s motion.¹ Question 26 asks the students to consider how the velocity of the box would be different if the woman were to push with twice as much force. This question requires knowledge of the relationship between frictional force and the weight of an object as well as understanding of how net force relates to velocity, but it does not test a student’s understanding of the presence of a force in the direction of motion.

Even question 28 on the original version of the FCI does not explicitly involve a force

---

¹This question is essentially a horizontal version of question 17.
in the direction of motion.$^2$

Questions 11, 13, and 30 fit more appropriately into these models. Question 11 asks students to reason about the forces on a hockey puck after it has been given a kick across a frictionless ice surface. The correct answer is that the only forces acting on the puck are a gravitational force down and a normal force up so that the net force is zero. Question 13 provides a scenario in which a boy throws a steel ball upward. This question is designed to get students to respond that the only forces acting on the ball are a gravitational force down and an air resistance force opposing the motion. Question 30 involves a tennis ball that has been hit across a court.$^3$ Students’ responses are considered correct if they indicate that only a gravitational force and a force due to air resistance are acting on the ball. In each of these three questions, students are considered incorrect if they indicate the lingering presence of the active agent (the “kick,” hand, and racquet or “hit,” respectively) by stating that a force exists in the direction of motion. I agree with Bao’s grouping in the sense that a force in the direction of motion could correspond with the notion that force is proportional to velocity or force is needed for motion to occur, but I do not feel that this cluster is justified for the model of force in the direction of motion.

With that said, I have divided the FCI into seven clusters of questions: Newton III, Circular Motion, After Impetus, Constant Force–1D, Constant Force–2D, “Strobe” Kinematics, and Two Objects in Free-fall. Table 3.3 shows how each question on the FCI has been categorized into one of these clusters, making the clusters mutually

$^2$See Appendix A.

$^3$The original version of the question used a golf ball.
exclusive. The question clusters on the FCI have been designed to be mutually exclusive to ensure that each question is given equal importance during data analysis. Examining Table 3.1, we can see that some questions (e.g. 3, 15, and 16) may elicit many incorrect mental models while others (e.g. 1, 2, and 20) only elicit one. This being the case, clusters based solely on these mental model categories would distribute questions unevenly. Some questions would be analyzed up to five times while others would only be analyzed once. For the initial attempt at a model analysis of the entire FCI, I decided to limit each question to a single cluster corresponding to one or two common incorrect mental models, thus weighting each question equally. In our analysis, question clusters were created considering the content and nature of each question as well as the various mental models that may be indicated by each possible response. Each question within a given cluster assesses similar content knowledge and elicits similar mental models. The following sections contain detailed explanations pertaining to the definitions and formulations of these clusters.

### 3.1.1 Newton’s Third Law Cluster

The Newton’s third law cluster (Newton III) of the FCI contains four questions that test students’ knowledge and understanding of Newton’s third law. All four of these questions can be found within the “Action/Reaction Pairs” section of Table 3.1. One of the questions (question 4) depicts a situation in which two objects collide and interact for a brief period of time. Two of the questions (questions 15 and 16) ask the students to reason about the forces exerted when one object pushes another object
Table 3.3: Clusters of questions on the FCI.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Question Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newton III</td>
<td>4, 15, 16, 28</td>
</tr>
<tr>
<td>Circular Motion</td>
<td>5, 6, 7, 18</td>
</tr>
<tr>
<td>After Impetus</td>
<td>8, 9, 10, 11, 13, 23, 24, 25</td>
</tr>
<tr>
<td>Constant Force–1D</td>
<td>3, 17, 25, 26, 27, 29</td>
</tr>
<tr>
<td>Constant Force–2D</td>
<td>12, 14, 21, 22</td>
</tr>
<tr>
<td>“Strobe” Kinematics</td>
<td>19, 20</td>
</tr>
<tr>
<td>Two Objects in Free-fall</td>
<td>1, 2</td>
</tr>
</tbody>
</table>

and the two are in contact for an extended period of time. Question 28 can be considered a pushing question in that one boy is pushing off of the other, but the two boys do not stay in contact for extended period of time, so this question is a bit different from questions 15 and 16. Table 3.4 shows how the responses of each of these questions can be classified by correct and incorrect student models.

Table 3.4: The “Newton III” cluster on the FCI.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Forces Equal</th>
<th>Mass Dependence</th>
<th>Action Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>e</td>
<td>a, d</td>
<td>XX</td>
</tr>
<tr>
<td>15</td>
<td>a</td>
<td>b</td>
<td>c, d</td>
</tr>
<tr>
<td>16</td>
<td>a*</td>
<td>b</td>
<td>a*, c, d</td>
</tr>
<tr>
<td>28</td>
<td>e</td>
<td>d†</td>
<td>b, d†</td>
</tr>
</tbody>
</table>

Some clarification should be made with regard to this table. First of all, more than one response for a given question can fit with each incorrect mental model. This
is a fairly common occurrence throughout the FCI. An example of this duality can be seen in question four in which both the “a” and “d” responses indicate that the heavy truck is exerting more force than the light car. The difference between these responses lies in whether or not the car exerts any force on the truck. This difference is highlighted in Table 3.1 by the “Ob” misconception, and one could argue that the idea that the car wouldn’t exert any force makes response “d” more problematic than response “a”; however, prior research into students’ understanding of Newton III has determined that this issue becomes insignificant when compared with mass dependence and action dependence categories.[31, 33–35] On rare occasions certain mental models will not be represented in a question. For example, in question 4 no information is known about the relative speeds of the vehicles prior to the collision. As a result there is no evidence that a student would be using an action dependence form of reasoning to answer the question. I must also define the meanings of the accentuated responses (*, †). A student who answers “a” for question 16(*) is only considered to be correct if she also answered “a” for question 15. The necessity for this correlation lies in the fact that the vehicles in question 16 are moving at a constant velocity. Research shows that students who answer that interacting forces between two objects are equal while the objects move at a constant velocity, cannot be said to have a robust understanding of Newton’s third law unless they answer similarly for interacting objects with nonzero acceleration.[33, 34] Furthermore, since question 28 depicts a situation in which the heavier body can be considered the active agent in the scenario, response “d”(†) cannot be completely classified into the mass or action...
dependence model. As such, students responding in this manner were recorded as reasoning using half mass dependence and half action dependence.

### 3.1.2 Circular Motion Cluster

The next cluster of questions displayed in Table 3.3 involves objects undergoing circular motion: balls in circular tracks, hammer throws, etc. Questions for this cluster were taken from the “I5” and “CF” misconception groups shown in Table 3.1 as these two groups deal explicitly with objects undergoing circular motion. Additionally, two questions that are not found on the original version of the FCI (5 and 18) were incorporated in this cluster since they both involve objects in circular motion.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Correct Response</th>
<th>Force in Direction of Motion</th>
<th>Force Outward</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>b</td>
<td>c, d</td>
<td>e</td>
</tr>
<tr>
<td>6</td>
<td>b</td>
<td>a</td>
<td>c, d</td>
</tr>
<tr>
<td>7</td>
<td>b</td>
<td>a</td>
<td>c</td>
</tr>
<tr>
<td>18</td>
<td>b</td>
<td>c, d</td>
<td>e</td>
</tr>
</tbody>
</table>

For all of these questions students are asked to reason about the forces acting on the objects while they are in the circular motion as well as the trajectory the objects would follow if they were to be “let go.” Interestingly, the two questions found on both versions of the FCI (6 and 7) ask the students to reason about the path the object would take if it were to be “let go,” while questions 5 and 18 ask the students explicitly
about the presence of various forces. The two most common incorrect student mental models for this section are the idea of a force in the direction of motion and that of an outward force (related to the “CF. centrifugal force” misconception). In my classification of these models, I can see why the authors of the revised FCI added question 5 in particular. It is often difficult to determine what a student is thinking based solely on the direction he thinks an object will take after being “freed” from its circular motion. Table 3.5 shows how the responses to each of the circular motion questions is classified based on these mental models. Once again, we can see several instances in which two different responses for a given question are categorized within the same mental model. We can also see that this cluster requires no correlation between questions to determine the correctness of a student’s answer.

3.1.3 After Impetus Cluster

The After Impetus cluster on the FCI contains eight questions that involve situations in which forces that had been exerted on various objects are removed. It is so named for students’ likelihood to use impetus models of physics rather than their Newtonian counterparts. The questions in this cluster can be primarily found in the “Impetus” group of misconceptions found in Table 3.1. The exception to this is question 9 which will be discussed in the following paragraphs. Situations in this cluster include a hockey puck after it has been kicked, a rocket ship after having burned up its fuel, etc. Students are asked about the forces being exerted on the objects, their velocities, as well as the directions of their motion. These questions are designed to elicit the
idea that impetus is supplied to an object by a “hit” or other short-lived external
force and that this impetus is sustained throughout the motion of the object even
when the source of the force is no longer in contact with the object.

Table 3.6: The “After Impetus” cluster on the FCI.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Correct Response</th>
<th>Impetus Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>b</td>
<td>d, e</td>
</tr>
<tr>
<td>9</td>
<td>e</td>
<td>b</td>
</tr>
<tr>
<td>10</td>
<td>a</td>
<td>b, d</td>
</tr>
<tr>
<td>11</td>
<td>d</td>
<td>b, c</td>
</tr>
<tr>
<td>13</td>
<td>d</td>
<td>a, b, c</td>
</tr>
<tr>
<td>23</td>
<td>b</td>
<td>d, e</td>
</tr>
<tr>
<td>24</td>
<td>a</td>
<td>b, d</td>
</tr>
<tr>
<td>30</td>
<td>c</td>
<td>b, e</td>
</tr>
</tbody>
</table>

The most common incorrect student model is one that incorporates this idea of
the force of a “hit” or impetus as continuing to be exerted after the object is in
motion. In some cases students believe this impetus dies away, corresponding with
misconception “I3” in Table 3.1, but they still have the idea that a force was imparted
to, and stays with, the object by whatever agent started its motion. Table 3.6 shows
how the responses to the questions in the After Impetus cluster can be classified based
on the idea of force from the “hit.”

Question 9 is a bit of an outlier in this cluster in that it doesn’t ask the students
to think about forces or changes in velocity. It is instead concerned with the manner
in which velocities are combined. Students must know that the original velocity of
the puck \( \vec{v}_o \) and that it would have received from the kick alone \( \vec{v}_k \) will combine in some manner and have an understanding of vector addition. As such, the presence of question 9 in this cluster is debatable, but there is not another cluster in which it would better fit.

3.1.4 Constant Force with One-dimensional Motion Cluster

The Constant Force with One-dimensional Motion cluster (Constant Force–1D) includes questions that involve objects under the influence of one or more constant linear forces. These objects move in straight line paths with zero or nonzero acceleration. The questions in this cluster can be found within the “Active Force” group of misconceptions in Table 3.1. Question 27 can also be found in the “Impetus” group of misconceptions and is in fact more frequent in that category. This begs the question, why is question 27 in the Constant Force–1D cluster?

Table 3.7: The “Constant Force–1D” cluster on the FCI.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Force follows Acceleration</th>
<th>Force follows Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>c</td>
<td>a, b</td>
</tr>
<tr>
<td>17</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>25</td>
<td>c</td>
<td>d</td>
</tr>
<tr>
<td>26</td>
<td>e</td>
<td>a, b</td>
</tr>
<tr>
<td>27</td>
<td>c</td>
<td>a</td>
</tr>
<tr>
<td>29</td>
<td>b</td>
<td>e</td>
</tr>
</tbody>
</table>

The answer to this question comes in two parts. The first part is that question 27 involves the same situation as question 25 and 26, which are both in this cluster.
Question 27 concerns the behavior of a box that has been pushed across a floor (in question 25) after the woman stops pushing. The second answer to the question of question 27’s placement is that the responses fit nicely with the mental models of force being proportional to acceleration vs. force being proportional to velocity. The misconception “AF4. velocity proportional to applied force” fits the latter model quite nicely. I think it is a mistake that question 27 is not listed with this misconception, as response “a” indicates that the velocity immediately goes to zero when the applied force (the woman’s push) is removed. The questions in this cluster deal with students’ understanding of Newton’s second law ($\Sigma F = ma$), particularly with what is meant by the sum of forces and how to use the net force on an object to find its acceleration or velocity.

### 3.1.5 Constant Force with Two-dimensional Motion Cluster

The Constant Force with Two-dimensional Motion cluster (Constant Force–2D) is an interesting combination of “Impetus” and “Active Force” concepts. In each of the questions an object has a constant initial velocity in one dimension and a constant force is applied perpendicular to the direction of the object’s motion. The forces in these questions exist in two forms: the Earth’s gravitational force (questions 12 and 14) or a force exerted by the gas being emitted by a rocket’s thrusters (questions 21 and 22). Arguments could be made to group these questions with those in the After Impetus cluster, but I feel that the two-dimensional motion and the similarity of these questions with each other warrant their own cluster. Table 3.8 shows the
questions contained in the Constant Force–2D cluster as well as how the responses to these questions can be categorized based on students’ mental models.

Table 3.8: The “Constant Force–2D” cluster on the FCI.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Correct Response</th>
<th>Overcoming Impetus</th>
<th>Motion Along Straight Path</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>b</td>
<td>c, d, e</td>
<td>a</td>
</tr>
<tr>
<td>14</td>
<td>d</td>
<td>e</td>
<td>b, c</td>
</tr>
<tr>
<td>21</td>
<td>e</td>
<td>d</td>
<td>b, c</td>
</tr>
<tr>
<td>22</td>
<td>b</td>
<td>d</td>
<td>a</td>
</tr>
</tbody>
</table>

The relationship between these questions and those in the After Impetus cluster is highlighted with the definition of the first incorrect mental model: overcoming impetus. This mental model incorporates several of the misconceptions mentioned in Table 3.1 including “I2” and “I3.” The original name for this mental model was the “time-delay force model” in that the responses indicate that there is a delay between when the external force is first exerted and when the object starts changing its course. It is as if the object wants to keep going in the direction of its initial velocity, and the force takes a little while to convince the object otherwise.

The second incorrect mental model associated with the Constant Force–2D cluster incorporates the idea that the object moves in a straight-line path. This model may derive from the notion that force is proportional to velocity. This concept combined with the addition of perpendicular vectors could create the idea that the objects would move in straight lines. Little research exists to support this claim for the structure of this model, but the straight-path model is pervasive nonetheless.
3.1.6 “Strobe” Kinematics Cluster

The Strobe Kinematics cluster is so named for its questions that present students with sequential snapshots of an object’s motion as if the student is watching it using a strobe light. The students are asked to compare the motion of two different objects in terms of their velocity (question 19) and acceleration (question 20). These questions are contained within the “Kinematics” group of misconceptions outlined in Table 3.1 and are listed in Table 3.9 along with the categorization of their responses.

Table 3.9: The “Strobe Kinematics” cluster on the FCI.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Correct Interpretation</th>
<th>Incorrect Comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>e</td>
<td>b, c, d</td>
</tr>
<tr>
<td>20</td>
<td>d</td>
<td>b, c</td>
</tr>
</tbody>
</table>

The most common student mental model for the Strobe Kinematics cluster is entitled “incorrect comparisons” simply for lack of a better name. These comparisons relate with misconceptions “K1” and “K2” and are derived from the findings of Trowbridge and McDermott concerning students’ inability to accurately discern velocity from position and acceleration from velocity.[22, 24] These questions on the FCI are in a way more challenging than those presented by Trowbridge and McDermott. On questions 19 and 20, the students must be able to visualize the motion represented by the strobe pictures and then reason about the objects’ relative velocities and accelerations. Dancy and Beichner have shown that students’ responses to these particular questions greatly improve when shown videos of the motion rather than relying on
them to be able to visualize it in their heads.[51] Unfortunately, the data for the current study were collected using the original static versions of the questions. This being the case, it may be impossible to determine whether student difficulties stem from kinematics concepts or visualization of the motion. However, since the questions are presented in the same manner throughout the collection of the data, comparisons of the results can be made.

### 3.1.7 Two Objects in Free-fall Cluster

The last cluster to be described for the FCI is the Two Objects in Free-fall cluster. As its name would suggest, these questions each deal with two objects that are in free-fall. These questions are related to the “G3. heavy objects fall faster” misconception as defined by Wells, Hestenes, and Swackhamer[50] and is closely related to the beliefs of Aristotelian physics. Table 3.10 shows how the responses to the questions in the Free-fall cluster correspond to either the correct or this incorrect mental model.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Same Acceleration</th>
<th>Heavier Falls Faster</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>c</td>
<td>a, d</td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td>b, d</td>
</tr>
</tbody>
</table>

The reader will notice that each of the questions contains two responses that can be categorized as the heavier-falls-faster mental model. These correspond with the responses that the ball that is twice as heavy will fall twice as fast (1a, 2b) and
that it will fall faster but not necessarily twice as fast (1d, 2d). This grouping is not unreasonable as the pervasive mental model in this case is that heavier objects fall faster, not necessarily that their velocity (or acceleration) is proportional to their mass. An argument could be made that question 2 should be placed in the Constant Force–2D cluster as it involves exactly the same type of scenario as all of those questions. The question asked, however, is virtually identical to that in question 1, causing the most sensible cluster to be the first two questions.

### 3.2 Categorization of Responses: FMCE

The *Force and Motion Conceptual Evaluation* lends itself much more easily to creating clusters of questions than the FCI. In fact, the questions on the FMCE are grouped into clusters by the structure of the test. The FMCE also differs from the FCI in terms of the number of models that appear in each cluster. The questions of the FCI often make three models available to students: the correct Newtonian model and two misconceptions. The majority of the questions of the FMCE, however, elicit only one common misconception along with the Newtonian model. The exception to this trend occurs within the Newton III cluster in which, once again, both the mass dependence and action dependence models are prevalent. More detailed discussions of the clusters and their corresponding models are contained in the following sections.
3.2.1 Force Sled Cluster

The Force Sled cluster on the FMCE contains questions that require students to select the force a person must exert to get an object to experience a described one-dimensional motion. These questions are often referenced as “Force Sled” questions due to the fact that the forces in the possible answers are being exerted on sleds that move across frictionless surfaces.[31] The defining feature of the questions in the Force Sled cluster is that they use plain language to describe the situations. For example, the first question asks students, “Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?” The possible responses for this question are a set of seven that are common to the entire cluster. They allow the students to detail their responses concerning the applied force for each situation in terms of direction (right, left, or no force) and the magnitude (increasing, decreasing, or constant). For the most part this is a very efficient way of structuring the answer choices. The only downfall is that students do not have
the option of selecting a force that changes its nature part way through the motion (e.g. “to the right and increasing” changing to “to the right and constant,” which is a common student response for open-ended versions of question 5). The most common student mental model for the Force Sled cluster is the idea that force is proportional to velocity, or force follows velocity to borrow from Thornton’s terminology.[31]

Table 3.12: The “Force Sled” cluster on the FMCE. *–Only considered a force-follows-velocity response if student chose “g” for question 4 as well.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Force follows Acceleration</th>
<th>Force follows Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>d</td>
<td>b</td>
</tr>
<tr>
<td>3</td>
<td>f</td>
<td>c, g</td>
</tr>
<tr>
<td>4</td>
<td>f</td>
<td>g</td>
</tr>
<tr>
<td>7</td>
<td>b</td>
<td>e, g*</td>
</tr>
</tbody>
</table>

An interesting aspect of the FMCE is that not all the questions are meant to be analyzed. Some are intended for diagnostic purposes, and some do not give a definitive indication as to whether a student is using Newtonian reasoning or not. Questions five and six in the Force Sled cluster fall under these categories. Question 5 describes two distinct parts of a motion (a sled speeding up then moving at constant velocity) and asks students to reason about the force applied to the sled during the second part (the constant velocity). Thornton uses this question as a measure of his students’ reading comprehension.[42] He suggests that a student who cannot differentiate the parts of this motion will have difficulties understanding other questions on the FMCE as well.
In this regard question 5 is used as a diagnostic for assessing students’ reading ability, but is not used for determining if they think in a Newtonian manner. Question 6 is also omitted from analyses of the FMCE due to question clarity.[42] According to Thornton and Sokoloff, up to 40% of physics faculty choose the incorrect response (typically response “f”). After discussion these faculty realize that their answers should be changed to “b,” but most are unable to suggest a modification that would make the question clearer. As a result of the specialized purposes for these questions, they are not included in traditional analysis of the FMCE and will therefore not be included in my analysis.

Table 3.12 displays the questions from the Force Sled cluster that will be included in my analysis as well as the responses that correspond with each of the student models. Special attention should be paid to question 7 in which the students are asked to reason about a sled traveling to the left and slowing down. Response “g” indicates that the applied force would be toward the left and increasing. At first glance this response to question 7 does not seems to fit either the force-follows-velocity model or the correct Newtonian model. However, referencing Thornton’s studies on students’ understanding of one-dimensional forces[31], students who choose response “g” for question 7 may be in the beginning phases of a transition from their current misconception to a correct Newtonian model. Students generally have the most difficult time reasoning about objects that are slowing down. The “g” response indicates the thinking that the force won’t be directly proportional to the velocity, that it may be opposite in magnitude. The correct Newtonian model is, of course, that the di-
rection of the applied force will be opposite to the direction of the velocity (slowing down), but that the magnitude will be proportional to the change in velocity (or acceleration). In this way, response “g” displays the beginnings of a transition to a Newtonian model, but it also indicates a student’s continued use of the force-follows-velocity model. Since response “g” for questions 4 clearly indicates a student’s use of the force-follows-velocity model, students who also chose “g” for question 7 were categorized as using a force-follows-velocity model for both. This type of correlation to determine student thinking is abundant throughout the FMCE and will be discussed further in subsequent sections.

3.2.2 Vertical Motion Cluster

The questions in the Vertical Motion cluster of the FMCE ask students to reason about the net force exerted on as well as the acceleration of objects that are tossed up (or rolled up a hill) throughout their up and down, free-fall motion. Each subset of questions in this cluster has its unique qualities. Questions 8–10 present a toy car that is given a quick push up a ramp and (under its own volition) moves up and down the ramp. The remaining questions ask the students to think about a coin being tossed in the air. Questions 11–13 ask the students about the force on the coin, while questions 27–29 ask the students about the coin’s acceleration. In each of these sub-clusters the students are asked about the object’s motion as it goes up (questions 8, 11, and 27), at the highest point in its journey (questions 9, 12, and 28), and as it comes back down (questions 10, 13, and 29). The most common incorrect student
model for the Vertical Motion cluster is the notion that there is always a force in the direction of motion. This idea corresponds with the impetus model of physics in that the hand that gave the object its initial motion has somehow imparted a force to that object which sustains it until the object reaches its highest point when gravity takes over.

Table 3.13: The “Vertical Motion” cluster on the FMCE.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Constant Downward Force/Acceleration</th>
<th>Force/Acceleration in Direction of Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>8–10</td>
<td>a–a–a</td>
<td>e, f, g·d–a, b, c</td>
</tr>
<tr>
<td>11–13</td>
<td>a–a–a</td>
<td>e, f, g·d–a, b, c</td>
</tr>
<tr>
<td>27–29</td>
<td>a–a–a</td>
<td>e, f, g·d–a, b, c</td>
</tr>
</tbody>
</table>

One of the most important considerations on the Vertical Motion cluster is the correlation of responses within each of the sub-clusters of questions. Thornton and Sokoloff state that a student cannot be considered to have a Newtonian model of vertical motion if he does not respond correctly for all three questions within a group.[42] The reasoning behind this correlation is that almost all students will indicate a downward force or acceleration while an object is moving downward. Furthermore, a good number of students who recognize a downward force or acceleration will also indicate its constancy. But what about when the object is going up, or “stopped” at the top of its motion? Only students who indicate a constant downward force or acceleration for all parts of the motion are truly considered to have a Newtonian outlook on the situation.
The previous paragraph explains the “Constant Downward Force/Acceleration” column in Table 3.13 and the purposes for combining sets of questions. Now let’s look at the “Force/Acceleration in Direction of Motion” column. To be sure, the most common incorrect student model for the Vertical Motion cluster would be a combination of a force-proportional-to-velocity model, combined with an inability to distinguish between acceleration and velocity. Students would indicate this modeling by responding that the force on/acceleration of the object would be up and decreasing, zero, and down and increasing (the g-d-b response) as the object moves up, pauses, and down, respectively. But what about the students who answer “g-d-a” or “f-d-b?” Do these students reason significantly differently than the students who choose “g-d-b” across the board? I believe it is more appropriate to relax the incorrect model a bit to allow the incorporation of students’ responses that indicate simply that there is a force in the direction of motion. In fact, Bao and Redish use this student model within the demonstration of their model analysis; and, while I’ve already posited why I do not agree with their question clustering or execution, the models they chose to analyze are very prominent in introductory physics.\[1\]

3.2.3 Force Graphs Cluster

The Force Graphs cluster of the FMCE asks students to reason about the force applied to a toy car undergoing various described motions. Many of the motions described in this cluster are identical to those found in the Force Sled cluster. The distinguishing factor between these clusters is that, while the Force Sled cluster describes the force
that is being exerted in plain language, the Force Graphs cluster supplies the students with various graphs of applied force vs. time throughout a given motion. Similarly to the Force Sled cluster, all of the correct responses indicate that the applied force is constant. Also, the most common incorrect student model on the Force Graphs cluster is the force-follows-velocity model. With all of these similarities, it is reasonable to wonder why the Force Sled and Force Graphs cluster were not grouped together. In light of Beichner’s studies of students’ understanding of graphs in kinematics[25] and Kohl and Finkelstein’s research into the effects of representation on students’ performance,[26, 27] the simple fact that the possible responses are formatted differently makes these clusters fundamentally different in terms of assessing students’ understanding of physics.

Table 3.14: The “Force Graphs” cluster on the FMCE.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Force follows Acceleration</th>
<th>Force follows Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>e</td>
<td>a</td>
</tr>
<tr>
<td>16</td>
<td>a</td>
<td>c</td>
</tr>
<tr>
<td>17</td>
<td>e</td>
<td>a, b</td>
</tr>
<tr>
<td>18</td>
<td>b</td>
<td>h</td>
</tr>
<tr>
<td>19</td>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>20</td>
<td>g</td>
<td>f</td>
</tr>
<tr>
<td>21</td>
<td>e</td>
<td>a, h</td>
</tr>
</tbody>
</table>

Similar to question 5, question 15 serves to diagnose student difficulties outside the realm of physics. Question 15 asks the students what the (horizontal) force would be on a car at rest. Thornton and Sokoloff suggest that incorrect responses to this
question reflect students’ inability to interpret graphs, not their incorrect notions concerning physics.[42] Additionally, over 90% of introductory physics students will answer this question correctly before instruction. For these reasons, question 15 is not considered a test of students’ Newtonian reasoning and is omitted analysis of the FMCE.

Another consideration that must be made is the designation of multiple responses for a single model on a given question. Unlike the FCI, the FMCE does not supply multiple responses for a given question that reiterate similar ideas. Each response is unique. On question 17, response “b” indicates a strictly force-follows-velocity model. Response “a” is not entirely different as it is congruent with this model in magnitude but not direction. A response of this type, however, could simply indicate a confusion between left and right as negative and positive, respectively. If a student who chooses response “a” displays an overall tendency to reason using a force-follows-velocity model, as indicated by her responses to other Force Graphs questions, then she is considered to be consistent in this regard.

Question 21 poses another interesting duality. This question asks students about the force on the car once it has been released (after being pushed). On one hand, response “a” seems to fit perfectly with the force-follows-velocity model: the car moves at a constant velocity so a constant force must be applied. On the other hand, what if the students don’t ignore friction (as they are explicitly told to do) or use an “impetus dies away” model? In this case, the car would slow down at a (more or less) steady rate, indicating a positive yet decreasing force (response “h”). Both of
these responses can be considered a force-follows-velocity model. They differ mainly in students’ abilities to picture an object actually moving at constant velocity without need of a force to counteract friction.

3.2.4 Acceleration Graphs Cluster

The Acceleration Graphs cluster is very similar to the Force Graphs cluster. In the Acceleration Graphs cluster students are asked about the acceleration of a toy car undergoing various types of motion. Again, students must choose the graph they believe best represents the acceleration of the car for each scenario. It should be noted that the parenthetical reminders of “(constant acceleration)” that are found in the questions of the Force Graphs and Force Sled clusters are omitted from these questions. The most common incorrect student model for the Acceleration Graphs cluster stems from students’ inability to distinguish between velocity and acceleration.[24] In this regard, students typically respond using an acceleration-follows-velocity model. Table 3.15 shows how this model is implemented within the various scenarios of the cluster.

Table 3.15: The “Acceleration Graphs” cluster on the FMCE.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Acceleration follows change in Velocity</th>
<th>Acceleration follows Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>22</td>
<td>a</td>
<td>e</td>
</tr>
<tr>
<td>23</td>
<td>b</td>
<td>g</td>
</tr>
<tr>
<td>24</td>
<td>c</td>
<td>b</td>
</tr>
<tr>
<td>25</td>
<td>b</td>
<td>f</td>
</tr>
<tr>
<td>26</td>
<td>c</td>
<td>a</td>
</tr>
</tbody>
</table>
3.2.5 Newton III Cluster

The cluster of questions concerning Newton’s third law is particularly interesting. It is the only cluster on the FMCE that elicits two different incorrect student models: the mass dependence model and the action dependence model. As defined in section 2.2, the mass dependence model is evident when students think that more massive objects always exert more force than less massive ones. On the other hand, the action dependence model states that “active” objects exert more force than “passive” ones. For example, if one car is pushing another, the car that is pushing will exert more force than the car that is being pushed. Also, in a collision the object that is initially going faster will exert more force. Table 3.16 shows how the responses to the questions in the Newton III cluster separate into these models.

Questions 33, 35, 37, and 39 are omitted from the analysis of the Newton III cluster. Thornton and Sokoloff report that question 33 should be omitted for similar reasons to question 15. In question 33, two vehicles of equal mass, traveling at equal velocity collide. Neither the mass dependence model nor the action dependence model is applicable in this situation. As such, almost all students will get this question correct. A student who does not answer correctly most likely has a problem with the language of the question, not necessarily the physics.[31] Question 35 is omitted because it tends to elicit the notion that no forces are exerted when objects aren’t moving. Question 37, similar to 33, has a tendency to elicit false positives due to the fact that the vehicles are moving at constant velocity. Finally, students using either
the mass dependence or action dependence model will choose the same response since
the boy doing the action (pushing) also has more mass.

Table 3.16: The “Newton III” cluster on the FMCE. *–If student chooses “a” for question 30 and “b” for question 32, a response of “e” or “f” for question 31 is counted as half MD and half AD. †–Considered half MD and half AD if a student chooses “a” for 30, “e” or “f” for 31, and “b” for 34. **–Categorization for this response depends on the student’s choice for question 36. ‡–Only considered an AD response if student chooses “c” for question 36.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Forces Equal</th>
<th>Mass Dependence</th>
<th>Action Dependence</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>e</td>
<td>a</td>
<td>XX</td>
</tr>
<tr>
<td>31</td>
<td>e</td>
<td>a, e*, f*</td>
<td>b, e*, f*</td>
</tr>
<tr>
<td>32</td>
<td>e</td>
<td>a, e‡</td>
<td>b, e‡</td>
</tr>
<tr>
<td>34</td>
<td>e</td>
<td>XX</td>
<td>b</td>
</tr>
<tr>
<td>36</td>
<td>a</td>
<td>b</td>
<td>c</td>
</tr>
<tr>
<td>38</td>
<td>a</td>
<td>b**</td>
<td>a‡, b**, c</td>
</tr>
</tbody>
</table>

Table 3.16 requires some explaining. Studies into students’ understanding of Newton III have shown that almost all students use both the mass dependence and action dependence models prior to instruction.[33, 34] Furthermore, students have been known to apply both models simultaneously such that they arrive at the correct answer without a true understanding of Newton’s third law. The phenomenon of false positives is most prevalent on free-response questions in which a small, fast car collides with a heavy, slow truck (as in question 31). With this in mind, a correct response for question 31 must be validated by correct responses for other questions. If, on the
other hand, the student does not respond correctly to other questions in the Newton III cluster, we must consider all of the mental models she is using.

A response of “a” on question 30 (a small car and a heavy truck moving at the same speed collide) indicates use of the mass dependence model. Furthermore, a response of “b” on question 32 (the truck is now still) indicates the student’s use of the action dependence model. If a student responds with the pattern “a-e-b” for questions 30–32, the student’s response for question 31 is considered to be the result of a compensation argument between the higher mass of the truck and the higher initial speed of the car. In that case, “e” for question 31 is coded as 1/2 mass dependence and 1/2 action dependence. Thornton also reports this discrepancy:

By examining the student written responses we find that additional students answer that the forces are equal in question (31) because they believe that the heavier mass of the truck just happens to balance the higher speed of the car.[31]

But what if the student doesn’t think the mass of the truck exactly balances speed of the car? In this case, a response of “a-f-b” for questions 30–32 would be coded the same way as the “a-e-b” response. I believe that these students are using the same compensation argument as the students who believe the two factors exactly compensate.

Another consideration must be made for students who answer “a-e-e-b” or “a-f-e-b” for questions 30–32, and 34. A response of “b” for question 34 (both vehicles have the same mass, but the truck is not moving) indicates an action dependence
type of reasoning. For these sets of responses, students display the use of both mass
dependence and action dependence models. Furthermore, an idea of compensation
similar to that described above seems to be in use while these students reason about
the situations in which the small car is faster than the large truck (questions 31 and
32).

The final consideration comes in question 38. If a student responds with response
“b,” it is impossible to tell whether he is using the mass dependence or the action
dependence model. The justification of action dependence reasoning is that the truck
is applying its brakes and causing the car to slow down. In order to determine this
student’s thinking, one must look at his answer to question 36. If on 36 he indicates
mass dependence reasoning (response “b”), it is likely that he is using the mass
dependence model again. If he uses action dependence reasoning for 36 (response
“c”), it is likely that he is using the action dependence model again on question 38.
Additionally, a student who chooses “a” for question 38 may not be using correct
Newtonian reasoning. If she responds using action dependence reasoning for question
36, it is likely that she is using a compensation argument between the action of the
car pushing and the action of the truck braking. In this case, response “a” would be
coded as the action dependence model.

The Newton III cluster highlights the down side to multiple choice assessments.
There is no way to know for certain why each student chooses the answers that she
does. The best we can do as researchers and instructors is relate the given responses
with previous results using free-response questions and correlate the responses for several questions to get a sense of her overall understanding of the concepts.

### 3.2.6 Velocity Graphs Cluster

The Velocity Graphs cluster is again very similar to the previously described “graphing” clusters. Students are presented with various descriptions of a car’s motion, and they must choose the correct velocity vs. time graphical representation for the motion. As with the Acceleration Graphs questions, the most common student model is derived from Trowbridge and McDermott’s studies of students’ understanding of kinematics and their inability to distinguish position from velocity.[22] This velocity-follows-position model is also closely related to Beichner’s proposition that students view graphs as a picture of the situation no matter what the axes indicate.[25] Table 3.17 shows how the responses in this cluster correspond with the various student models.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Correct view of Velocity</th>
<th>Velocity follows Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>a</td>
<td>d</td>
</tr>
<tr>
<td>41</td>
<td>f</td>
<td>g</td>
</tr>
<tr>
<td>42</td>
<td>b</td>
<td>c, h</td>
</tr>
<tr>
<td>43</td>
<td>d</td>
<td>XX</td>
</tr>
</tbody>
</table>

On question 42, the toy car is said to be moving at a constant velocity toward the left (toward the origin). Responses “c” and “h” both indicate a graph that gets
steadily closer to the horizontal axis as time progresses. For response “c” a student could be picturing the car starting at the right and moving toward “0,” and students choosing “h” could be triggered by the word ”left” to choose a graph that depicts negative velocity.

### 3.2.7 Energy Cluster

The Energy cluster on the FMCE contains questions that ask students to reason about the speed and kinetic energy of a sled after sliding down a hill. In each question the student is asked to compare the speed or kinetic energy of the sled after it has gone down a specific hill with the same quantity after the sled had gone down a reference hill that differs in slope and/or height. The correct answer is that a higher hill would cause the sled to have a greater change in potential energy, in turn causing a greater change in kinetic energy. This also means that the higher hill allows the sled to have a greater velocity at the bottom. Contrary to popular belief, the speed (and/or kinetic energy) does not depend on the steepness of the hill, which is the most common incorrect student model. Table 3.18 shows how the question in this cluster are divided amongst the correct and this incorrect model.

Once again we must correlate several responses to determine what the student is thinking. In particular, response “b” on questions 46 and 47 most likely indicates a compensation argument between the increased height and the decreased steepness. In this case the combination of the two responses would yield a coding of each of the responses as height dependence and slope dependence. However, the student’s
Table 3.18: The “Energy” cluster on the FMCE. *–Categorization of the “b” response for questions 46 and 47 depends on the student’s responses to question 44 and 45.

<table>
<thead>
<tr>
<th>Question Number</th>
<th>Energy/Velocity Depends on Height</th>
<th>Energy/Velocity Depends on Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>44</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>45</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>46</td>
<td>a, b*</td>
<td>b*, c</td>
</tr>
<tr>
<td>47</td>
<td>a, b*</td>
<td>b*, c</td>
</tr>
</tbody>
</table>

responses to questions 44 and 45 can be used to determine which, if either, is the case. If a student strictly uses a steepness argument for 44 and 45, it is likely he would use it again. The same applies for the student who reasons using strictly height arguments for the first two questions.

3.2.8 Old Clusters vs. New Clusters

The majority of research that has been conducted using the FMCE has not employed the clusters that have been defined in the previous sections. Instead, researchers have used a set of clusters defined by Wittmann in the late 1990’s.[52] The questions that comprise these clusters are shown in Table 3.19.

At first look these clusters seem ideal for analyzing the FMCE. The clusters are concise and group the questions in terms of the type of material in each. In fact the Velocity, Newton III, and Energy clusters are the same as those defined in Table 3.11; however, there are several reasons for dividing the Force (Newton I & II) and Accel-
Table 3.19: Clusters of questions on the FMCE as defined by Wittmann.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity</td>
<td>40–43</td>
</tr>
<tr>
<td>Acceleration</td>
<td>22–29</td>
</tr>
<tr>
<td>Force (Newton I &amp; II)</td>
<td>1–4, 7–14, 16–21</td>
</tr>
<tr>
<td>Newton III</td>
<td>30–32, 34, 36, 38</td>
</tr>
<tr>
<td>Energy</td>
<td>44–47</td>
</tr>
</tbody>
</table>

eration clusters. First of all, studies by Beichner[25] and Kohl and Finkelstein[26, 27] have shown that the manner with which material is presented significantly effects students’ abilities to demonstrate their understanding. With this in mind, the questions contained in the Force Graphs and Acceleration Graphs clusters should be considered inherently different than those that do not involve graphical interpretation. The questions contained in the Force Sled cluster contain situations in which a person in direct contact with an object exerts a horizontal force, while those contained in the Vertical motion cluster involve non-contact, vertical forces being exerted on objects in free-fall. Moreover, Dykstra proposes that students have trouble reasoning about motion in which an object has a turning point; that is, when an object under the influence of a constant force moves in a particular direction while slowing down and then reverses direction and speeds up.[53] Dykstra used examples of horizontal motion in his study, but the questions in the Vertical Motion cluster (both force and acceleration) are directly analogous to his scenarios.
In addition to the differences in cluster definitions, typical analysis of the FMCE does not consider students’ incorrect mental models, nor does it correlate responses to determine students’ true understanding.\(^4\) If all we are concerned with is whether or not students are correct on a given question, this is not a problem; however, in order to properly assess students’ understanding we must consider their prior knowledge and modes of thinking. Additionally, for the reasons stated in the previous sections, it is sometimes important to correlate students’ responses on certain question in order to properly determine their understanding. Without these correlations, it is likely that students’ response will be incorrectly coded as “right” or “wrong.”

\(^4\)The exception to this is the questions contained in the Vertical Motion cluster. These questions are typically grouped into the three sub-clusters defined above.
CHAPTER 4

DATA AND ANALYSIS

Data for this study come from two very different sources. The College of DuPage in Glen Ellyn, IL is a two-year community college located approximately 25 miles west of downtown Chicago. As such, the 34,000-student population of the College of DuPage consists of students from the suburbs of Chicago with approximately 25%–50% minority students. The University of Maine is almost a polar opposite. Located in a small town of approximately 9,000, the university’s population of 11,000, over 90% of which is white, overwhelms the community with students from various nations and states, while most of the student population is from throughout the state of Maine. The University of Maine offers a wide variety of four-year undergraduate degrees and graduate degrees from five different colleges. The University of Maine also maintains a student-faculty ratio of 16:1 compared with the College of DuPage’s 21:1. Furthermore, this value for the University of Maine does not consider the numerous teaching assistants provided by its Graduate School. Additionally, the University of Maine is organized under a semester-long course structure while the College of DuPage operates using quarter-long courses. As a result of these demographic differences and those concerning the classes from which data were collected that will be discussed
in the following sections, it is beneficial to examine the data from each of these institutions separately.

4.1 College of DuPage

Data were collected at the College of DuPage (COD) within an algebra-based introductory mechanics course using the FMCE. Data exist from 2000–2004. In 2000 and 2001, two sections of the algebra-based course were offered: a morning and an afternoon section. The morning and afternoon sections were virtually identical for each year in terms of method of teaching. Beginning in 2002 the morning algebra-based section was replaced by an evening calculus-based course. Data from the calculus-based course were collected using the FCI. From 2000–2004 various changes were made to the ways in which each of the courses was taught. Table 4.1 and Table 4.2 depict how each of the courses changed through the years. These changes will be explained in more detail in the following sections.

4.1.1 Algebra-based Course

The algebra-based course at COD runs for three quarter-long sessions each year and covers a variety of physics topics including linear and rotational kinematics and dynamics, thermodynamics, waves and sound, and electrostatics. The course set-up is fairly standard: four hours of lecture and problem solving accompanied by two hours of laboratory work each week. Within these lab periods students were occasionally required to perform problem solving tasks within small groups that were not neces-
sarily directly related with their current experiment. Beginning in 2000, the *Tools for Scientific Teaching* (TST)[54, 55] laboratory materials were implemented to replace the confirmation-style labs that had been previously used. Approximately ten of the TST labs were completed throughout the course.

Many aspects of the course changed in 2001. First of all, the TST labs were replaced by *RealTime Physics* (RTP)[57, 58] versions. In addition, the number of lab sessions was increased from ten to fourteen giving students more hands-on experience. Furthermore, the small-group problem solving activities were moved from the lab periods to the lecture periods. Students were required to complete reading assignments prior to lecture periods to familiarize themselves with the basics of the material to be covered, and all tests and examples were modified to be gender-neutral.¹ In the afternoon section in 2001, a Pupil Response System (PRS) was employed as recommended by *Interactive Lecture Demonstrations* curricula. The PRS allows each student to see how their ideas compare with the ideas of the entire class without being singled out or identified. This practice enables both the professor and the students to gain insight on the views of the class both before as well as during the lesson. Additionally, in-class examples were taken from Mazur’s book[59], which provides examples of how to use peer instruction effectively.

The biggest change in 2002 was the implementation of the web-based homework system, TYCHO. The only other change this year was that students were no longer

¹Research into gender biases in physics suggests that gender-neutral scenarios in examples and assignments may be beneficial to both male and female students.[61]
Table 4.1: Changes that were made each year to the algebra-based physics course at the College of DuPage: 2000–2004.

<table>
<thead>
<tr>
<th>Year</th>
<th>Changes from Previous Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>00–Morning</td>
<td></td>
</tr>
<tr>
<td>· Lecture-style course, 4 hours per week, 2 hour labs</td>
<td></td>
</tr>
<tr>
<td>· Some small-group problems in lab</td>
<td></td>
</tr>
<tr>
<td>· No enforced structure to group problems</td>
<td></td>
</tr>
<tr>
<td>· Began implementing <em>Tools for Scientific Thinking</em> (TST) labs[54, 55]</td>
<td></td>
</tr>
<tr>
<td>· Completed approximately 10 labs</td>
<td></td>
</tr>
<tr>
<td>· Used Cutnell &amp; Johnson text[56]</td>
<td></td>
</tr>
<tr>
<td>00–Afternoon</td>
<td></td>
</tr>
<tr>
<td>· No difference from morning section</td>
<td></td>
</tr>
<tr>
<td>01–Morning</td>
<td></td>
</tr>
<tr>
<td>· Replaced TST labs with <em>RealTime Physics</em> (RTP) materials[57, 58]</td>
<td></td>
</tr>
<tr>
<td>· Increased number of lab sessions to 14</td>
<td></td>
</tr>
<tr>
<td>· Reading assignments required students to answer simple questions about each chapter before the material was discussed in class</td>
<td></td>
</tr>
<tr>
<td>· Tests and examples modified to be gender-neutral</td>
<td></td>
</tr>
<tr>
<td>· Moved small-group interactions from labs to class periods</td>
<td></td>
</tr>
<tr>
<td>· Additional peer instruction using a Pupil Response System (PRS)</td>
<td></td>
</tr>
<tr>
<td>· Questions taken from Mazur’s book[59] and/or modified locally</td>
<td></td>
</tr>
<tr>
<td>01–Afternoon</td>
<td></td>
</tr>
<tr>
<td>· Implemented web-based homework system (TYCHO)</td>
<td></td>
</tr>
<tr>
<td>· Students not allowed to drop course after halfway point</td>
<td></td>
</tr>
<tr>
<td>· Peer instruction continued</td>
<td></td>
</tr>
<tr>
<td>02–Afternoon</td>
<td></td>
</tr>
<tr>
<td>· Students allowed to drop course if they spoke with a COD counsellor (similar to previous policy)</td>
<td></td>
</tr>
<tr>
<td>03–Afternoon</td>
<td></td>
</tr>
<tr>
<td>· Cutnell &amp; Johnson text replaced by Walker text[60]</td>
<td></td>
</tr>
<tr>
<td>· Only one small-group session throughout the course</td>
<td></td>
</tr>
<tr>
<td>· Lesson timing adversely affected by new text</td>
<td></td>
</tr>
</tbody>
</table>
allowed to drop the course after the halfway point. This is a change that was modified in 2003 when students were allowed to drop but only if they spoke with a COD counsellor first.

In 2004 the Cutnell & Johnson[56] text was replaced by the Walker[60] text. Both of these textbooks are quite well-known and widely used, and both are quite good for introductory, algebra-based physics courses. The process of the change, however, had adverse effects in and of itself. The timing of the lessons was compromised, and, as a result, fewer small-group problem solving assignments were completed. Figure D.1–Figure D.8 in Appendix D show the model plots from the FMCE data collected in the algebra-based course throughout the years.

As can be seen from most of these figures, the data didn’t change much from year to year. The plots that depict the data from the Force Sled, Force Graphs, Acceleration Graphs, and Energy clusters (Figure D.1, Figure D.3, Figure D.4, and Figure D.8) are particularly noticeable in the similarities among both the pre-instruction data and the post-instruction data from each year. The data from the Velocity Graphs cluster also vary little from year to year, but this cluster has the unique feature of having very high pretest scores. Most students have a fairly good grasp of the concepts presented in the Velocity Graphs cluster before any instruction occurs. It should also be noticed that for all plots, the pretest data from each year are nearly identical. This means that the \textit{a priori} ideas about physics held by the student population of this course at COD changed very little from year to year. This is convenient for the study in that it allows a meaningful comparison between the post-instruction data from each year.
Furthermore, the conceptual gains depicted throughout the data are incredibly large as shown by pretest data in the lower right-hand corner of the plot and post-test data in the upper left-hand corner.\(^2\)

The truly interesting data are found within the Vertical Motion and Newton III clusters (Figure D.2, Figure D.6, and Figure D.7. On the Vertical Motion cluster, the data from the 2000 afternoon section, both 2001 sections, and the 2003 and 2004 afternoon sections all seem fairly large and relatively similar. However, the data from the morning section of the 2000 course and that from the 2002 section are less encouraging. These data points indicate approximately half the improvement of the other years/sections. The 2000 morning point is particularly interesting in its difference from that of the 2000 afternoon section. The reader will recall that there were no intentional differences between the instruction of these two sections.

The data for the Newton III cluster are similar in their dichotomy. The reader should first notice that there are two sets of data for the Newton III cluster: one for the Mass Dependence student model (Figure D.6) and one for the Action Dependence model (Figure D.7). Each of these was created using the same set of raw data, so they have many similarities. The major difference between the two plots is the higher prevalence of the Action Dependence mental model. This higher prevalence may be a result of the Action Dependence model being applied differently in collision and

\(^2\)The reader will recall that a data point in the lower right-hand corner of the model plot indicates a class of students where almost everyone exclusively uses the most common incorrect model, while a data point in the upper left-hand corner indicates a class populated by students who almost exclusively responds using the correct Newtonian model.
pushing situations (faster object exerts more force, and pusher exerts more force, respectively), while the Mass Dependence model is the same for each. The relative positions of the data points from year to year are similar for each plot. The data from the 2000 sections indicate significantly less conceptual change than those of the following years. Additionally, the afternoon section in 2000 appears to do a bit better than the morning section. This difference is not nearly as pronounced as that found in the Vertical Motion cluster, but it is still noticeable. It should be noted that data on the Energy cluster were not collected in 2000, disallowing any search for a similar pattern in that cluster. In many ways, the Newton III data make more sense than those in the Vertical Motion cluster. Student learning of Newton’s third law improved as the course was modified. But why did student learning of vertical motion concepts drop so drastically in 2002? Possible responses are discussed in Chapter 5.

4.1.2 Calculus-based Course

Several changes were made to the calculus-based course at COD throughout the years as well. This course also spanned three quarter-long sessions. The calculus-based course generally covers the same topics that are included in the algebra-based course. The description of the course in 1998 displayed in Table 4.2 is purely for comparison purposes. No data exist from this year. The course in 1998 consisted of: 1) evening classes in which all time was spent on lecturing, and 2) laboratory sessions in which students performed experiments to confirm various physics principles. Every other
week, small-group problem solving was completed using end-of-chapter problems from
the Halliday, Resnick, & Walker text[62].

Table 4.2: Changes that were made each year to the calculus-based

<table>
<thead>
<tr>
<th>Year</th>
<th>Changes from Previous Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>98–Evening</td>
<td>· Lecture-based course</td>
</tr>
<tr>
<td></td>
<td>· Evening classes</td>
</tr>
<tr>
<td></td>
<td>· Used Halliday, Resnick &amp; Walker text[62]</td>
</tr>
<tr>
<td></td>
<td>· Non-computer-based “confirmation” lab sessions</td>
</tr>
<tr>
<td></td>
<td>· Group-work completed using end-of-chapter (EOC) problems during labs on alternating weeks</td>
</tr>
<tr>
<td>02–Evening</td>
<td>· Implemented Peer Instruction and PRS use</td>
</tr>
<tr>
<td></td>
<td>· Switched to RTP labs</td>
</tr>
<tr>
<td></td>
<td>· Group-work on EOC problems completed in class but not on a regular basis</td>
</tr>
<tr>
<td>03–Evening</td>
<td>· Switched to web-based homework assignments</td>
</tr>
<tr>
<td></td>
<td>· 10 TYCHO problems plus 3–4 CyberTutor[63] problems assigned each week</td>
</tr>
<tr>
<td></td>
<td>· Group-work still completed in class on an irregular basis</td>
</tr>
<tr>
<td>04–Evening</td>
<td>· All homework assignments from TYCHO</td>
</tr>
<tr>
<td></td>
<td>· Switched to Cummings, Redish &amp; Laws text[64]</td>
</tr>
<tr>
<td></td>
<td>· Fewer group-work sessions (approx. one per month)</td>
</tr>
</tbody>
</table>

In 2002 the lecture sessions of the course were modified to include Peer Instruction\(^3\) with the use of a PRS. The small-group problem solving sessions using EOC problems were moved to the lecture periods to allow more laboratory time and more opportunities for social learning. In addition to the lecture period modifications, the

\(^3\)Similar to ILDs.[59]
confirmation-based labs were replaced by RTP lab activities which promote exploratory rather than procedural experiments.

The most significant change in 2003 was the switch to web-based homework assignment from the previous paper-based versions. Ten of these assignments were distributed through the TYCHO system, and an additional three or four came from the CyberTutor program. The CyberTutor program is designed to be more interactive than the TYCHO system by assisting students as they complete their homework. Small-group work was still completed within the lecture portion of the class, still not on a regular basis.

Anecdotal reports from students in the 2003 course indicated a severe degree of animosity toward the CyberTutor problems. As such, the CyberTutor homework assignments were removed in 2004 and replaced by TYCHO counterparts. In addition, the text was changed from the standard Halliday, Resnick, & Walker text to the version edited by Cummings, Redish, & Laws. This text is designed to be aligned with physics education research findings to provide good learning opportunities and experiences for the students. The number of small-group problem solving sessions decreased in 2004 to approximately one each month. Figure E.2–Figure E.10 in Appendix E show the model plots of the FCI data collected within the calculus-based course.

---

4The CyberTutor program has been renamed MasteringPhysics.[63]
5The professor of the course received comments from many of his students such as, “PLEASE DON’T USE CYBERTUTOR EVER AGAIN!!!!!!!!!!”
One of the most interesting changes in the data from the calculus-based course from that gathered from the algebra-based course is the difference of the students’ initial knowledge state from year to year. Data exist only for a three-year period (2002–2004), but each year in this period seems to draw a different student population. Moreover, the data are inconsistent as to which year depicts the students with the most correct initial knowledge base. For all question clusters, with the exception of Newton III and Circular Motion, the students from the 2004 course seem to begin the course with more appropriate physics knowledge than the students from 2003, who in turn seem to have more initial knowledge than the students from 2002. The data for the Newton III cluster, however, indicate that the students from 2003 began the course with fewer correct mental models than the students had in 2002. And in the Circular Motion cluster, the students in the 2004 class began the year with more incorrect mental models than either of the other two years. With all of these differences in the initial knowledge state of the students from year to year, it is inappropriate to base conclusions solely on the post-instruction results. We must instead consider each class’s “movement” through the model plot from pre- to post-instruction.

The arrows shown on the model plots depict this “movement.” An arrow that points primarily toward the left indicates that students in the class abandoned their use of the incorrect mental model, but did not typically adopt the correct Newtonian view. On the other hand, an arrow that points mostly upward indicates a class that accepted the Newtonian model very well, but did not radically change their views concerning the most common incorrect model. Depending on the initial state of the
class’s knowledge a vertical arrow could be considered either good or bad, where a horizontal arrow is generally not considered a triumph.

It is noted above that the 2003 data on the Newton III cluster depict a lower pre-instruction score than either of the other two years. It is interesting then that this year shows the highest post-instruction score. In fact, every student in the 2003 class responded correctly to all four of the Newton III questions on the FCI. The arrows for the 2002 and 2004 data are similar in length and pitch with the 2004 slightly out-performing its 2002 counterpart. The difference in post-instruction scores seems only slightly larger than the difference in those gathered pre-instruction.

The data for the Circular Motion cluster shown in Figure E.4 & Figure E.5 depict a more typical representation of the data from the calculus-based course. In both figures, the pre-instruction data from 2002 indicate that these students displayed less tendency to utilize the appropriate Newtonian mode of thinking than the students from either of the subsequent years. Furthermore, the 2002 students also display less conceptual development as indicated by the shorter length of the data arrow and its more horizontal trajectory. It should also be noted that the data from the 2003 and 2004 classes are very similar. For each incorrect model in the Circular Motion cluster, the students from 2004 started the course with a slightly higher understanding of the correct Newtonian model of physics. This understanding gap does not change much from 2003 to 2004 indicating that the students in both years improved by approximately the same amount.
This tendency for the 2003 class to outperform the 2002 is replicated throughout the remaining clusters. It is visually simple to see that for each cluster, the arrow representing the conceptual change undergone by the students in 2003 moves farther upward and just as much toward the left as the arrow for the 2002 students. Furthermore, the arrow for the 2004 students is typically very similar to that of the 2003 students or slightly more vertical (as in the Strobe Kinematics cluster and the Circular Motion cluster for the force-in-the-direction-of-motion model). The only remaining difference in these trends occurs within the Free-fall cluster in which the 2002 and 2003 data are virtually identical in terms of conceptual development, while the 2004 data depict a significantly steeper improvement. It should be noted that the students in the 2004 class began the course with less frequent use of the most common incorrect model. If the model plots are examined with this in mind, evaluating solely the increase in Newtonian thinking (vertical change), the data for the 2003 and 2004 classes become very similar.

### 4.2 University of Maine

The General Physics course (PHY 111/112) at UMaine is a two-semester sequence that covers linear and rotational kinematics and dynamics, work and energy, gravitation, waves and sound, electrostatics, circuits, electrodynamics/magnetics, and optics. Data for this study were gathered at the beginning and the end of the first semester (PHY 111) by which time the students had completed their study of classical mechanics. The course is designed with two hours of lecture, two hours of recitation (or
tutorial), and two hours of laboratory work each week. Table 4.3 shows how each of these aspects of the course have changed from 2003–2005.

Table 4.3: Changes that were made each year to the algebra-based physics course at the University of Maine: 2003–2005

<table>
<thead>
<tr>
<th>Changes from Previous Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
</tr>
<tr>
<td>· Standard lecture format</td>
</tr>
<tr>
<td>· <em>Tutorials in Introductory Physics</em> (TIP)<em>[43]</em> used during all “recitation” periods</td>
</tr>
<tr>
<td>· Confirmation-style laboratory work</td>
</tr>
<tr>
<td>2004</td>
</tr>
<tr>
<td>· New primary instructor</td>
</tr>
<tr>
<td>· <em>Interactive Lecture Demonstrations</em> (ILDs)<em>[42]</em> implemented for some lectures</td>
</tr>
<tr>
<td>· Studies conducted in Newton III and energy tutorial periods in which portions of students were given tutorials other than those from the TIP materials</td>
</tr>
<tr>
<td>· Study conducted in lab periods in which a modified version of the modeling laboratory methods*[41]* was utilized</td>
</tr>
<tr>
<td>2005</td>
</tr>
<tr>
<td>· More ILDs used</td>
</tr>
<tr>
<td>· More TIP tutorials traded for alternate versions</td>
</tr>
<tr>
<td>· Modified modeling labs implemented for all sections</td>
</tr>
<tr>
<td>· Lecturer trained TAs for tutorial and laboratory sessions</td>
</tr>
</tbody>
</table>

The first aspect to notice in Table 4.3 is the change in the primary instructor between 2003 and 2004. This has become a habit of the PHY 111 course. From 1999–2006, the PHY 111 course has had six different primary instructors. In addition to the change in face of the lecturer, each primary instructor provided a different level of involvement in the course ranging from personally preparing each teaching
assistant (TA) and peer facilitator\textsuperscript{6} to passing complete control of the laboratory periods to another party. This instability in and of itself could prove damaging to the course, but the data for this particular study come from years in which this confounding factor is minimized. In addition to the primary instructor, the TAs and peer facilitators who teach the tutorial and laboratory sections for the PHY 111 course change almost completely from year to year. Fortunately, the relative stability of the primary instructor during two of the three years of this study allowed these transitions to be smooth and consistent with the previous year’s instructional methods.

Personnel changes aside, the PHY 111 course has made significant changes to its structure in all aspects of the course. In the lecture portion of the course, ILDs were first introduced in 2004 with the use of a PRS and became more frequent and integral in 2005. These changes required the students to be more engaged in the lecture periods by asking them to make predictions and justifications rather than simply “soak in” the knowledge that is presented by the lecturer.

In addition to these changes made to the lecture portion of the course, modifications also occurred within the recitation sessions. By the mid-1990’s, traditional recitation sessions of TAs standing at a chalk board and solving problems had been abolished in PHY 111 in favor of preliminary versions of the tutorials from the University of Washington and later the TIP materials in their published form. In 2004 several studies were conducted by undergraduate seniors within the PHY 111 tutorial

\textsuperscript{6}An undergraduate student who has previously taken PHY 111 or an equivalent course and has performed particularly well. Peer facilitators assist the TAs during tutorial sessions.
sections that replaced the TIP tutorials with other varieties for some students. One study substituted the TIP Newton III tutorial with a version from the *Activity-Based Tutorials (ABT)*[45] for a third of the class and the *Maryland Open-source Physics Tutorials (OPT)*[46] for another third. The remaining students still used the TIP version. The results of this study showed that the OPT tutorial was more effective than either the ABT or the TIP version.[34] Another study took place during the work-energy portion of the course in which half of the students were given a locally developed tutorial while the other half were given the standard TIP energy tutorial. The results of this second study showed that the students who had used the local tutorial performed significantly better on post-instruction assessments than their TIP counterparts.[65]

In 2005 the tutorial portion of the PHY 111 course was modified even further. The results of the previous year’s studies spawned the implementation of OPT tutorials for many portions of the course. The 2004 Newton III study was conducted as a second trial during this year but without the same degree of significance in the results. The differentiation of instruction that occurred in 2004 within the work-energy tutorial sessions was also repeated in 2005.

The laboratory sections of PHY 111 also changed significantly in 2004. A third study conducted by an undergraduate senior implemented a modified version of Wells, Hestenes, and Swackhamer’s modeling lab procedures for half of the student population.7 The success of this study[66] caused the implementation of these modified

---

7This half of the student body is not correlated in any way with the half that comprised the experimental group during the energy-tutorial study, nor do either of these groups correspond with
modeling methods within all laboratory sections starting in 2005. By this point, research-based curricula were being implemented in all portions of the PHY 111 course. Figure F.1–Figure F.8 in Appendix F show the data collected for each of these years and how the various curricular changes have impacted students’ conceptual development.

Once again, the reader will notice that the pre-instruction data for each of the three years are almost identical for each cluster allowing a meaningful comparison based solely on the post-instruction data. Interestingly, these data points are also very similar to the pre-instruction data from the algebra-based course at the College of DuPage. The post-instruction results, however are very different.

The data for the Vertical Motion, Force Graphs, Acceleration Graphs, and Velocity Graphs clusters (Figure F.2–Figure F.4, & Figure F.5) each show little difference from year to year. The data from the Force Sled cluster (Figure F.1) show improvement from 2003–2004, but that improvement is diminished in 2005. The most significant results, however, are found in the data for the Newton III and Energy clusters.

The data for the Newton III cluster depicted in Figure F.6 & Figure F.7 show almost no gain in Newtonian reasoning during the 2003 semester. It appears as though some students abandoned their action dependence models, but many of these students did not move to a Newtonian view as evidenced by the almost horizontal change from pretest to post-test data, i.e. many students abandoned their previously held incorrect models but did not fully accept the appropriate Newtonian model. In either of the experimental groups for the Newton III study.
2004 and 2005, however, much better results are depicted. In fact by the end of each of these semesters, virtually none of the students were using the mass dependence model and most of them were using Newtonian reasoning in some capacity or another. It is of particular interest that these are the years in which the TIP Newton III tutorial was replaced by the ABT and OPT versions for some students.[34]

Within the data for the Energy cluster we can see another interesting change. The 2003 and 2004 data are practically identical. This is interesting considering the fact that a significant difference was found between the 2004 students who had used the TIP tutorial and those who had used the locally developed tutorial.[65] This may indicate that the half that performed worse in 2004 also performed worse than the class average in 2003. The interesting aspect of the Energy data comes in 2005 which displays significantly less gain in students’ conceptual development than either 2003 or 2004. This loss occurred in spite of all teaching methods being the same as in 2004.

4.3 Less Common Models

One of the more interesting aspects of the data analysis for this study was the discovery of additional common student mental models. These models would become apparent when a particular pattern of responses (e.g. a-b-d-g-d for the Force Sled cluster) was observed repeatedly. It then became my goal to determine what type of reasoning these responses might represent and if they were common enough to be in-
cluded in my analysis of the data. This section describes several of these “less common” models and the responses with which they correspond.

The first of these less common models is that described for the Force Sled cluster above: a-b-d-g-d (questions 1–4, and 7). The “d” responses for questions 3 and 7 indicate the idea that no force is needed for an object to slow down. This in turn corresponds with Aristotelean notion that all objects come to rest. In students’ everyday observations they notice that objects being pushed eventually slow to a stop if the pushing agent is removed. For the most part students abandon this idea when asked about a frictionless environment, but some do not, particularly during pre-instruction assessment.

Another of these less common models is found in the data for the Vertical Motion cluster. On questions 27–29, students will occasionally respond c-d-f or b-d-f. These responses indicate that there is a downward acceleration while the coin goes up, zero acceleration when the coin is at the peak of its path and an acceleration upward when the coin is going back down. The “c” or “b” response for question 27 could be related to another less common model in the Force Sled cluster in which the students respond a-b-e-g-c. In these cases, the questions 3, 7, and 27 present an object that is slowing down. These responses indicate that the acceleration (or force in the Force Sled cluster) is opposite the direction of motion. This idea in and of itself is correct, but these responses also indicate that the acceleration/force is decreasing (increasing in the case of “b” for question 27). This seems to be a combination between the cor-
rect Newtonian idea (in terms of direction) and the common model of force/acceleration-follows-velocity (in terms of magnitude).

Table 4.4: Percentages of students who use each of the described “less common” models on pre-instruction assessments.*–The response indicating a force-follows-velocity model has been included for comparison purposes.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Response indicative of model</th>
<th>Percentage of usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–4, 7</td>
<td>a-b-c-g-e*</td>
<td>49.8%</td>
</tr>
<tr>
<td>1–4, 7</td>
<td>a-b-d-g-d</td>
<td>4.7%</td>
</tr>
<tr>
<td>27–29</td>
<td>c/b-d-f</td>
<td>12.5%</td>
</tr>
<tr>
<td>1–4, 7</td>
<td>a-b-e-g-c</td>
<td>1.9%</td>
</tr>
</tbody>
</table>

The “f” response for question 29, however, reveals an interesting perspective on this combination. In this case, the direction of the acceleration is still opposite the direction of motion, but the magnitude agrees with the acceleration-follows-velocity model. It is possible that these students are considering the fact that air resistance is always opposite to the direction of motion while responding to question 29, but it is yet unclear why they would choose response “f.” Table 4.4 displays the percentages of students that typically hold each of these less common models according to the pre-instruction data. The percentage of students who hold the force-follows-velocity model on the Force Sled cluster is shown for comparison.
CHAPTER 5

DISCUSSION OF RESULTS

The data that are reported in the previous chapter raise many questions as to the causes in the differences between each of the years. One of the most obvious observations is the marked difference in the data from the College of DuPage with those from the University of Maine. With this being the case, the discussions of the results from each of these institutions have been appropriately separated. Additionally, the data from each of the courses (the algebra-based and the calculus-based) at the College of DuPage will be discussed individually.

5.1 College of DuPage Algebra-based Course

As reported in section 4.1.1 and shown in Appendix D the data from the COD algebra-based course show that the instructional methods for each year yielded nearly identical results in terms of student learning as measured by the FMCE. The data for some clusters, however, yield much more interesting results.

Data from the Newton III cluster (Figure D.6 & Figure D.7) display significant trends for both the mass dependence and action dependence models during 2000 and 2001. The data for this cluster show that the methods used from 2001–2004 were more effective than those used in 2000. The reader will also note that the data on the New-
ton III cluster from 2001–2004 are virtually identical leading to the conclusion that
the curricular modifications during those years (the implementation of the TYCHO
homework system in 2002, course withdrawal policy changes in 2003, and textbook
switch in 2004) had little to no effect on students’ learning of Newton’s third law.
This also leads to the assertion that the changes made between 2000 and 2001 caused
this gain in students’ conceptual development. However, since so many changes were
implemented in 2001 (RTP labs instead of TST, required textbook reading, etc.), it is
difficult to determine precisely which of the modifications is responsible for students’
increased conceptual gains. Several of these factors, though seem more plausible than
others. For instance, the increased number of laboratory sessions (from 10 in 2000 to
14 in 2001), the greater time allowed during each laboratory session (due to the group
problem-solving sessions being moved into the lecture periods), and the required text-
book reading most likely did not cause the observed gain in students’ understanding
of Newton’s third law. These curricular modification affect the course in a very broad
manner. It is possible that each of these changes could have significantly impacted
the amount of time the students spent learning Newton’s third law, but it is likely
that the benefits of these particular changes would have been observed throughout all
aspects of the course. That is, if these changes were the cause of students’ increased
conceptual gains within the realm of Newton’s third law, these same improvements
should have been observed on all clusters of the FMCE, which is not the case. With
this in mind, the most likely candidate for students’ improved understanding of New-
ton’s third law is the switch from the TST to the RTP laboratory materials. While
the switch to the RTP materials, like previously mentioned modifications, does affect
the course in a very broad manner, the RTP labs can be seen as a series of modifica-
tions, one for each laboratory session. This was not a change in the structure of the
course, but an alteration of the content of each of the laboratory sessions. Work by
Smith and Wittmann[34] and Abbott, Saul, Parker, and Beichner[67] has shown that
changing a single lesson or laboratory session can have a dramatic effect on students’
success in learning the material in question. It is likely that the change from the TST
to RTP lab activity for learning Newton’s third law, caused students to gain a better
understanding of the physics.

Another aspect shown within the data from the Newton III cluster is the tendency
for the afternoon section to outperform the morning section in 2000. Furthermore,
this trend is replicated throughout most of the clusters in the FMCE (particularly the
Vertical Motion cluster, Figure D.2). This result is particularly interesting considering
the fact that the instruction for each of those sections was identical. Why then, do
the afternoon students seem to learn more than the morning students? One cause
could be the instructor’s familiarity with the instructional materials and techniques.
The TST laboratory materials were first implemented in 2000, and, like any teacher
trying out a new tool, the instructor of the course most likely experienced some
difficulties and unexpected pitfalls at first. The afternoon session in each case would
reap the benefits of the misfires that had occurred earlier in the morning. Another
factor that could contribute to the difference between the morning and afternoon
sections is the student population of each. Students have reasons for choosing to
take physics in either the morning or the afternoon. It is possible that the afternoon students possessed a higher aptitude for learning physics. It is also possible that the afternoon students were more alert and prepared to learn physics than their morning counterparts. If, however, the students’ alertness truly did affect their ability to learn physics, we would expect to observe a similar trend in the data from 2001. Instead, the morning and afternoon data from 2001 are virtually identical for all clusters of the FMCE, leading to the assertion that the distinguishing factor between the morning and afternoon sections in 2000 is the instructor’s familiarity with the tools being employed within the class. This being said, it may also be reasonable to expect that the instructor’s familiarity with the TST materials may carry over to the RTP versions (as they are designed by the same researchers), causing another increase in student conceptual gains from 2000 to 2001, as seen in the Newton III cluster. As such, studies should be conducted comparing the TST and RTP laboratory activities for teaching Newton’s third law (and all other aspects of the curricula) to determine the differences in their structure and effectiveness in terms of student learning.

Data from the Vertical Motion cluster (Figure D.2) also display very interesting results. As mentioned previously, a profound difference exists between the morning and afternoon sections in 2000. In fact, the afternoon section displays as much conceptual gain as the 2001, 2003, and 2004 classes. The similarities between these years provide support for the likelihood that instructor familiarity has a significant impact on student learning. The Vertical Motion data also show a substantial drop in students’ conceptual gains in 2002 (as compared to the 2000-afternoon section and both
sections in 2001). Since the only significant instructional change in 2002 was the implementation of a web-based homework system (TYCHO), it is reasonable to assume that the students’ use of TYCHO hindered their learning of vertical motion concepts. This trend, however, is not observed in the data from any of the other question clusters. Moreover, in 2003 and 2004, students displayed at least as much conceptual gain on the Vertical Motion cluster as the 2000 and 2001 students, suggesting that whatever had caused students’ poor performance in 2002 was rectified the following year. All factors considered, the low achievement of the 2002 students on the Vertical Motion cluster seems insufficient to consider TYCHO a hazard to student learning. In fact, in the vast majority of cases, students who took the COD algebra-based course in years in which TYCHO was implemented displayed as much conceptual gain on the FMCE as students who had taken the course using homework questions from the Cutnell & Johnson text.[56] This result suggests that the poor performance of the 2002 students on the Vertical Motion cluster may have also stemmed from a lack of familiarity with the TYCHO system as it pertains to vertical motion concepts on the part of the instructor. It is obvious from the data, however, that any effect that a lack of familiarity may have had in 2002 was completely remedied in 2003.

Examining the data from the morning and afternoon sections of the algebra-based course in 2001 yield one striking result: they are virtually identical across the board. The afternoon section in 2001 saw the implementation of Peer Instruction within the lecture periods. Peer Instruction has been shown to be a very effective form of physics instruction,[59] but the data from COD shows no improvement. Before
dismissing Peer Instruction, however, we must consider the fact that data from the morning section in 2001 display incredibly substantial gains. This being the case, it is very possible that the effects of the Peer Instruction are nearly undetectable when accompanied by the effects of implementing the RTP curriculum and requiring group problem-solving within lecture periods.

### 5.2 College of DuPage Calculus-based Course

The data from the calculus-based course at COD display more consistent trends than those from the algebra-based course. As discussed in section 4.1.2, the data from the 2002 course show that these students did not improve as much throughout the course as their 2003 and 2004 counterparts. Furthermore, the data from 2003 and 2004 are almost identical in terms of student improvement throughout the course. The main difference between these years appears to be students’ initial understanding of physics. In fact, from 2002 to 2004 students’ initial knowledge base as measured by their pre-instruction scores on the FCI steadily improves. Coletta and Phillips report findings that students’ initial understanding of physics can have a substantial impact on learning.\[^{[68]}\] In most cases, they found a significant, positive correlation between students’ pretest scores on the FCI and their normalized gains.\[^{1}\] This correlation was also observed for class average pre-instruction scores and average normalized gains. The results of Coletta and Phillips’ research suggest that the differences in conceptual gain from 2002 to 2004 in the calculus-based course at COD could be

\[^{1}\text{See section 2.4.2.}\]
more of an effect of the students’ initial understanding of physics rather than the result of any particular curricular changes.

Another aspect to consider is the absence of data prior to 2002. Comparing Table 4.1 and Table 4.2 shows that the curricular changes made to the calculus-based course in 2002 and 2003 are almost identical to those made to the algebra-base course in 2001 and 2002, respectively. Within the algebra-based course, the most interesting results occurred between 2000 and 2001 with the implementation of RTP laboratory materials and problem-solving sessions during lectures. Unfortunately, the first set of data from the calculus-based course describes a year in which these educational strategies had already been implemented. As such, it is difficult to assess what effects the curricular changes carried out in 2002 may have had from previous years. Additionally, considering the results reported by Coletta and Phillips,[68] the curricular changes for which data do exist (implementation of web-based homework systems TYCHO and CyberTutor and change in text from Halliday, Resnick, & Walker[62] to Cummings, Redish, & Laws[64]) appear to have had very little effect on students’ learning of introductory mechanics. Moreover, the data displayed in Appendix E show that the majority of students from all three years leave the course with a very good understanding of physics, as indicated by data points near the upper, left-hand corner of the model plot.

The final factor to consider regarding the data from the calculus-based course is one pertaining to the incredible gain on understanding of Newton’s third law shown in 2003 when every student responded correctly to all of the Newton III questions.
Upon communicating with the instructor of these courses it became apparent that during 2003 he ran an informal experiment himself in which he repeatedly told the students that interacting objects always exert equal forces on each other. In essence he stated Newton’s third law for them over and over again. This tactic obviously had an effect. A question that could be posed is whether or not these students truly understand Newton’s third law. One could argue that a student who realizes that interacting objects exert equal and opposite forces for all circumstances does, in fact, understand Newton’s third law. It is undeniable that the students from the 2003 calculus-based course at COD can state Newton’s third law and choose responses on the FCI that indicate equal and opposite interaction forces; but can they apply Newton’s third law in a more in depth manner?

The instructor of the calculus-based course at COD has recently conducted a study to answer this particular question. He noticed that many of his students, even at the beginning of the course, were choosing the correct responses for the Newton III questions on the FCI. He decided to add two questions to the post-instruction version of the Newton III portion of the FCI. The first of these questions presents a situation in which a ball undergoes a perfectly elastic collision with the floor and asks the students to compare the force on the ball due to gravity with that due to the floor during the collision. This question has nothing to do with Newton’s third law, since it only asks to compare the forces on a single object (the ball), but the question was surrounded by questions that did directly relate to Newton’s third law. The instructor wanted to find out if the students would still respond that the forces in question are
“equal and opposite” even though that response would be incorrect. The data has not been fully analyzed, but the preliminary results indicate that the majority of students do, in fact, respond that the forces on the ball are “equal and opposite.” This fascinating result begs the question, can students understand Newton’s third law as well as a physicist without first developing a robust understanding of Newton’s second law? For that matter, how does a student’s understanding of the conservation of momentum throughout a collision affect his understanding of Newton’s third law? And finally, can students’ responses that interacting objects exert equal and opposite forces be trusted to indicate understanding if these students do not respond correctly in other physics contexts? These issues should be studied in depth to help confirm or undermine the validity of assessments used by researchers and educators.

5.3 University of Maine

The data from UMaine displayed in Appendix F are similar to those from the algebra-based course at COD in that the curricular changes made each year yield few trends in terms of differences in students’ conceptual development between the years of the study, and none that span the entirety of the physics content assessed by the FMCE. In fact, the only feature of the data that is consistent throughout the clusters is that the students from 2004 appear to perform better on post-instruction assessments than the students from either of the other two years. The only exception to this statement is seen in Figure F.2 (the Velocity Graphs cluster) for which the 2005 data are slightly superior to those from 2004. This particular cluster, however, shows
very strong similarities between the data sets from all three years making the small
differences almost meaningless. With so few trends observed, the data from UMaine
must be analyzed on a cluster-by-cluster basis.

The data from the Vertical Motion, Force Graphs, Acceleration Graphs, and Ve-
locity Graphs clusters show very little difference from year-to-year. In each of these
clusters, the students seemed to start and end the semester with approximately the
same understanding of physics. In fact, it is particularly interesting to note that the
pre-instruction data for all of the clusters on the FMCE are remarkably similar. This
fact supports the implicit assertion that students who take PHY 111 at UMaine are
similar each year in terms of their initial understanding of physics.

The remaining clusters from the UMaine data, however, do display trends that
students’ conceptual gains change from year-to-year. In fact, the data from 2004 on
the Force Sled cluster shown in Figure F.1 are noticeably better than those from 2003.
The Force Sled data from 2005 also show better conceptual gains than the 2003 data,
but less conceptual gain than the 2004 data. It is possible that the increase in 2004
stems from the introduction of ILDs into the curriculum. Thornton and Sokoloff have
shown that students’ understanding of the material tested specifically by the Force
Sled cluster will improve significantly with the use of ILDs.[42] These results were
obtained using only ILDs pertaining to kinematics without including those designed
to supplement dynamics instruction. As such, it is likely that the ILDs used in
2004 helped those students learn the Force Sled concepts more effectively than their
predecessors. Unfortunately, this explanation will not account for the drop in the
data in 2005. It is unlikely that the increased usage of ILDs reported in 2005 would cause student understanding to decrease. We must, therefore, examine other aspects of the course in 2005 to determine the cause of this loss of conceptual development.

As discussed in section 4.2, 2005 saw the implementation of tutorials from the OPT set of materials for many aspects of the PHY 111 course as part of a pilot study being conducted at the University of Maryland. These tutorials emphasize students’ development of their own epistemological beliefs and views toward science rather than focusing primarily on physics content. The tasks asked of students within the OPT materials are very different from those expected of students while using a corresponding tutorial from the TIP set. Students were required to switch between the types of tutorials as frequently as once a week during some portions of the course. Reports from teaching assistants, as well as the primary instructor of the course, indicated a high level of unrest and confusion among the students concerning what was expected of them during tutorial periods. It was difficult for the students to feel comfortable in a routine when that routine was constantly being altered in terms of the activities being completed and the types of thinking being asked of them (from comparing free body diagrams to asking themselves how they know a certain fact to be true). This general feeling of unrest throughout the student population in 2005 may have contributed to the decrease in conceptual development from 2004 depicted in the data for the Force Sled cluster.

The data for the Newton III cluster display an even more significant trend than those found on the Force Sled cluster. The students in 2003 displayed almost no
gain in understanding of Newton’s third law throughout the semester. In 2004 and 2005, however, significant improvements are observed. As reported in section 4.2, OPT and ABT materials were implemented for some students in place of the TIP versions during the Newton III tutorial periods for these years. A study on the effects of this implementation in 2004 showed that the OPT tutorials affected a much larger conceptual gain than either of the other versions.[34]

A study conducted within a workshop for in-service elementary school teachers also yielded results that the Newton III tutorial from the OPT set of materials can affect significant gains in understanding, even over short periods of time.[69] In this particular study, in-service teachers participated in a week-long workshop with the goal of better understanding physics so as to become more effective science teachers. All of the participants engaged in the OPT Newton III tutorial for learning Newton’s third law. By the end of the workshop, the in-service teachers in the study displayed remarkable gains in understanding of Newton’s third law, as measured by the FMCE, but these gains were absent from their understanding of Newton’s first or second law assessed with other questions on the FMCE. As discussed with the calculus-based students at COD, however, we must wonder whether these in-service teachers have a robust understanding of Newton’s third law or if they have been superficially trained to respond that all interaction forces are “equal and opposite.” Further studies must be conducted to determine the validity of students’ responses to the Newton III cluster, as it may be impossible to determine the answer to this question using data from the FMCE alone.
Further investigation of the current data from the PHY 111 course pertaining to the Newton III cluster should also be conducted to determine whether or not the implementation of the ABT and OPT materials truly was the cause of the increase in conceptual development from 2003 to 2004. The results of Smith and Wittmann’s study that discusses the beneficial effects of the OPT Newton III materials show that significant differences existed between the students who had used each type of tutorial, in terms of student learning of Newton’s third law. This being the case, it is difficult to determine what factor(s) may have caused the average conceptual development of all the students in PHY 111 to rise so drastically in 2004. The first step to examining this quandary more closely is to separate the data from 2004 to create three sets that correspond to the students who used each of the three tutorials. The data for TIP students must be compared with the data from 2003 (when all students used TIP version) to determine whether a significant difference exists between students who went through the same tutorial instruction. This process should be replicated for the 2005 data so that comparisons can be made for all students based on the tutorial they completed between 2004 and 2005. These comparisons may help illuminate the causes behind the difference in students’ average conceptual development between 2003 and 2004.

The Energy cluster also displays an interesting result. The study in 2004 by Dudley and Bianchi showed that the students who had used a tutorial on energy conservation and equipotentials developed at UMaine displayed a better understanding of energy concepts than the students who had used the TIP version.[65] The average
over all of the students, however, shows approximately the same gain in conceptual understanding as the students from 2003 (see Figure F.8). The students from 2005, however, show significantly less improvement. This result is quite curious considering the fact that each version of the energy tutorial was once again administered to half of the students in 2005. If, however, we consider other aspects of the tutorial portion of the PHY 111 course in 2005, the reasoning behind this result becomes a bit clearer. As mentioned previously, the students in 2005 felt a high level of discomfort during tutorial sessions resulting from frequent changes between TIP and OPT activities and types of thinking. When students were then given a tutorial that did not conform with either of the previously established models, they most likely became more confused and less engaged due to the lack of consistency. Students in 2004 had completed tutorials from the TIP almost exclusively throughout the course. As such, they may have been surprised by the change of attitude of the UMaine tutorial, but they were most likely able to handle the transition with little stress and effort. The 2005 students, however, had already been required to change their expectations repeatedly throughout the semester, and with energy being one of the later topics covered, they may have become overly frustrated with the prospect of learning a completely new set of expectations. This being the case, it is likely that the new tutorial structure had adverse effects on student learning of energy concepts in 2005.
5.4 Comparing Results Between Institutions

One of the most obvious aspects of the data is the marked differences between the results from each of the courses in this study. Of particular interest are the differences between the algebra-based course at COD and the PHY 111 course at UMaine. Since both are algebra-based courses and serve similar non-physicist populations, it would be expected that the two courses would yield similar results when research-based instruction is implemented. Looking at the data, however, we can easily see that this isn’t the case. The data from the algebra-based course at COD display much larger conceptual gains than those from UMaine. Furthermore, it is also easy to see that the pre-instruction data from each of these courses each year are virtually identical. As such, it is reasonable (almost necessary) to wonder why the data are so different between the institutions.

The first difference to consider is the stability of the instructor at COD. One instructor has been responsible for teaching all aspects of the course throughout the entirety of this study. He chose the curricular modifications that were implemented, and he was personally able to implement them to the best of his ability. The instructors at UMaine, on the other hand, are continually in flux. None of the instructors (neither the primary instructor, nor the TAs) taught for more than two years during the span of the study. Additionally, the primary instructor in 2003 (the first year of the study) was not the same person as had taught the course 2002. In fact, the primary instructor for PHY 111 has almost constantly been changing since 1999. The
primary instructor in 2004 and 2005, however, was a senior PER graduate student who had prior experience teaching high school and was well versed in research-based curricula. Even so, each instructor (primary or assistant) at UMaine is typically responsible for only one aspect of the course (lecture, tutorial, or laboratory) adding a degree of instability. Occasionally, a TA will be in charge of both tutorial and laboratory sections, but this is not always the case.

Smaller class sizes at COD may have also contributed to increased student learning. The combination of small class sizes (between 30 and 50 students) and greater consistency in terms of instructor throughout a week of class time (six hours per week with a single instructor at COD compared with an average of two hours per week per instructor at UMaine) may have served to create a social environment that is very conducive to student learning. Many aspects of research-based curricula (tutorials, peer instruction, laboratory modifications) advocate Vygotsky’s ideas that learning is a social experience and should be done with others. The personal relationships fostered between instructors and students could be vital to students’ success in learning physics. These relationships could easily be enhanced through the smaller class sizes and greater number of contact hours that occur within the courses at COD. An additional effect of the course structure at COD may be that students gain a better understanding of science as a whole, i.e. they hold more of a physicist’s view of physics as a coherent set of knowledge and understanding rather than a collection of isolated facts and formulas. This effect would best be assessed using the *Maryland Physics Expectations Survey* or the Epistemological Beliefs Assessment for Physical Science.
discussed in section 2.1.4. With all of the differences between the course structures at COD and UMaine, it is nearly impossible to separate each of the factors to determine why the results from the two institutions are so different.

Another possible explanation of UMaine’s apparent lack of success is the use of *Tutorials in Introductory Physics* as the basis of its research-based curriculum. The TIP materials have been shown to be a very effective supplement to lecture-based instruction, but they are intended for a calculus-based course. As a result, the assumptions that are implicitly made about the student population may be unavoidably detrimental to their use in an algebra-based course in which the population consists primarily of life science and earth science majors rather than future physicists and engineers. One possible remedy to this dilemma would be the complete implementation of the OPT curriculum. These tutorials, however, with their focus on epistemological development over physics content, have not yet been tested to determine whether they are as effective as other research-based curricula. As such, caution must be exercised before making a wholesale switch away from the TIP materials.
CHAPTER 6

CONCLUSIONS

Many of the changes made to the courses studied seem to have had little effect on students’ general understanding of introductory mechanics. The effects that can be observed from the data are mostly localized to specific physics content areas, especially those pertaining to the algebra-based course at the College of DuPage and the PHY 111 course at the University of Maine. One such example seen in the data from the algebra-based course at COD is the dramatic improvement in students’ conceptual gains from 2000 to 2001 on the topic of Newton’s third law. Starting in 2001, the RealTime Physics[57, 58] curriculum was implemented within the laboratory sessions in place of the Tools for Scientific Thinking[54, 55] curriculum, students were required to read portions of relevant textbook chapters before coming to lecture, and more time was available during laboratory periods due to group problem-solving sessions being moved into lecture periods. Studies showing the effectiveness of modifying a single lesson,[34, 67] however, support the conclusion that the implementation of the RTP curriculum most likely caused the observed increase in students’ conceptual gains in 2001. In 2002, the TYCHO web-based homework system replaced standard textbook homework assignments for the algebra-based course, and students’ gains in understanding of vertical motion concepts decreased. In 2003, however, students’
understanding of vertical motion was as high as or higher than it had been in 2000 and 2001. It is likely that the first time using the TYCHO system uncovered some unsuspected difficulties that may have adversely affected student learning of concepts pertaining to objects undergoing vertical motion. These difficulties, however, were overcome the following year as evidenced by the high level of achievement of the 2003 students.

Within the PHY 111 course at UMaine, Interactive Lecture Demonstrations were first implemented in 2004, and improvement was seen from 2003. This trend is particularly evident in the data from the Force Sled cluster but is minimal for other clusters, e.g. the Vertical Motion and Force Graphs clusters. Furthermore, the implementation of the Activity-Based Tutorials[45] and Maryland Open-source Physics Tutorials[46] in place of the Tutorials in Introductory Physics[43] for some students within the tutorial portion of the course for Newton’s third law in 2004 accompanied an incredible improvement in student learning of Newton III concepts. This level of accomplishment was seen in 2005 as well, when the ABT and OPT materials were used again for some of the students. In 2005, the primary instructor of the PHY 111 course began training the teaching assistants for both the tutorial and laboratory sessions, adding a greater level of coherence between the instructors for various aspects of the course. In previous years, TA coordination had been conducted by a different professor, not otherwise connected with the course. Also in 2005, additional tutorials from the OPT materials were implemented throughout the course, creating high levels of confusion and frustration among the students. This frustration can be
inferred from the data from the Force Sled and Energy clusters in the form of lower conceptual gains in 2005 as compared with 2004. This decrease in the Energy cluster may have been enhanced by the implementation of a tutorial created at UMaine that differs in style and method from both the TIP and OPT materials.

Despite the similarities in the types of trends observed for both courses, data from the algebra-based course at COD show significantly more conceptual development than those from the PHY 111 course at UMaine even though the content covered and initial students' understanding of physics are almost identical for the two courses. Several factors may have influenced these differences. First of all, the single-instructor set up at COD may have created more personal connections between the students and their instructor. The single instructor spent six hours each week with the students and was (obviously familiar) with the material that had been covered during each session. UMaine, on the other hand, uses many TAs for the tutorial and laboratory portions of PHY 111. Each instructor (including the primary lecturer) spends two hours each week with the students, and the TAs rarely have direct knowledge of what has been covered in sessions for which they are not responsible. Furthermore, while each TA at UMaine interacts with approximately 25 students at a time (comparable to class sizes at COD), the class as a whole consists of over 100 students, and populations of the laboratory and tutorial sections are independently determined meaning that a particular group of students only interacts with each other for approximately two hours of class time each week. These differences between the composition of the courses may have contributed to develop a greater sense of community within the algebra-
based course at COD. These students spent six hours each week with each other and their instructor, learning physics together. This social environment, combined with a highly motivated instructor who is familiar with the research-based curricula being implemented, may have enabled a much higher achievement in the students in that course. Additional studies should be conducted using the MPEX or EBAPS to determine if these students’ beliefs and attitudes toward the nature of physics, and science in general, also benefitted from the social environment established within the algebra-based course at COD.

The results from the calculus-based course at COD are very different from those corresponding to either the algebra-based course at that institution or the PHY 111 course at UMaine. First, students begin the calculus-based course with a good understanding of physics. Furthermore, this high level of initial understanding increases throughout the years of the study. In virtually all content areas assessed, the students from 2004 begin the year with a better understanding of physics than their 2003 counterparts, who in turn had a higher initial understanding than the students in 2002. The data also show that the students in 2003 and 2004 (when the web-based homework systems TYCHO and CyberTutor[63] were implemented) make greater conceptual gains throughout the course than the students from 2002. Studies by Coletta and Phillips,[68] however, show that students’ normalized gains on the FCI are directly related to their pre-instructional scores. This relation also holds when discussing the average gains of an entire class compared to the average initial score, suggesting that the improved conceptual development of the 2003 and 2004 students
could be due to the students’ heightened initial understanding rather than a superiority of the instructional methods used.

More interesting results pertaining to the data from the calculus-based course can be seen in the Newton III cluster in 2003 in which all students answered perfectly on the post-instruction FCI. As discussed in Chapter 5, preliminary results from current studies being conducted within the calculus-based course at COD suggest that students may be cued by a collision or pushing scenario to respond that two forces are “equal and opposite” regardless of whether or not the forces in question constitute an interaction pair. These results emphasize the question; do students who always answer “equal and opposite” to appropriate Newton III scenarios have a true understanding of Newton’s third law? An additional study that may shed light on this quandary would be to ask students to answer a series of Newton III questions once as themselves and again as they think their professors would respond. This method mimics studies conducted at the University of Maryland pertaining to students views of the nature of physics as measured by the MPEX.[19]

One of the biggest conclusions from my research is the necessity to examine data from the FCI and FMCE in terms of carefully defined clusters of questions that correspond to specific correct and incorrect mental models. Throughout the data from all three courses, improvement from year-to-year (when observed) is localized within a handful of question clusters. If the data from this study were analyzed by solely examining students’ pre- and post-instruction scores over an entire assessment, many of the trends and intricacies of the results would have been undetectable. Only by
carefully examining students’ responses to determine the manners in which they are thinking (both correct and incorrect), and comparing students’ responses on particular questions to their responses on similar questions, can we determine how each instructional factor affected the ways in which students learned.

Using model analysis facilitated the examination of students’ mental models and comparison of responses between questions, as both are required for the analysis. Model analysis provides a method for concisely executing both of these elements. The implementation of model analysis, however, is not an easy task. Each question must be carefully analyzed to determine which mental models would be indicated by a student’s choice of each possible response. Questions must then be grouped into clusters based on the mental models which may be used while choosing a response. The FMCE has proved to be much more conducive to model analysis than the FCI. Primarily, the FMCE is grouped into similar-question clusters by the nature of its structure. Groups of questions pertaining to the same type of physical concept are presented with a bank of choices which may be used for each question in a group. This structure makes the clustering of questions very simple. Furthermore, most questions on the FMCE are conducive to the use of two common mental models: the correct model, and a common misconception. The questions on the FCI, however, often lend themselves to the use of three models, i.e. the correct model and two common misconceptions. The more complicated nature of these responses makes results from the FCI more difficult to analyze using model analysis. Additionally, the questions on the FCI are not grouped into nice clusters. The broader range of topics covered on the
FCI creates questions that are less similar to one another than the questions within the FMCE. Table 3.1 shows the vast number of misconceptions confronted within the FCI. As such, the clustering of questions becomes quite difficult. Reflecting on my own analysis, I feel as though my choice of clusters was excessively rigid, requiring that each question reside in one (and only one) cluster. Instead, I now recommend that all questions pertaining to a particular mental model (correct or incorrect) be analyzed together regardless of additional mental model(s) that may be elicited by each of the questions. In this manner a more complete picture can be formed of how the students in a class progress throughout the semester in terms of each particular mental model. The representation of such analyses using model plots may be very difficult or impossible. As such, this suggested modification may go beyond the capabilities of model analysis in its current form.


[27] Patrick B. Kohl and Noah D. Finkelstein. Student representational competence


[34] Trevor I. Smith and Michael C. Wittmann. Comparing the effectiveness of three methods for teaching Newton’s Third Law. conditionally accepted for publication.


APPENDIX A

The Force Concept Inventory

Table A.1: FCI—Original vs. Revised: This table shows how the questions on the FCI as described in ref [50] correspond with the revised version that was distributed to students at the College of DuPage. A listing of N/A indicates no corresponding question, and * indicates a question that is similar in content but different in form to its corresponding question. Furthermore, question 28 on the original version of the FCI was broken into two parts for the revised version (questions 25 and 26).

<table>
<thead>
<tr>
<th>Original</th>
<th>Revised</th>
<th>Original</th>
<th>Revised</th>
<th>Original</th>
<th>Revised</th>
<th>Original</th>
<th>Revised</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>9</td>
<td>11</td>
<td>17</td>
<td>3</td>
<td>25</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>10</td>
<td>6</td>
<td>18</td>
<td>17</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>11</td>
<td>28</td>
<td>19</td>
<td>N/A</td>
<td>27</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>12</td>
<td>29*</td>
<td>20</td>
<td>19</td>
<td>28</td>
<td>25, 26</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>13</td>
<td>15</td>
<td>21</td>
<td>20</td>
<td>29</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>14</td>
<td>16</td>
<td>22</td>
<td>30*</td>
<td>30</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>15</td>
<td>N/A</td>
<td>23</td>
<td>14</td>
<td>N/A</td>
<td>5</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>16</td>
<td>12</td>
<td>24</td>
<td>21</td>
<td>N/A</td>
<td>18</td>
</tr>
</tbody>
</table>

Due to copyright reasons, the Force Concept Inventory is not available in the online version of this thesis. If you are interested in obtaining a copy of the modified FCI, please contact me at: Trevor.I.Smith@umit.maine.edu
APPENDIX B

The Force and Motion Conceptual Evaluation

A sled on ice moves in the ways described in questions 1-7 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.

You may use a choice more than once or not at all but choose only one answer for each blank. If you think that none is correct, answer choice J.

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).

1. Which force would keep the sled moving toward the right and speeding up at a steady rate (constant acceleration)?
2. Which force would keep the sled moving toward the right at a steady (constant) velocity?
3. The sled is moving toward the right. Which force would slow it down at a steady rate (constant acceleration)?
4. Which force would keep the sled moving toward the left and speeding up at a steady rate (constant acceleration)?
5. 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?
6. The sled is slowing down at a steady rate and has an acceleration to the right. Which force would account for this motion?
7. The sled is moving toward the left. Which force would slow it down at a steady rate (constant acceleration)?
Questions 8-10 refer to a toy car which is given a quick push so that it rolls up an inclined ramp. After it is released, it rolls up, reaches its highest point and rolls back down again. *Friction is so small it can be ignored.*

![Diagram of a car on an inclined ramp]

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. Answer choice J if you think that none is correct.

**A** Net constant force down ramp  
**B** Net increasing force down ramp  
**C** Net decreasing force down ramp  
**D** Net force zero  
**E** Net constant force up ramp  
**F** Net increasing force up ramp  
**G** Net decreasing force up ramp

8. The car is moving up the ramp after it is released.  
9. The car is at its highest point.  
10. The car is moving down the ramp.

Questions 11-13 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 the cases described below. Answer choice J if you think that none is correct. *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

11. The coin is moving upward after it is released.  
12. The coin is at its highest point.  
13. The coin is moving downward.
Questions 14-21 refer to a toy car which can move to the right or left along a horizontal line (the positive part of the distance axis).

Assume that friction is so small that it can be ignored.

A force is applied to the car. Choose the one force graph (A through H) for each statement below which could allow the described motion of the car to continue. You may use a choice more than once or not at all. If you think that none is correct, answer choice J.

14. The car moves toward the right (away from the origin) with a steady (constant) velocity.

15. The car is at rest.

16. The car moves toward the right and is speeding up at a steady rate (constant acceleration).

17. The car moves toward the left (toward the origin) with a steady (constant) velocity.

18. The car moves toward the right and is slowing down at a steady rate (constant acceleration).

19. The car moves toward the left and is speeding up at a steady rate (constant acceleration).

20. The car moves toward the right, speeds up and then slows down.

21. The car was pushed toward the right and then released. Which graph describes the force after the car is released.

J None of these graphs is correct.
Questions 22-26 refer to a toy car which can move to the right or left on a horizontal surface along a straight line (the + distance axis). The positive direction is to the right.

Different motions of the car are described below. Choose the letter (A to G) of the acceleration-time graph which corresponds to the motion of the car described in each statement.

You may use a choice more than once or not at all. If you think that none is correct, answer choice J.

22. The car moves toward the right (away from the origin), speeding up at a steady rate.
23. The car moves toward the right, slowing down at a steady rate.
24. The car moves toward the left (toward the origin) at a constant velocity.
25. The car moves toward the left, speeding up at a steady rate.
26. The car moves toward the right at a constant velocity.
Questions 27-29 refer to a coin that 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 acceleration of the coin during each of the stages of the coin's motion described below. Take up to be the positive direction. Answer choice J if you think that none is correct.

A. The acceleration is in the negative direction and constant.
B. The acceleration is in the negative direction and increasing
C. The acceleration is in the negative direction and decreasing
D. The acceleration is zero.
E. The acceleration is in the positive direction and constant.
F. The acceleration is in the positive direction and increasing
G. The acceleration is in the positive direction and decreasing

27. The coin is moving upward after it is released.
28. The coin is at its highest point.
29. The coin is moving downward.

Questions 30-34 refer to collisions between a car and trucks. For each description of a collision (30-34) below, choose the one answer from the possibilities A through J that best describes the forces between the car and the truck.

A. The truck exerts a greater amount of force on the car than the car exerts on the truck.
B. The car exerts a greater amount of force on the truck than the truck exerts on the car.
C. Neither exerts a force on the other; the car gets smashed simply because it is in the way of the truck.
D. The truck exerts a force on the car but the car doesn't exert a force on the truck.
E. The truck exerts the same amount of force on the car as the car exerts on the truck.
F. Not enough information is given to pick one of the answers above.
J. None of the answers above describes the situation correctly.

In questions 30 through 32 the truck is much heavier than the car.

30. They are both moving at the same speed when they collide. Which choice describes the forces?
31. The car is moving much faster than the heavier truck when they collide. Which choice describes the forces?
32. The heavier truck is standing still when the car hits it. Which choice describes the forces?
In questions 33 and 34 the truck is a small pickup and is the same weight as the car.

_____ 33. Both the truck and the car are moving at the same speed when they collide. Which choice describes the forces?
_____ 34. The truck is standing still when the car hits it. Which choice describes the forces?

Questions 35-38 refer to a large truck which breaks down out on the road and receives a push back to town by a small compact car.

Pick one of the choices A through J below which correctly describes the forces between the car and the truck for each of the descriptions (35-38).

A. The force of the car pushing against the truck is equal to that of the truck pushing back against the car.
B. The force of the car pushing against the truck is less than that of the truck pushing back against the car.
C. The force of the car pushing against the truck is greater than that of the truck pushing back against the car.
D. The car’s engine is running so it applies a force as it pushes against the truck, but the truck’s engine isn’t running so it can’t push back with a force against the car.
E. Neither the car nor the truck exert any force on each other. The truck is pushed forward simply because it is in the way of the car.
J. None of these descriptions is correct.

_____ 35. The car is pushing on the truck, but not hard enough to make the truck move.
_____ 36. The car, still pushing the truck, is speeding up to get to cruising speed.
_____ 37. The car, still pushing the truck, is at cruising speed and continues to travel at the same speed.
_____ 38. The car, still pushing the truck, is at cruising speed when the truck puts on its brakes and causes the car to slow down.
39. Two students sit in identical office chairs facing each other. Bob has a mass of 95 kg, while Jim has a mass of 77 kg. Bob places his bare feet on Jim’s knees, as shown to the right. Bob then suddenly pushes outward with his feet, causing both chairs to move. In this situation, while Bob’s feet are in contact with Jim’s knees,

A. Neither student exerts a force on the other.
B. Bob exerts a force on Jim, but Jim doesn’t exert any force on Bob.
C. Each student exerts a force on the other, but Jim exerts the larger force.
D. Each student exerts a force on the other, but Bob exerts the larger force.
E. Each student exerts the same amount of force on the other.
J. None of these answers is correct.

Questions 40-43 refer to a toy car which can move to the right or left along a horizontal line (the positive portion of the distance axis). The positive direction is to the right.

Choose the correct velocity-time graph (A - G) for each of the following questions. You may use a graph more than once or not at all. If you think that none is correct, answer choice J.

40. Which velocity graph shows the car moving toward the right (away from the origin) at a steady (constant) velocity?
41. Which velocity graph shows the car reversing direction?
42. Which velocity graph shows the car moving toward the left (toward the origin) at a steady (constant) velocity?
43. Which velocity graph shows the car increasing its speed at a steady (constant) rate?
A sled is pulled up to the top of a hill. The sketch above indicates the shape of the hill. At the top of the hill the sled is released from rest and allowed to coast down the hill. At the bottom of the hill the sled has a speed \( v \) and a kinetic energy \( E \) (the energy due to the sled's motion). Answer the following questions. **In every case friction and air resistance are so small they can be ignored.**

44. The sled is pulled up a steeper hill of the same height as the hill described above. How will the velocity of the sled at the bottom of the hill (after it has slid down) compare to that of the sled at the bottom of the original hill? Choose the best answer below.
   A. The speed at the bottom is greater for the steeper hill.
   B. The speed at the bottom is the same for both hills.
   C. The speed at the bottom is greater for the original hill because the sled travels further.
   D. There is not enough information given to say which speed at the bottom is faster.
   J. None of these descriptions is correct.

45. Compare the kinetic energy (energy of motion) of the sled at the bottom for the original hill and the steeper hill in the previous problem. Choose the best answer below.
   A. The kinetic energy of the sled at the bottom is greater for the steeper hill.
   B. The kinetic energy of the sled at the bottom is the same for both hills.
   C. The kinetic energy at the bottom is greater for the original hill.
   D. There is not enough information given to say which kinetic energy is greater.
   J. None of these descriptions is correct.

46. The sled is pulled up a higher hill that is less steep than the original hill described before question 44. How does the speed of the sled at the bottom of the hill (after it has slid down) compare to that of the sled at the bottom of the original hill?
   A. The speed at the bottom is greater for the higher but less steep hill than for the original.
   B. The speed at the bottom is the same for both hills.
   C. The speed at the bottom is greater for the original hill.
   D. There is not enough information given to say which speed at the bottom is faster.
   J. None of these descriptions is correct.

46a. Describe in words your reasoning in reaching your answer to question 46. (Answer on the answer sheet and use as much space as you need)

47. For the higher hill that is less steep, how does the kinetic energy of the sled at the bottom of the hill after it has slid down compare to that of the original hill?
   A. The kinetic energy of the sled at the bottom is greater for the higher but less steep hill.
   B. The kinetic energy of the sled at the bottom is the same for both hills.
   C. The kinetic energy at the bottom is greater for the original hill.
   D. There is not enough information given to say which kinetic energy is greater.
   J. None of these descriptions is correct.
APPENDIX C

Sample Data Analysis

Table C.1: Sample data from the Fall 2003 semester at UMaine

<table>
<thead>
<tr>
<th>Pre-instruction</th>
<th>Post-instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question number</td>
<td>Question number</td>
</tr>
<tr>
<td>Student</td>
<td></td>
</tr>
<tr>
<td>F03-005</td>
<td>a b c g e</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-030</td>
<td>a b c g e</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-042</td>
<td>a b c g e</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-047</td>
<td>a b d g d</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-065</td>
<td>b d f f b</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-080</td>
<td>a b c g e</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-090</td>
<td>b d f f b</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>F03-107</td>
<td>a b e g c</td>
</tr>
</tbody>
</table>

138
Table C.2: Categorization of responses into the three models: 1) Force is proportional to the change in velocity; 2) Force is proportional to velocity; and 3) Other. See section 3.2 for correlation between responses and models.

<table>
<thead>
<tr>
<th>Student</th>
<th>Pre-instruction</th>
<th>Post-instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>F03-005</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>F03-030</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>F03-042</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>F03-047</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>F03-065</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>F03-080</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>F03-090</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>F03-107</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Density Matrices

The “Sample” student shows how a density matrix is created from the coded responses. The variable $\eta_n$ is defined as the probability that a given student will use the $n$th model. These individual matrices are averaged to determine the class density matrix shown below.

$$
\begin{align*}
\text{Sample} & = \begin{bmatrix}
\sqrt{\eta_1^2} & \sqrt{\eta_1 \eta_2} & \sqrt{\eta_1 \eta_3} \\
\sqrt{\eta_1 \eta_2} & \sqrt{\eta_2^2} & \sqrt{\eta_2 \eta_3} \\
\sqrt{\eta_1 \eta_3} & \sqrt{\eta_2 \eta_3} & \sqrt{\eta_3^2}
\end{bmatrix} \\
F03-005-pre & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix} \\
F03-030-pre & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix} \\
F03-042-pre & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix} \\
F03-005-post & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix} \\
F03-030-post & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix} \\
F03-042-post & = \begin{bmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{bmatrix}
\end{align*}
$$
Below are displayed the density matrices that represents the knowledge state of the entire class of students pre-instruction and post-instruction. This matrix is created by summing the individual students’ density matrices and dividing each entry by the number of the students in the study.

Pre-instruction

\[
\begin{bmatrix}
0.25 & 0 & 0 \\
0 & 0.65 & 0.12 \\
0 & 0.12 & 0.10 \\
\end{bmatrix}
\]

Post-instruction

\[
\begin{bmatrix}
0.53 & 0.16 & 0.06 \\
0.16 & 0.43 & 0 \\
0.06 & 0 & 0.05 \\
\end{bmatrix}
\]
Eigenvalues, Eigenvectors and the Model Plot

To continue the analysis, we must find the eigenvalues of the class density matrix and their corresponding eigenvectors:

\[
\begin{array}{c}
\text{Eigenvalues} \\
0.68 & 0.25 & 0.08
\end{array}
\]

\[
\begin{pmatrix}
0.00 \\
0.98 \\
0.20
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 \\
0 \\
0
\end{pmatrix}
\]

\[
\begin{pmatrix}
0.00 \\
0.20 \\
-0.98
\end{pmatrix}
\]

\[
\begin{array}{c}
\text{Post – instruction} \\
\text{Eigenvalues} \\
0.65 & 0.32 & 0.04
\end{array}
\]

\[
\begin{pmatrix}
0.81 \\
0.58 \\
0.08
\end{pmatrix}
\]

\[
\begin{pmatrix}
0.57 \\
-0.81 \\
0.13
\end{pmatrix}
\]

\[
\begin{pmatrix}
-0.14 \\
0.06 \\
0.99
\end{pmatrix}
\]

We use the primary eigenvalues (0.68 and 0.65 for this data set) and their corresponding eigenvectors to create our model plot. The horizontal and vertical coordinates for the model plot are defined as:

\[
x = \sigma^2_{\mu} \nu^2_{2\mu}
\]

\[
y = \sigma^2_{\mu} \nu^2_{1\mu}
\]

where \( \sigma_{\mu} \) is defined as the \( \mu \)th eigenvalue and \( \nu^2_{i\mu} \) is the \( i \)th component of the \( \mu \)th
eigenvector. Furthermore, \(i = 1\) corresponds with the correct Newtonian model while \(i = 2\) corresponds with the most common incorrect model. For our current data, the pre-instruction coordinates would be given by:

\[
x = (0.68)^2 (0.98) = 0.45
\]

\[
y = (0.68)^2 (0.00) = 0.00
\]

and the post-instruction coordinates would be given by:

\[
x = (0.65)^2 (0.58) = 0.25
\]

\[
y = (0.65)^2 (0.81) = 0.34
\]

Using these coordinates, we can create our model plot of the data for the Force Sled cluster of the FMCE:

Figure C.1: Model plot for the sample data from the Force Sled cluster of the FMCE.
APPENDIX D

FMCE Data from the Algebra-based Course at the College of DuPage

Figure D.1: Results from the algebra-based course on the Force Sled cluster of the FMCE: 2000–2004
Figure D.2: Results from the algebra-based course on the Vertical Motion cluster of the FMCE: 2000–2004

Figure D.3: Results from the algebra-based course on the Force Graphs cluster of the FMCE: 2000–2004
Figure D.4: Results from the algebra-based course on the Acceleration Graphs cluster of the FMCE: 2000–2004

Figure D.5: Results from the algebra-based course on the Velocity Graphs cluster of the FMCE: 2000–2004
Figure D.6: Results from the algebra-based course on the Newton III cluster of the FMCE, Mass Dependence: 2000–2004

Figure D.7: Results from the algebra-based course on the Newton III cluster of the FMCE, Action Dependence: 2000–2004
Figure D.8: Results from the algebra-based course on the Energy cluster of the FMCE: 2001–2004
APPENDIX E

FCI Data from the Calculus-based course at the

College of DuPage

Figure E.1: Results from the calculus-based course on the After Impetus cluster of the FCI: 2002–2004
Figure E.2: Results from the calculus-based course on the Newton III cluster of the FCI, Mass Dependence model: 2002–2004

Figure E.3: Results from the calculus-based course on the Newton III cluster of the FCI, Action Dependence model: 2002–2004
Figure E.4: Results from the calculus-based course on the Circular Motion cluster of the FCI, Force in Direction of Motion model: 2002–2004

Figure E.5: Results from the calculus-based course on the Circular Motion cluster of the FCI, Force Outward from Center model: 2002–2004
Figure E.6: Results from the calculus-based course on the Constant Force–1D cluster of the FCI: 2002–2004

Figure E.7: Results from the calculus-based course on the Strobe Kinematics cluster of the FCI: 2002–2004
Figure E.8: Results from the calculus-based course on the Constant Force–2D cluster of the FCI, Overcoming Impetus model: 2002–2004

Figure E.9: Results from the calculus-based course on the Constant Force–2D cluster of the FCI, Straight Path model: 2002–2004
Figure E.10: Results from the calculus-based course on the Two Objects in Free-fall cluster of the FCI: 2002–2004
APPENDIX F

FMCE Data from the Algebra-based Course at the University of Maine

Figure F.1: Results from the PHY 111 course on the Force Sled cluster of the FMCE: 2003–2005
Figure F.2: Results from the PHY 111 course on the Vertical Motion cluster of the FMCE: 2003–2005

Figure F.3: Results from the PHY 111 course on the Force Graphs cluster of the FMCE: 2003–2005
Figure F.4: Results from the PHY 111 course on the Acceleration Graphs cluster of the FMCE: 2003–2005

Figure F.5: Results from the PHY 111 course on the Velocity Graphs cluster of the FMCE: 2003–2005
Figure F.6: Results from the PHY 111 course on the Newton III cluster of the FMCE, Mass Dependence: 2003–2005

Figure F.7: Results from the PHY 111 course on the Newton III cluster of the FMCE, Action Dependence: 2003–2005
Figure F.8: Results from the PHY 111 course on the Energy cluster of the FMCE: 2003–2005
Trevor I. Smith was born in Harrisburg, Pennsylvania on July 15, 1983. For the majority of his childhood, Trevor lived in Carlisle, Pennsylvania, and he graduated from Carlisle High School in 2001. After earning a Bachelor of Science degree in Physics from the University of Maine in the spring of 2005, he began his work in the Master of Science in Teaching program at that institution.

From 2002–2005, Trevor was a member of the percussion section of the Bluecoats Drum & Bugle Corps from Canton, OH. He continues to be very involved with music as head instructor of the drum and cymbal lines for the University of Maine’s Black Bear Marching Band.

Trevor has recently been accepted to the Doctor of Philosophy in Physics program at the University of Maine and will begin his work in that vein in the fall of 2007. Trevor is a candidate for the degree of Master of Science in Teaching from the University of Maine in May, 2007.