MODELING DISCOURSE MANAGEMENT COMPARED TO OTHER CLASSROOM MANAGEMENT STYLES IN UNIVERSITY PHYSICS

by

Dwain Michael Desbien

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree Doctor of Philosophy
ARIZONA STATE UNIVERSITY August 2002
has been approved
July 2002

ABSTRACT

A classroom management technique called modeling discourse management was developed to enhance the modeling theory of physics. Modeling discourse management is a student-centered management that focuses on the epistemology of science. Modeling discourse is social constructivist in nature and was designed to encourage students to present classroom material to each other. In modeling discourse management, the instructor's primary role is of questioner rather than provider of knowledge. Literature is presented that helps validate the components of modeling discourse. Modeling discourse management was compared to other classroom management styles using multiple measures. Both regular and honors university physics classes were investigated. This style of management was found to enhance student understanding of forces, problem-solving skills, and student views of science compared to traditional classroom management styles for both honors and regular students. Compared to other reformed physics classrooms, modeling discourse classes performed as well or better on student understanding of forces. Outside evaluators viewed modeling discourse classes to be reformed, and it was determined that modeling discourse could be effectively disseminated.

ACKNOWLEDGMENTS

I would like to thank my committee for their help and guidance during my research. I would especially like to thank David Hestenes for his willingness in letting me develop modeling discourse management. I would also like to thank my wife, Diane for her patience and support during the writing process. Without her editing and support I would have never finished.

TABLE OF CONTENTS

Page

CHAPTER

APPENDIX (see accompanying document)

LIST OF TABLES

LIST OF FIGURES

Chapter 1
Introduction

Studies into student conceptual understanding of Newtonian mechanics consistently reveal that traditional physics courses do little to change students' many misconceptions (Clement, 1977, 1981, 1982, Clement, Lockhead, & Monk, 1981, Champagne, Klopfer, & Anderson, 1980, Champagne, Klopfer, & Gunstone, 1981, Halloun, 1984, Halloun, Hestenes, 1985a, 1985b, Hake 1998). Traditional instruction – lecture, recitation, laboratory – typically involves little student interaction and follows a standard textbook (Halloun, 1984). These standard physics texts place a strong emphasis on problem solving and often rely on mathematical constructs to explain underlying physical concepts (Clement, 1981, Hudson & Liberman, 1982). Because of this emphasis on mathematics at the expense of physical concepts, little conceptual improvement occurs in the students.

Many researchers have implemented reforms to address the lack of student conceptual understanding with varying degrees of success (Halloun, 1984, Laws, 1991, Mazur, 1997). Some of these reforms depart radically from the traditional paradigm of instruction (Laws 1991, C. Deleone, personal communication, November, 2000) while others fit within the traditional framework (Mazur 1997). Examples of this latter type of reform include microcomputer-based labs (MBL) (Thornton & Sokoloff, 1990), peer interaction during lecture (Mazur, 1997), and interactive demonstrations during lecture (Sokoloff & Thornton, 1997). More radical reforms include the switch to a studio classroom. Studio classes make no distinction between lecture and lab and usually meet in a block of 2 hours at a time, three times a week. Students work in groups, and labs for this type of reform typically use MBL activities. All of the reforms mentioned above share the following goal: increase conceptual understanding.

The reform efforts mentioned are built on the premise that involving students more actively in their education will increase conceptual understanding (Wells, 1987, Wells, Hestenes & Swackhamer, 1995, Laws 1991, Mazur 1997, Thornton & Sokoloff, 1990, Thornton & Sokoloff, 1998). These reform efforts are based on the constructivist idea that students construct knowledge through interactions with each other and the instructor (Piaget, 1964, 1970 Vygotsky, 1962). The modeling method is one such constructivist reform (Wells et al. 1995). Overall, these reforms have shown significant improvement in student conceptual understanding; however, better student conceptual understanding is still desirable.

Historical Background for Current Research

Currently there are many groups involved in Physics Education Research (PER). These PER programs reflect many different viewpoints and approaches to physics education. A list of various PER groups, their primary focus, and primary investigator is given in **Appendix A**. Many other groups, while not performing PER, are active in the development of curriculum based on the results of PER. Two prominent examples are *Workshop Physics* by Priscilla Laws of Dickinson College (Laws, 1991) and *Spiral Physics* by Paul D'Alessandris of Monroe Community College (D'Alessandris, 2000). My dissertation was framed within the modeling theory of physics developed by David Hestenes at Arizona State University (Hestenes, 1987, 1992, 1996, 1998, 2000).

The Modeling Research Group at Arizona State University has been actively engaged in the reform of introductory physics education for the past 20 years. Eight years ago this group published the Force Concept Inventory (FCI) (Hestenes, Wells, & Swackhamer, 1992). The FCI is a multiple-choice inventory designed to measure a student's conceptual understanding of forces. Many PER groups and curriculum developers report that the FCI or its predecessors have changed the researchers' view of what the student learns in a mechanics course (Laws, 1991, Van Heuvelen, 1991a, 1991b, Mazur, 1997, Maloney, 1990). As a result, the FCI has become a standard measuring tool for the success of a physics course covering mechanics.

The ASU group is currently working to remodel University Physics. Much of this work was evaluated with the FCI, the Mechanics Baseline Test (MBT) (Hestenes & Wells, 1992), the Views About Science Survey (VASS) (Halloun & Hestenes, 1996, Halloun, 1997, 2001), and the Conceptual Survey in Electricity and Magnetism (CSEM) (Maloney, O'Kuma, Hieggelke, & Van Heuvelen, 2001). This remodeling effort is based on the idea that modeling skills underlie both understanding of physics and problem solving skills. In fact, the following question could be asked: What significance is a numerical solution if the conceptual understanding of its meaning is missing? Therefore, problem solving plays less of a role in the modeling curriculum even though it is still an important aspect of introductory physics. The ASU PER group is currently researching how to improve problem solving within the modeling physics framework (Hestenes & Politano, 1999, Brewe, personal communication, Fall 2000). However, I focused not on improving problem-solving directly, but on how classroom management issues of modeling effect student understanding.

Overview of Research and Justifications

This section gives a brief overview of the research I performed. Included is a brief description of the classroom management technique I developed (modeling discourse management). This model of instruction is contrasted with Socratic discourse commonly used in PER classrooms. The research questions are then presented with a brief description of how they were answered.

Introduction to Modeling Discourse Management

Modeling discourse management is the classroom management style I developed and investigated for this dissertation. Modeling discourse management attempts to create classroom discourse that is more student-centered than other PER based management styles. The instructor encourages students to bring new ideas and concepts to the class rather than using lecture or whole class questioning.

Figure 1 compares modeling discourse with Socratic discourse, which is common in PER courses (including previous modeling courses). Socratic discourse in physics education literature is often defined by an instructor questioning small collaborative groups or the whole class (Hake, 1992). The primary discussion occurs between the instructor and the students. Key features of modeling discourse such as seeding, questioning, and a learning community are demonstrated in **Figure 1**. Seeding/questioning is a common technique for motivating students to bring their ideas to the classroom community rather than to the instructor. The instructor is outside the community but interacts to provide activities, materials (tools), terminology, and guidance. The instructor extracts information from student discussions for use in formative evaluation of class progress and understanding. Terms used in **Figure 1** such as Learning Community and seeding are based on my own ideas and elaborated in Chapter 4. More details and a narrative example of modeling discourse management can be found in Chapter 4

Modeling discourse management was developed within the framework of the modeling theory of physics (Hestenes, 1992). The modeling theory of physics states that physics is based on a small set of models that represent the structure seen in the world. Modeling discourse management was designed to enhance the curriculum developed for the modeling theory of physics.

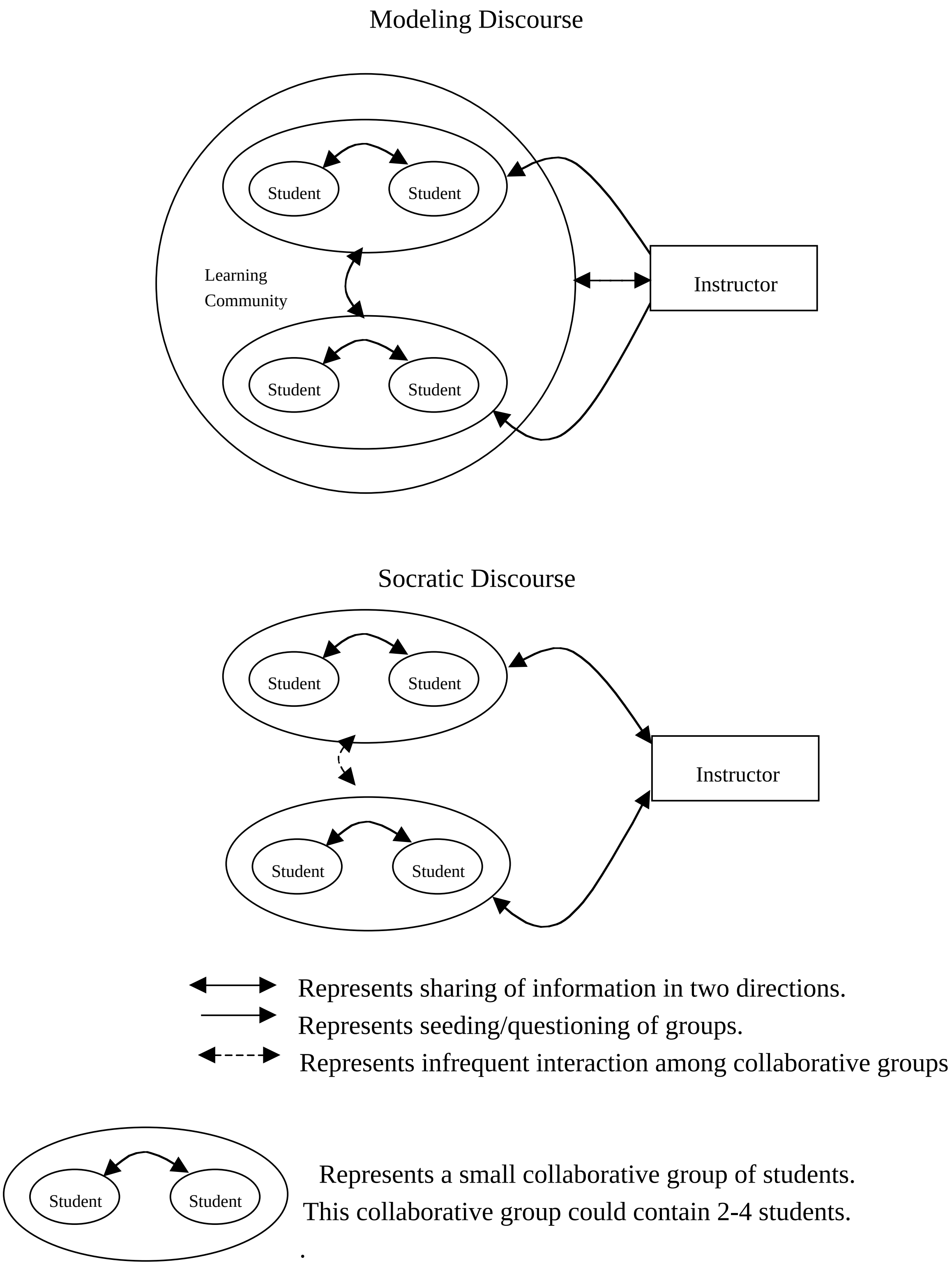


Figure 1. Modeling Discourse vs. Socratic Management Styles.

Research Description and Research Question

This dissertation has at its heart one basic research question: Does the inclusion of modeling discourse management in a modeling course improve student understanding when compared to other classroom management styles? Modeling discourse management was used in several university physics courses (calculus based) and compared to courses not using modeling discourse management.

Courses involved in this study can be broken into two categories: honors university physics courses and regular university physics courses. Within each category courses were studied that used modeling discourse management with and without using the modeling theory of physics. Also, within each category were courses using the traditional method of physics instruction. This last group of courses did not use the modeling theory of physics.

Within each category the classes were compared and contrasted on several different measures. Five smaller, more focused research questions were examined. Results from these five questions were combined to answer the basic research question listed above.

Question 1.

Does the inclusion of modeling discourse management into the classroom enhance student conceptual understanding of forces? I compared several first semester classes in university physics in an attempt to answer this question. The Force Concept Inventory (FCI) was used to measure student understanding of forces.

Question 2.

Does modeling discourse management improve student views about science? The Views About Science Survey (VASS) was administered to two modeling classes to help answer this question. The VASS was selected because students who understand and use the modeling theory of physics should score high. VASS data from these classes was compared to existing data to determine if modeling discourse management improved students' views of the process of science.

Question 3.

Do educators outside the modeling research project view courses taught using modeling discourse management as reformed? Third party observers evaluated the level of reform for most of the courses using the Reformed Teaching Observation Protocol (RTOP) (Piburn & Swada, 2000). This information helped determine the level of reform of a modeling discourse management class as seen through the eyes of trained observers.

Question 4.

Does a modeling discourse class develop problem-solving skills to the same extent as other physics courses? An instrument needed to be chosen that looked at basic problem-solving skills in mechanics. Therefore, the Mechanics Baseline Test (Hestenes & Wells, 1992) was selected because of its acceptance in the PER community as looking at problem-solving skills. The MBT was administered to modeling discourse courses only. The MBT data was evaluated to determine if a class with modeling discourse management has better problem-solving skills than those from baseline MBT data.

Question 5.

Is the modeling discourse management technique transferable to other instructors? I observed the quality of modeling discourse in several of the courses to determine if these courses were able to employ aspects of modeling discourse management. The Modeling Observation Protocol (MOP) was developed to help ensure consistent comparisons among classes. This data was evaluated to see how transferable the technique was from the developer to adopters.

Looking at the overall results of these five questions I will make a general conclusion about the effectiveness of modeling discourse management. That conclusion will be tempered based on the fact that the major question was not directly investigated. However, I will suggest future research that will hope to address that shortcoming and provide a clearer picture of the effectiveness of modeling discourse management. **Table 1** lists the five evaluation instruments used and the research question that they help to answer.

Table 1. Evaluation Instruments and Related Research Questions

Instrument	Acronym	Research Question (RQ)	Reason for Use
Force Concept Inventory	FCI	Used for RQ 1	Evaluation of student conceptual understanding of Forces.
Mechanics Baseline Test	MBT	RQ 4	Evaluation of student problem solving ability.
Reformed Teaching Observation Protocol	RTOP	RQ 3	Evaluation of modeling discourse as a reformed teaching technique.
Modeling Observation Protocol	MOP	RQ 5	Evaluation of the modeling discourse use in classes.
Views About Science Survey	VASS	RQ 2	Evaluation of student views towards science after a modeling discourse course.

Research Justification

The inclusion of modeling discourse management into the remodeled course was an attempt to improve student understanding beyond what had already been achieved by successful modeling courses. This effort was not a modification of an already successful curriculum but rather a non-traditional modification of the classroom management style. While the modeling method (Wells 1987, Wells et al. 1995) was designed for a constructivist classroom, the remodeled classroom created a learning community that was more social constructivist (Vygotsky, 1962, Roth & Roychoudhury, 1992) in nature. This research determined that a more unconventional constructivist approach improved student conceptual understanding and attitudes toward science. My work also sought to validate the position that improvement of classroom management can enhance an existing research-based curriculum (Halloun, 1984). These positive results led to an argument for future research into the inclusion of such reforms into other physics classrooms.

Dissertation Organization and Overview

Chapter 1 includes a brief description of the research performed. It outlines the research questions and presents a brief description of modeling discourse management. Chapter 2 consists of a literature review. The review presents research from within the PER community and from the general education research community. The general education articles reviewed present evidence of success using aspects of modeling discourse management in other contexts. Chapter 3 outlines the research design and how each of the five research questions will be answered. Chapter 4 is a detailed description of modeling discourse management and a narrative example. Chapter 5 contains the analysis and results for each of the research questions. Chapter 6 discusses the conclusions reached and future research plans

Chapter 2

Literature Review

This literature review outlines the framework in which modeling discourse management was developed. Studies or research on each major construct of modeling discourse management are discussed. The literature review is divided into two components: physics-related research and classroom management-related research. The physics sections will deal with modeling theory, student misconceptions in mechanics, the Force Concept Inventory, the VASS, and the MBT. The classroom management sections will discuss discourse management, classroom atmosphere, seeding, and student-student interactions. While there is some overlap of the two major themes, emphasis will be placed on how these topics relate specifically to modeling discourse management.

A recent study found that the inclusion of a discourse perspective within an existing successful reform curriculum enhanced student performance beyond what either reform could do individually (Duit, Roth, Komorek, & Wilbers 1998). Duit et al. define discourse perspective as the idea that the class should consist of dialogue among the students as the primary means of communication. This result is particularly relevant here, since I am attempting very similar research. Duit et al. (1998) concluded that the unification of discourse and curriculum shows promise and merits future research.

Practicing classroom teachers have designed new learning environments based on research in science and learning (Roth, 1998). While developing these new environments, teachers studied the effects of their teaching on student learning and the classroom environment. Formative evaluation of this sort plays an important role in modeling discourse management and was applied during the development phase. Both of these works lend credence to further research into formative evaluation of classroom environments and the combination of reformed curriculum and a discourse perspective. This work is built on both of these premises.

Modeling

Modeling is the process by which science attempts to explain the world (Bower & Morrow, 1990, Johnson, Johnson, & Johnson 1983, Hestenes, 1992). Reviewing the literature led to several articles that help define and clarify the modeling theory of physics and scientific models (Redish, 1994, Hestenes, 1987, 1992,1996, 2000, Halloun, 1998a). These articles can be grouped into the following two categories: those describing models/modeling theory and those describing how to teach modeling. My dissertation deals with models in physics and introduces a new classroom management technique into the modeling classroom; therefore, a discussion of each category is appropriate.

Modeling Theory of Physics

A modeling curriculum and a constructivist classroom management style (Socratic discourse) are principal components of the modeling method in physics education (Hestenes, 1987, 1992). The philosophy, curriculum, and classroom management style of the modeling method continue to evolve. The Modeling Research Group at ASU is currently involved in a large-scale effort to reform university physics. Education research in the group is concentrated along the following lines:

1. Building long-range coherence into the course content by
   - Design and evaluation of an energy thread
   - Design and evaluation of a structure of matter thread
2. Enhancing student learning through
   - Modeling discourse management
   - A modeling approach to problem solving

This dissertation focuses on enhancing student learning through the implementation of modeling discourse management. As part of the introduction of modeling discourse into the classroom an energy thread was also introduced into the curriculum. The evaluation of the energy thread is left for future research.

Model Definition

Science is the process of modeling the physical world (Hestenes, 1992); therefore, making models is fundamental to science. Since science is about making models, physics curriculum should be based on the same idea. Students should be brought to the understanding of the process of science (making models) and learn the explicit rules of modeling. The modeling theory of instruction is based on the premise that students need to understand the world in terms of the models science commonly uses. Thus, to understand modeling, an understanding of what constitutes a model is required.

A model is a representation of structure in a physical system. The structure gives it predictive and/or explanatory capabilities. There are four kinds of model structures: systemic, geometric, temporal, and interactive (Hestenes, 1992). A model may have any mixture of the four kinds of structure. An overview of the structure of a model can be seen in **Figure 2**. Each of the model structures has subparts that are discussed below.

I. Systemic Structure specifies
- Composition
- Environment
- Connections

II. Geometric Structure specifies
- Position with respect to a reference frame
- Configuration

III. Temporal Structure specifies changes in system properties via
- Descriptive Models
- Causal Models

IV. Interaction Structure specifies interaction laws.

Figure 2. Model Specification Overview

The systemic structure of a model specifies what constitutes the system of interest. This structure will vary depending on the situation being modeled. It also contains information on what objects outside the system of interest will influence the system. Lastly, the systemic structure specifies connections between objects. These are causal links that will cause changes to, or within, the system.

The geometric structure includes specifying the configuration of a system. It gives the geometric relations of the objects inside and outside the system. It also defines the position of the objects within the system with respect to an external frame of reference. The frame of reference allows connection to the geometry of the world outside the system.

The temporal structure of a model defines how the state variables of the system change. State variables are properties of the system or its parts. Examples of state variables include position and velocity. These state variables may describe how the system changes with respect to time. State variables can also be causally determined by differential equations from interaction laws.

The interaction structure of a model specifies the interaction laws mentioned in the previous paragraph. They could include force laws or potential energy equations. These laws describe quantitatively the causal links from the systemic structure. The equations from this structure are typically in terms of functions of the state variables from the temporal structure.

Figure 3 shows a simple representation of the relationship of the physical world to the kinds of scientific models described above (Hestenes, 1992).

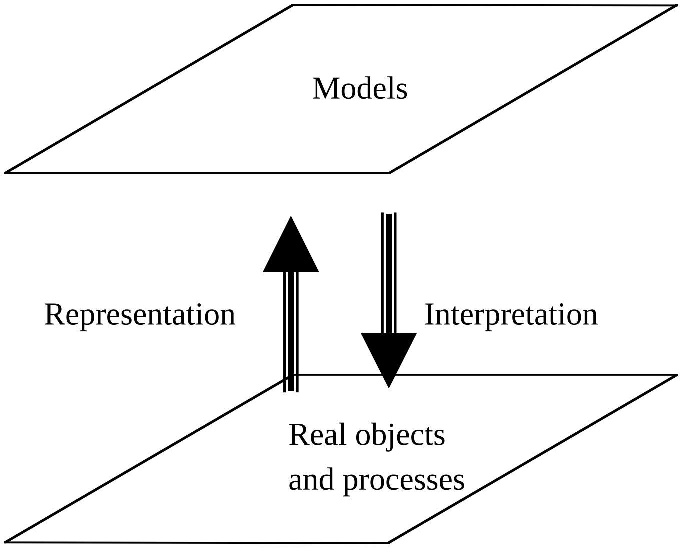


Figure 3. Model and the Real World Diagram

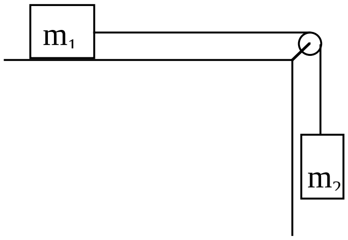
Particle Models

A first semester university physics course, like those being studied in this research, focuses on a small number of models. Most of the standard first semester curriculum is based on particle models of an object. **Figure 4** lists the common particle models used in university physics (Hestenes, 1996). Other commonly discussed object models are the rigid body and elastic body models. The FCI and MBT primarily focus on particle models and thus the emphasis here is on those models. **Figure 5** shows the structure of a model of a modified Atwood's Machine (Hestenes, 1996). The model is shown to emphasize the structures shown in **Figure 4**.

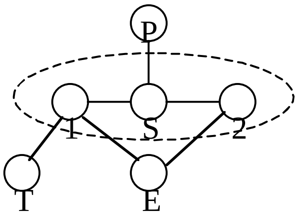
Kinematical Models	**Causal Models**
Constant Velocity	Free Particle
Constant Acceleration	Constant Force
Simple Harmonic Oscillator	Linear Binding Force
Uniform Circular Motion	Central Force
Collision	Impulsive Force

Figure 4. Basic Particle Models in Mechanics (see also Hestenes, 1996, Box 3.)

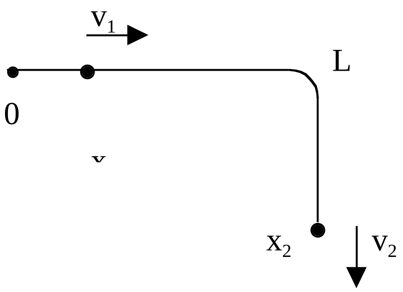
SITUATION MAP



SYSTEM SCHEMA



MOTION MAP



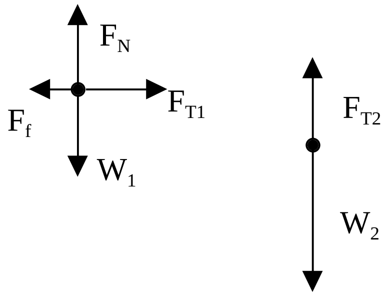
GEOMETRIC

$x_2 = x_1 + L$

$v_1 = v_2$

$a_1 = a_2$

INTERACTION MAP



INTERACTION LAWS

Internal: $F_{T1} = F_{T2}$

External: $F_f = \mu F_N$

$W_1 = m_1 g$

$W_2 = m_2 g$

TEMPORAL

For single particle subsystems:

$$F_{T1} - \mu F_N = m_1 a_1$$

$$m_2 g - F_{T2} = m_2 a_2$$

$$F_N = m_1 g$$

For two particle system

$$m_2 g - \mu m_1 g = (m_1 + m_2) a_1$$

Figure 5. Representation of Structure in a Model for the Modified Atwood’s Machine (Same as Hestenes, 1996, Box 5.)

The Modeling Classroom Without Modeling Discourse Management

While an understanding of a scientific model is important, more is required in order to manage a modeling classroom. In a modeling classroom, students are traditionally guided through a modeling cycle (Hestenes, 1992), which is a refinement of the learning cycle developed by Robert Karplus (Karplus, 1977). An outline of the Modeling Method can be found in **Figure 6** (Hestenes & Politano, 1999). The first half of **Figure 6** describes what should be taught and the second half describes how to conduct a modeling classroom.

Coherent Instructional Objectives

- To engage students in understanding the physical world by constructing and using scientific models to describe, to explain, to predict and to control physical phenomena.
- To provide students with basic conceptual tools for modeling physical objects and processes, especially mathematical, graphical and diagrammatic representations.
- To familiarize students with a small set of basic models as the content core of physics.
- To develop insight into the structure of scientific knowledge by examining how models fit into theories.
- To show how scientific knowledge is validated by engaging in evaluating scientific models through comparison with empirical data.
- To develop skill in all aspects of modeling as the procedural core of scientific knowledge.

Student –Centered Instructional Design

- Instruction is organized into modeling cycles which move students through all phases of model development, evaluation and application in concrete situations –thus promoting an integrated understanding of modeling processes and acquisition of coordinated modeling skills.
- The teacher sets the stage for student activities, typically with a demonstration and class discussion to establish common understanding of a question to be asked of nature. Then, in small groups, students collaborate in planning and conducting experiments to answer or clarify the question.
- Students are required to present and justify their conclusions in oral and/or written form, including a formulation of models for the phenomena in question and evaluation of the models by comparison with data.
- Technical terms and concepts are introduced by the teacher only as they are needed to sharpen models, facilitate modeling activities and improve the quality of discourse.
- The teacher is prepared with a definite agenda for student progress and guides student inquiry and discussion in that direction with “Socratic” questioning and remarks.
- The teacher is equipped with a taxonomy of typical student misconceptions to be addressed as students are induced to articulate, analyze and justify their personal beliefs.

Figure 6. The Modeling Method (same as Hestenes, 1996, Box 2)

Model-Centered Curriculum

One distinctive feature of the modeling method outlined in **Figure 6** is the focus on models as the content core of physics. These models are to be validated and evaluated using empirical data. Developing the understanding of the relationships of models to scientific theories is an important part of the modeling method. Lastly the students should learn that modeling is the procedural core of scientific knowledge.

Since its development, the modeling method has focused on a student-centered instructional design. Instructional time is organized around modeling cycles. Modeling cycles involve model development, evaluation, and deployment. Students work in small collaborative groups while performing each of these phases of the modeling cycle. The instructor is the guide for the class and keeps the discussion moving by Socratic questioning. The instructor has a specific agenda for each phase of the modeling cycle and by questioning ensures that the agenda is addressed.

Modeling Cycle and Classroom Management

The modeling cycle typically begins with a demonstration by the instructor. The class, with guidance from the instructor, decides what aspects of the demonstration are to be investigated. A discussion of how to investigate the areas of interest is agreed upon by the students and instructor. Collaborative groups then conduct experiments, deduce relationships from the data, and begin to formulate a model (or aspects of a model).

Next, students create white boards to share with the rest of the class and the instructor. These white boards contain the collaborative groups' experimental results along with models inferred or developed from the data. Groups then use the white boards in presentations to the rest of the class and instructor. Typically students present their white boards while facing the rest of the class. **Figure 7** shows a typical classroom presentation. Most of the questioning comes from the instructor, so students rarely directly question other students. Often when other groups have questions, they direct the questions to the instructor who either redirects the question to the group presenting or answers the question directly. All groups present their results with questioning and discussion following each presentation. The instructor introduces terms, ideas, and representational tools to improve the models and classroom discussion. These introductions take place during the small collaborative group work and during the large classroom discussion. At the end of the discussion a conclusion is reached and validated by the data collected by the various groups.

Figure 7. Typical modeling classroom whiteboard presentation (no picture here – too many bytes)

Following the discussion, students are given activities that require deployment of the model developed. This deployment can be in the form of in-class activities or homework problems. White boards are often based on this deployment to allow for the sharing of ideas and resolve questions about the activity. Students are often assigned different problems to white board so that a variety of deployments can be discussed among the class. **Appendix B** contains a sample lesson plan from the high school modeling curriculum.

This classroom management style has been effective in improving student understanding of physics (for example Hestenes et. al. 2000 and Hestenes & Politano, 1999). However, the modeling method is continuing to develop and adapt as new research is conducted. Modeling discourse management is one such development.

Student Misconceptions

The literature on student misconceptions in physics is long and varied. Many studies investigated student conceptual change in physics as a result of instruction (Clement, 1977, 1981, 1982, Clement et al. 1981,Champagne et al. 1980, Champagne et al. 1981, Trowbridge & McDermott, 1980, Halloun, 1984, Wells, 1987). This short list of prominent and important historical studies into student misconceptions continues to be widely cited and used in designing classroom materials (D'Alessandris, 2000, Laws, 1991a, 1991b, Thornton & Sokoloff, 1990, VanHeuvelen 1991a, 1991b). A full listing and description of the literature would be a book unto itself; however, this review will focus on work on misconceptions in mechanics.

Many PER groups focus their research on student misconceptions. Groups, such as the University of Washington, investigate a misconception and then create a curricular material to help address the problem (McDermott & Shaffer, 1998). This practice of investigating a limited number of Newtonian concepts is common in the literature (Clement, 1977, 1981, 1982, Clement et al. 1981, McKloskey, Caramazza & Green, 1980, Minstrell, 1982, Trowbridge & McDermott, 1980…). All of these investigations review a limited number of Newtonian concepts and report student understanding before and after instruction. All report virtually the same dismal result-- traditional instruction does not transform students into Newtonian thinkers. Many other studies have followed in the physics literature with virtually the same findings and conclusions (Mazur, 1997, Hake, 1998).

Recent studies tend to focus on misconceptions beyond those found in a traditional mechanics course. Recent efforts have included creation of a FCI-like instrument for electricity and magnetism (Maloney et al. 2000). Other efforts have looked at misconceptions in modern physics (Escalada, 1997, Scheer, Shaffer, & McDermott, 2000). These efforts examine student beliefs in quantum mechanics and ways to address these alternative conceptions. Solutions have included the use of computer visualization tools such as Visual Quantum Mechanics (Zollman, 2000) and tutorials addressed specifically at the misconception. These projects have seen some success, but like earlier research into misconceptions, have a narrow focus and often miss the coherence of the physics curriculum as a whole (Halloun, 1984).

In all studies the misconceptions found in students about physics are not easy to change. The studies mentioned have had some success at changing student ideas. Even though specifically addressed, the misconception often remains. Most of the concepts covered in the FCI are explained or demonstrated in a traditional lecture course; however, students' beliefs still do not change. The resiliency of student beliefs is not a new notion and has been written about for over a century (Pierce, 1877). Pierce makes the argument that unless a student's premises about a demonstration are questioned, students will have no need to change their ideas about the world around them. This resistance to change is commonly seen in the misconception literature. A student can recite Newton's third law, but still believes that a larger object exerts a larger force on a smaller object than the smaller object exerts on the larger object. Persistence of student misconceptions was recently found to be a common reason for "cooking" experimental data (Lawson, Lewis, & Birk, 2000). This study found that students would change data so that the experimental results would match their misconceptions. Even when confronted with data that did not match their beliefs, students would "correct" the data to match their belief. The resiliency of misconceptions was not limited to undergraduates as recent evidence was reported in science education literature reported (Lawson, Drake, Johnson, Kwon, & Scarpone, 2000).

Force Concept Inventory

The Force Concept Inventory (FCI) was first published in 1992 (Hestenes, Wells, & Swackhamer, 1992). The FCI was designed to examine student understanding of the Newtonian conception of force. It consisted of 29 multiple-choice questions aimed at addressing all aspects of the Newtonian force concept. Data was given comparing the performance of students in modeling courses to traditional lecture-recitation-lab classrooms. The FCI was also used to measure the effectiveness of modeling courses in high school with an exemplary teacher (Hestenes et al. 1992). Data was given showing the performance was better when using modeling in the introductory physics classroom compared to a traditional course. A taxonomy of the Newtonian force concept and corresponding questions on the FCI are given in the article.

Since its publication, the FCI has gone through one minor revision. This occurred in 1995. At that time, one question was omitted, another two added, and several modified. These modifications occurred because of feedback from the physics community. The FCI is now a 30 question inventory and is still aimed at addressing student understanding of the Newtonian force concept.

Since its publication, many articles have been published using the FCI as a tool to measure the effectiveness of teaching mechanics (too many to list but a few recent examples include: Yamamoto, 1998, Adams & Chiappetta, 1998, Cummings, Marx, &

Thornton, 1999, Francis, Adams, Noonan, 1998, Saperstein, 1995). Many American Association of Physics Teacher talks use the FCI as a measure of the effectiveness of instruction (AAPT Announcer, Fall 2001). Since the FCI is so widely used, many have evaluated what it is measuring (Huffman & Heller 1995, Heller & Huffman, 1995). Hestenes and others have responded in defense of the FCI (Hestenes & Halloun, 1995), and the FCI continues to be the most widely accepted tool for measuring the force concept (Hestenes, 1998).

Meta-analyses have been done using the FCI to compare many classes that would have never been compared otherwise (Hake, 1998). This study found that interactive-engagement courses outperformed traditional courses on the FCI when evaluating the normalized gain. The Hake study involved over 6000 students and Hake continues to update the data and results on his website. The wide acceptance of the FCI makes it a good tool to compare classes covering mechanics and to make simple cursory comparisons to classes outside the study of interest.

Views About Science Survey

Research has shown that student views about science can effect their ability to learn science (Halloun & Hestenes, 1996, Halloun, 2001, Aikenhead, 1987). The VASS consists of two types of questions. The first type of question is dichotomous in nature. Responses are one extreme or the other. The second type uses a contrasting alternatives scale. This second type of question makes up the majority of VASS questions and carries the greatest weight in determining the VASS score.

The VASS was developed to help instructors evaluate instruction aimed at changing student views about science (Halloun & Hestenes 1996, Halloun, 1997, 2001). The VASS categorizes students into four profiles: folk, low transitional, high transitional and expert. For example, a student with an expert profile answers like a physicist. Research using the VASS has found that little or no change in student profiles occurs as a result of traditional physics instruction but a correlation exists between achievement and student profiles. The same findings are true of modeling instruction without modeling discourse management. VASS has been modified to make scoring easier and to make distinctions between the four student profiles more distinct (Halloun, 2001). This latest version is the one used for this dissertation. However, the percentage of students in each category remains fairly constant from previous versions to the new version. The newest version has the advantage that a numerical score is assigned to each student and thus average scores for a group can be found and statistical comparisons can be made among groups. Reliability and validity of this VASS version are addressed by non-traditional techniques (Halloun, 2001).

Mechanics Baseline Test

The MBT was designed to be the "next step" in the understanding of mechanics from the FCI (Hestenes & Wells, 1992). The MBT requires mechanics knowledge that can only be obtained with instruction. Thus, the MBT is typically only given as a post-test, as is the case in this dissertation. Because it is not appropriate for a pre-test, fewer studies use the MBT as an assessment tool (Hestenes & Wells, 1992). Unlike the FCI, the MBT requires some simple computations to obtain the correct answers. However, it is still designed to test conceptual understanding. The MBT scope is that of the skills required in mechanics problem solving. Thus the FCI and MBT together give a good picture of student understanding of basic mechanics.

Discourse Management

As previously stated, Duit suggested that the deliberate inclusion of a discourse perspective into a reformed curriculum can enhance the course more than either reform individually (Duit et al. 1998). A review of studies of science classroom discourse suggests that curricular changes are insufficient and that students should understand that science is the "ongoing flow of discourse" (Scott, 1998). Management of student discourse within a successful curriculum is central to this work. How to manage discourse in the classroom so that it is productive and effective is a nontrivial question. The imposition of a discourse management style onto a quality curriculum allows for formative evaluation of both the curricula and the management style.

The next few sections detail research into aspects of discourse and discourse management that are important in modeling discourse management. Studies are included that used techniques similar to those in modeling discourse management. These studies are not the only uses of discourse and discourse management, but they do provide a theoretical framework for modeling discourse management to be built upon.

Structuring the Classroom

How does one structure a course that is to have a strong discourse perspective? Several authors have made suggestions based on classroom research (Roth & Roychoudhury, 1992, Day, 1995, Johnston, Woodside, &Day 1992). For discourse to be effective it must be allowed to reach a conclusion, which may mean that the discourse needs to be quite long and uninterrupted. Extended discourse is required if students are going to be able to work out differences and reach consensus on an idea or concept. Extended discussions will often require more time than is allotted in a standard class time frame. Block scheduling with back-to-back class time slots allows more time for discussions and makes the management of the classroom discourse easier. With an extended period available, the instructor can allow the discussion to follow its natural path to a conclusion without having to interrupt and hasten the discussion. This natural progression allows students to question and address their own ideas without feeling rushed.

For effective discussion, the physical location and orientation of the students is critical (Malcolm, 1991). How the students interact is directly related to their ability to be seen and heard. A student's position within the classroom will also affect his or her likelihood of contributing to the discussion. Ideally, all students should face all of their classmates, which can be accomplished by positioning the students in a circle. This arrangement also helps keep students on task, as it is much harder for students to drift from the conversation when all their peers can see them.

Managing Discussions

Arranging the layout of the discussion environment is straightforward. However, getting the students to actually discuss in a productive manner requires careful management of the classroom. Classroom discussion should be designed with these three questions in mind (Tobin, 1990):

- How do students negotiate meaning?
- What is involved in shared meaning?
- What is involved in consensus building?

A discussion or activity should be designed with an understanding of how the class will address these questions. Without knowing how the students will negotiate meaning or build consensus makes it difficult to manage a discussion that helps students confront and resolve their differences. Knowing when a class is reaching a shared meaning is also very important. Thus, a careful study of the individual class is required to determine how each discussion should be framed to meet that class's unique setting. Suggestions for activities that help instructors and researchers answer these questions about a class have been suggested by Roth and Roychoudhury (1993). During a discussion, students will develop temporary alliances that are dynamic. The instructor must be able to pick out these alliances and what role they play in the progress of a discussion.

Instructors must also be aware that their responses to these questions will directly affect their students' responses as well (Johnston et al. 1992). The epistemology that the instructor uses to orchestrate the class will likely be the one that the students will use to relate to the subject matter. Thus an instructor who feels confident in his/her answers to the questions will likely have students who are confident in a classroom that focuses on such questions.

However, the important point is that to manage classroom discourse effectively the instructor must understand how students view the three aforementioned questions. Thus, activities that are not adding material but helping the students understand how they address consensus building and help the instructor answer these questions are needed early in the course if discourse is to effectively occur.

Conflict in the classroom is inevitable. How conflict is resolved is particularly important in a classroom that emphasizes classroom discourse. Studies have found that when conflict is mediated very slightly in a classroom discourse course, the result is a greater understanding for the entire community (Mortimer & Machado, 2000, Leonard, 1999). In classrooms where discourse is fundamental and accepted by the students, perceptions of conflict are often different than those of society. Conflict can be seen as positive and healthy instead of adversarial. When students view conflict in this manner, they feel it is acceptable to disagree and they seek to find common understanding.

Emphasizing Instructor Confidence

Besides the physical environment and suitable activities, instructor confidence in the subject and in managing the discourse itself is required for a productive class to occur (Tomanek, 1994). Without content knowledge confidence, the instructor may not be able to effectively resolve discrepancies within the discussion. Good discussion exceeds the bounds the instructor originally intended. The instructor needs the confidence to deal with unanticipated issues. If the instructor is uncertain about how to resolve an issue, the instructor should admit it openly, and the students should be challenged to improve everyone’s understanding.

Besides content confidence, the instructor must convey confidence in the efficacy of discussion. To keep the discourse running smoothly instructors also need to know how to handle interference from student discourse styles learned in other contexts (Roth, 1997a). Other instructors conduct instruction according to their beliefs, so students bring preconceived notions about discourse into the classroom. The instructor should create the climate and clear expectations for discourse from the first day of class.

Beyond the Classroom

Outside the class, students will continue to learn and grow. Continued learning is facilitated by the ability to recognize discourse practices and effectively participate in them (Roth, 1997c). To the extent that it becomes internalized in student behaviors effective discourse management goes beyond the classroom and helps prepare students to be effective members of society.

Classroom Atmosphere/Learning Community

A science classroom should help students understand that science is tentative and continually evolving (Elby & Hammer, 2000). Science should be understood as being subject to human perspectives and as such it is socially constructed by scientists. However, the creation of a classroom where such ideas are understood would require a radical change from the traditional course. Lecturing as an authority on what science is and its tentativeness is inconsistent with the notion that science is socially constructed. Students must be immersed in a classroom atmosphere that facilitates the development of scientific thinking. Students need to recognize changes in their beliefs and convince others of the merits of their new ideas. These new ideas should be shared and agreed upon as a foundation for future development.

Formative Evaluation

Creation of a learning community as described requires the instructor to continually evaluate the classroom and modify the atmosphere as needed (Roth, 1998, Roth and McGinn 1998). Formative evaluation of the class must not only consider student understanding of the material but also the classroom environment. Forming the class into an effective social learning community must begin immediately in the course, and evolution of this community must continue throughout the course. Without continued evolution the atmosphere will become stagnate and discourse management will become more difficult. The instructor cannot neglect either the content or the learning environment. Continued development of each by formative evaluation is required for a class to grow and learn. Formative evaluation involves reflection by the instructor after each class. This reflection is facilitated by taking notes during classroom discussions, evaluating understanding by listening to the quality of argumentation, and evaluating class responses to seeded questions.

Student Views in the Learning Community

Research has shown that unless the students see the need for a learning community, no amount of effort by the instructor will continue classroom growth (Roth et al. 1999). Students need to understand the importance of building a consensus. They must understand the need for an open and supportive environment to make progress. Students therefore must see the explicit need for an agreement on the conduct of social interactions. Within this setting they can learn to see science as tentative and evolving.

Students' views of the classroom learning community determine its success (Lucas & Roth, 1996). Unless the students recognize the need for a community, its benefits are lost. Lucas and Roth found that as students' views of the nature and need for the community evolve, so does their understanding of the course content. A well-formed classroom atmosphere helps develop student understanding of science. Inherent in this type of environment is the notion that knowledge is constructed and shared through dialogue (Stage, Muller, Kinzie, & Simmons, 1998). Students must recognize this idea for this type of classroom environment to be successful. In addition, students must understand the process by which they are to learn science to optimize its success.

For an effective learning community, students must come to share the view that truth requires coherence (Starver, 1998). Students must negotiate a common understanding of what is observed and on how to interpret what they observe. The community needs to resolve differences in opinions and ideas to make a coherent picture of the content of the course. Only through sharing and seeing a need for that sharing will development of a coherent consensus occur. As the course evolves, student views of a particular situation will change along with the community view. A more in depth look at an activity may lead to a greater coherence and a deeper understanding for the students. However, without the explicit understanding by the students of the need for consensus, incompatible views of the activity will remain unresolved.

The Role of Participation

For a learning community to be successful everyone must feel comfortable about participating (Reynolds & Nunn, 2000, Lopez-Freman, 1995). That does not mean that everyone must participate in every discussion. A classroom atmosphere that is conducive to participation requires an environment that is supportive and empathetic rather than competitive. Reynolds and Lopez-Freeman found this to be especially true for women and minorities. Another study found that women may still be uncomfortable, especially pre-medicine majors (Laws, Rosborough, & Poodry, 1999). A supportive environment included peer encouragement to share ideas and consider alternative points of view. Peer questioning needs to be non-threatening and aimed at developing group consensus and understanding. The instructor should try to correct problems in a manner that does not embarrass individuals or impair the overall tenor of classroom discussions.

Research has shown that such collaborative atmospheres provide insightful experiences for the students and improve student understanding of course content (Roychoudhury & Roth, 1992, Younger, 1999). Younger used an athletic team analogy to help describe the atmosphere desired in the classroom. On a team, everyone works toward one goal. No one person is more important than another. Occasionally, certain individuals will provide more of a boost than others; however, subsequent activities should be designed to highlight other students' strengths. If one person on the team feels he or she is more important than the rest of the members, the whole team will suffer. This is true in a collaborative classroom also. As the class works together more effectively, the understanding of the whole class will improve. As the classroom becomes more supportive and less competitive, better understanding of differences develop and the class will become a more congenial place.

Student Behavior

As the students understand the need for cooperation and sharing, their behavior improves. DeDobbelaere, Leaf, & Ziegler (1998) found that creation of a learning community by including activities aimed at improving student behavior was successful within the classroom. Unfortunately these improvements were short-term and there was no transferability to other situations. This research found that in the classroom where the learning community exists, behavior is better. The sooner the creation of a cooperative classroom occurs, the easier conflict resolution will be (DeVries & Zan, 1995). This study found that the earlier the development of shared meaning and a social atmosphere took place the easier it was to diffuse conflicts when they arose. Lastly, it has been demonstrated that behavior of students is important to improving their critical thinking skills (Browne, 1986-87). If it is possible to modify student behavior in the classroom it is possible then to start to improve students' critical thinking skills.

Communication Tools

For a community to function effectively, it must also have the tools to communicate. Shared tools for communication need to be developed by the class (Roth, 1995, Roth, & Bowen 1995, Roth & McGinn, 1998). Tools, in this sense, are anything that allow for the representation of ideas during the discussion. The more people involved in this development, the more robust the tool use will be. The more tools used in discussions, the more knowledge is transferred to the individual students. Some students may understand an idea by one representation while others may require a second or even a third. These representational tools must be socially agreed upon and available for all. If an individual brings a new representational tool to the group, that individual must be willing to share and explain it to the satisfaction of the whole community. Without communication tools the community would not be able to progress. One such common tool used in a modeling classroom is the whiteboard (Wells, 1987).

Creation of a learning community is critical to the success of a collaborative classroom. Unless the students see the need for and the advantage of cooperation, the class will not progress effectively. Explicit understanding by the students of the learning community being developed is required, along with positive peer encouragement. Such a community improves student behavior. Also, conflict resolution is easier because of the spirit of cooperation. Students must understand the process of the development of science and be able to communicate it to their classmates. Tools must be socially available for aiding in discussion of ideas and reaching consensus.

Seeding

Seeding involves two interrelated activities: planting and questioning. The first activity is to plant an idea or concept with a small collaborative group of students so the students can bring that idea to the larger learning community (Green, 2000). The second activity is for the instructor to question a small collaborative group to guide the group's thinking. The collaborative group can then bring its answer to the instructor-supplied question to the learning community (Leonard, 1999). Both of these techniques to introduce concepts or questions enhance classroom discussions and the level of student involvement (Maloch, 1999, Townsend, 1998). Students guided in this manner often find the course material more interesting, and they feel greater ownership of the ideas (Jetton & Alexander, 1997).

Student-Student Interactions

Social constructivism places the process of learning squarely in a dynamic social context (Confrey, 1995, Ernest, 1995, Prawat & Folden, 1994, Vygotsky, 1962). A consensus on scientific knowledge is socially constructed through journals and conferences. The community shares ideas and results, modifies the results based on feedback and experimentation, and continues the process of modeling the world. Students should be initiated into this process (Driver, Asoko, Leach, Mortimer & Scott, 1994). To understand the process of science, students must take an active role in developing their understanding (Piaget, 1964). Therefore, student-student interactions are vital to initiating students into the scientific process. The need for student-student interactions must be explicit and the rules for such interactions must be shared and agreed upon.

Group Structures and Language

Student-student interactions are facilitated by an environment with structures for discourse and tools to aid the students in communication. Within a suitable environment, students develop group structures that afford quality student-student interactions in time (Roth & Bowen 1995). These structures take time to develop, and students, not the instructor, must develop them. The instructor guides the development by seeding small groups to enrich the larger classroom community. In this manner the ideas develop from the students' peers and not the authority figure. During the development of group structures there must also be group development of tools to aid in communication among the students (Roth & Roychoudhury, 1992). Tools must be agreed upon and the learning community must understand the context in which the tools are to be used. No amount of time will allow for group structures to develop that enable students to have effective peer interaction in developing content knowledge, unless proper tools are developed.

The structures developed by student-student interactions are dependent on students' use of language and their relationships to each other (Roth, 1997c). Thus, it is important for a social community to be created within the classroom in which scientific ideas can emerge. The language of science is to be provided by the instructor while the concepts and discussion are to be developed by the student. When the student-student discussion has reached a consensus, it is the instructor's role to introduce the appropriate language for the concept or idea. In this manner the construction is framed by student-student interactions, but the language is that of the scientific community. Since student relationships also influence student-student interactions, it is vital to create a community in which cooperation, not competition, is rewarded.

Constructive Student-Student Interactions

Studies have found that students must place emphasis on empathy and understanding for peer interaction to occur and be constructive (Reed, Mcleod, & McAllister, 1999. Henry, Reed, & McAllister, 1995). Interviews and surveys with students consistently found that for students to participate in peer interaction, the addressee must be empathic and understanding of the ideas being presented. Disagreements about ideas can be resolved, but the feeling of ideas not being considered by peers will destroy the learning community and effectively end discussion. Students must feel that their peers are interested in their contributions and that differences can be resolved in a constructive manner. These studies investigated how students viewed student-student interactions. Therefore, to have successful student-student interactions, the classroom atmosphere must address the issues that the students consider important for participation in discussions.

Developing student-student interactions is also vital since world-views are shaped by the social circle in which they are created (Clark, 1998). The atmosphere in which the knowledge is created will greatly affect the knowledge. Therefore, to make sure all opinions and ideas about the world are understood and melded into the social understanding of the classroom, students must feel comfortable sharing with their peers. Social constructivism is built on this idea that each person makes sense of the world by his or her experiences and the influences of the social company kept. Student-student interactions allow for the most diversity of ideas to be shared and questioned because ideas from fellow students are more likely to be questioned than ideas from the instructor. Therefore, students must question their own beliefs based on those suggested by their peers. When ideas come from an authority they are readily accepted as correct and real changes in student beliefs do not often occur. Student-student interactions help to combine many ideas into a body of knowledge that each individual has helped to create. This new knowledge is only created when students confront their own beliefs and modify or abandon them.

The Role of the Instructor

With emphasis placed on student-student interactions, what role does the instructor play in a social constructivist classroom? The teacher is the guide but not the leader of the discussions (Roth, 1996). The instructor keeps the discussion focused and on task but should avoid being the leader. When the discussion gets off task the instructor should intervene and redirect the discussion with an appropriate question. However, the instructor should be flexible in his or her schedule to allow student discussions to continue when the path is appropriate. The students may be ready to discuss a concept before the teacher anticipated and these discussions provide an opportunity to move the class forward more quickly. Discussion diversions into the students' daily lives should be encouraged when appropriate. The instructor ensures that the students explore the consequences of their ideas by interjecting questions to the group or introducing an activity that builds on the discussion.

The instructor is also charged with creating an atmosphere in which student-student interactions are fostered and nourished (Ben-Ari & Kedem-Friedrich, 2000).

The teacher must develop situations that require student-student interactions. This can be accomplished by developing activities that are rich enough that a wide variety of ideas are required to accomplish a task or solve a problem. Also, during small group work the instructor can seed ideas for the small groups to bring to the whole class during discussion. Through seeding, the students, not the instructor, present the ideas to the whole class. Instructor evaluation of the classroom atmosphere is vital to continued student-student interactions. Activities must be included that continue to develop the community and foster student-student interactions. As a result, these activities may lack course content but ultimately will aid the class by improving student discussions and enriching the class understanding of the material.

Interaction and Cognitive Conflict

The constructivist philosophy places emphasis on the creation of cognitive conflict in the students to overcome existing beliefs (Piaget, 1964). Cognitive conflict has been researched and used in developing physics curriculum for several years (Dykstra, Boyle, & Monarch, 1992, Mestre & Touger, 1989, Redish, 1994, Redish, Saul, & Steinberg, 1998, Thornton, 1997). However, misconceptions are tenacious and more than cognitive conflict is required to change student beliefs (Guzzetti, Williams, & Skeels, 1997). Guzzetti et al. found that the combination of cognitive conflict with student-student discussion is more effective in changing student beliefs than conflict alone. Students will often find parts of activities that support their misconceptions and will try to reinforce their beliefs even though the activity conflicts their notions of the world. Therefore, student-student interaction is necessary, in addition to conceptual conflict, to help change student conceptions. Peer interactions allow students with many different ideas to discuss and work out differences. Many different student ideas that must be reconciled reinforce the activity and aid in changing student ideas. While conceptual conflict is required for change it is not sufficient and must be supplemented with discussion (Guzzetti et al., 1997).

Conclusion

The studies discussed have been described to highlight the practices and philosophies of modeling and modeling discourse management. While not concerned with modeling discourse explicitly, each study from the discourse management section provides evidence for ideas or practices that have contributed to its modeling discourse management's development. Chapter Four describes modeling discourse management in detail and synthesizes the ideas reviewed here into modeling discourse management.

Chapter 3

Research Design and Analysis Methods

This chapter describes the design and analysis of an investigation into the effectiveness of modeling discourse management. It includes a description of the classes involved in this study and how the data was collected, a description of the evaluation instruments used, and an outline of the statistical data analysis, including a description of the comparisons to be made and how they will be evaluated. The implications of each comparison are briefly discussed at the end of this chapter.

Course Descriptions

Most of the data I used was collected specifically for this research. However, some data was collected previously and not specifically for this research. This data was collected in the same manner as that specifically for this dissertation and therefore appropriate to use. The Modeling PER group has been collecting data from the FCI, MBT, and VASS for several years since their publication. When appropriate random samples from the pool were used for comparisons. The samples were taken from university physics courses that had supplied their data to the Modeling PER group. To aid the reader **Table 2** contains names and associated acronyms used throughout this chapter.

The classes and data collected for this dissertation are listed in **Table 3**. Two groups of calculus-based physics classes have been studied – honors university physics and regular university physics. Unless stated otherwise, all classes were regular university physics courses. For all classes in **Table 3**, FCI data includes both pre- and post-data.

Table 2. Various Names and Associated Acronyms

Full Name	Acronym
North Carolina State University	NCSU
Chandler-Gilbert Community College	CGCC
Arizona State University	ASU
Force Concept Inventory	FCI
Mechanics Baseline Test	MBT
Views About Science Survey	VASS
Reformed Teaching Observation Protocol	RTOP
Modeling Observation Protocol	MOP
Foundation Coalition	FC

Table 3. Data Collected for this dissertation

Institution	Year	FCI	MBT	VASS	RTOP	MOP
ASU FC	2000-2001	X	X	X	X	X
ASU FC	1999-2000	X	X			X
ASU Lecture	2000-2001	X			X	X
NCSU PER	2000-2001	X				
CGCC	2000-2001	X	X	X	X	
CGCC	1998-1999	X	X		X	
ASU Honors	2000-2001	X			X	X
ASU Honors	1999-2000	X	X		X	X
ASU Honors	1996-1997	X	X			

The courses I chose to study represented several instructional styles and classroom management styles. General information on the courses is given in **Table 4**. *Level of Class* in **Table 4** refers to whether the course was an honors or regular university physics course. The courses labeled *FC* are regular university physics courses and are part of an integrated first year program for engineering students. *Number of students* contains only those students who completed the course and who took the FCI pre and post. The column labeled *Modeling Curriculum* indicates if the curriculum was centered on the models developed by the modeling theory of physics. The last column describes the management style employed in the class. Determination of the management style was made by classroom observations with the exception of the NCSU data which was self reported.

I considered the instructional style of a class to be the primary treatment for comparisons. The NCSU data used a PER based curriculum and used interactive/Socratic discourse for classroom management and therefore was used as a control group in making comparisons to modeling courses. The honors courses provide an ideal series of comparisons since one used traditional methods, one used modeling without modeling discourse, and one used modeling with modeling discourse. Following **Table 4** is a description of each course.

Table 4. General University Physics Course Information

Institution	Location	Level of Class	Year	Number of Students[1]	Modeling Curriculum	Instructional Style
NCSU	Raleigh, NC	Regular	2000-2001	55	No	Interactive/Socratic
CGCC	Chandler, AZ	Regular	1998-1999	15	Yes	Modeling Discourse
CGCC	Chandler, AZ	Regular	2000-2001	20	Yes	Modeling Discourse
ASU	Tempe, AZ	Regular	2000-2001	116	No	Lecture/Lab
ASU	Tempe, AZ	FC[2]	1999-2000	44	Yes	Modeling Discourse Inconsistently[3]
ASU	Tempe, AZ	FC[2]	2000-2001	60	Yes	Modeling Discourse
ASU	Tempe, AZ	Honors	1996-1997	13	Yes	Interactive/Socratic
ASU	Tempe, AZ	Honors	1999-2000	19	Yes	Modeling Discourse
ASU	Tempe, AZ	Honors	2000-2001	12	No	Lecture/Lab

[1] This number is only for FCI and MBT data where appropriate. The numbers for VASS data can be found in chapter 5.

[2] FC refers to a program for freshman engineering students at ASU. There are no requirements for admission other than desire and the usual course prerequisites. This is a regular university physics course.

[3] This determination was based on MOP scores. See page 118 in Chapter 5.

ASU Foundation Coalition (FC) Courses

Both of the ASU FC courses were part of a first year program for engineering students. The program was an integrated first-year experience in which student's took all of their courses together. A special room was designed to facilitate group work, and all classes in the program used interactive classroom techniques. There were no special requirements for admittance into the program other than an expressed desire. On standard measures of success (high school GPA, SAT scores), the FC students did not differ from the other first-year engineering students at ASU (D. Evans, personal communication, Fall 2000). All students in the program were entering freshmen. The 2000-2001 course was team taught by three instructors. This course consistently used modeling discourse management. One of the three instructors was the developer of modeling discourse management and the author of this dissertation. The second instructor was a graduate teaching assistant working with the modeling research group. The third instructor was a visiting assistant professor also working with the modeling research group. All three instructors had previous experience with modeling discourse management, but had differing levels of comfort and experience. Duties were shared equally among the instructors so that no hierarchy existed. Each instructor worked exclusively with one set of students so that intra-class comparisons of student performance are possible. This also enabled study of the transferability of modeling discourse management.

Three instructors also taught the 1999-2000 FC course. Two of the three instructors were the same as the 2000-2001 course, the author of this dissertation and the graduate student. An associate professor was the third instructor. Modeling discourse management was used inconsistently in that course based on MOP scores. The reader is referred to page 118 in Chapter 5 for more details. Neither the graduate student nor professor had classroom experience with modeling discourse management at that time. However, both had attended a two-week workshop on modeling discourse management. Like the 2000-2001 course, class duties were shared and no hierarchy among the instructors existed. Unlike the 2000-2001 class, each instructor worked with all the students so within class comparisons are not possible.

ASU Honors Courses

A professor new to ASU taught the 2000-2001 ASU honors course. The professor had several years of teaching experience at another institution. This course was taught in a traditional manner even though the class-time was scheduled as a studio class. The students were members of the ASU honors college and thus had a strong academic background. The sample size for post-test FCI data for this class is very small because one half of the students withdrew during the first semester.

The other two honors courses were taught by the same visiting assistant professor who worked with the 2000-2001 FC course. In the 1999-2000 course the professor used modeling discourse management for the first time. In the 1996-1997 course the professor used modeling without modeling discourse management. These two courses were both taught using a studio format. Like the 2000-2001 course, the students had strong academic backgrounds. However, unlike the 2000-2001 honors course these two courses did not have a high attrition rate. The same visiting professor also taught the honors course for the years 1997-1999. During that time the professor learned modeling discourse management. The year 1999-2000 was the first time the professor had used modeling discourse management without the assistance of the developer. In all courses the students were traditional university students (students who went from high school directly to college).

ASU Lecture Course

A well-respected professor using traditional methods taught the 2000-2001 ASU lecture course. Thus, student interactive classroom techniques were not used. The class consisted of three one-hour lectures per week and one hour of recitation taught by various graduate students. The lab portion of this university physics course was completely divorced from the lecture portion. Coordination between the lecture and lab was not attempted and the labs had no bearing on the grade in the lecture portion of the class.

NCSU Physics Education Course

The NCSU PER course was a class of approximately 60 students and was interactive in nature. This course was NCSU's attempt to reform their university physics class (Beichner & Saul, 2000). The students in this course were similar to those in the ASU FC course and ASU lecture course. NSCU's reform efforts did not focus on modeling and thus make a good comparison group for the effectiveness of modeling and modeling discourse management. Modeling discourse management was not used in this course. More information on NCSU reform efforts is available in the reference for SCALE-UP (Beichner, 2001).

CGCC Courses

I, the developer of modeling discourse management, taught both of the CGCC university physics courses. The CGCC courses met twice a week for two and one half hours a class period. Students were mixed between traditional and non-traditional and a wide range of ages existed. All courses were taught using modeling discourse management.

Other Courses

In addition to the classes mentioned one more source of data was used in this dissertation. The modeling research group at ASU has been collecting data from the FCI, MBT and VASS for several years as part of its continuing research efforts. This is a useful data pool for comparisons. Random samples were selected from university physics courses within the pool data. This data was then treated as any other data set.

Evaluation Instruments

Evaluation instruments were chosen to investigate specific aspects of the courses involved in this study and to answer specific research questions. This section details the instruments and explains what each instrument measured and why it was chosen. **Table 5** gives a summary of the reliabilities for the instruments

FCI

Many studies have been conducted using the FCI to measure the effectiveness of a physics course at improving student conceptual understanding (Hake, 1998, Adams & Chiappetta, 1998, Yamamoto, 1998, Mazur, 1997, Cummings et al. 1999, Francis et al. 1998, Saperstein, 1995). Comparisons of courses taught using the modeling method to those taught using traditional methods and other reform efforts show the modeling method compares favorably. This fact can be seen in **Figure 8**, which shows normalized gains for various types of courses (Hake, 1998, S. Osborne-Popp personal communication, Fall 2000). The normalized gain is the ratio of the actual average gain to the maximum possible average gain (Hake, 1998). While the modeling courses performed as well as other reform efforts, better student performance should be possible.

Table 5. Reliabilities for instruments used

Instrument	Acronym	Reliability (r squared)
Force Concept Inventory	FCI	0.90
Mechanics Baseline Test	MBT	0.85
Reformed Teaching Observation Protocol	RTOP	0.95
Modeling Observation Protocol	MOP	NA
Views About Science Survey	VASS	0.68

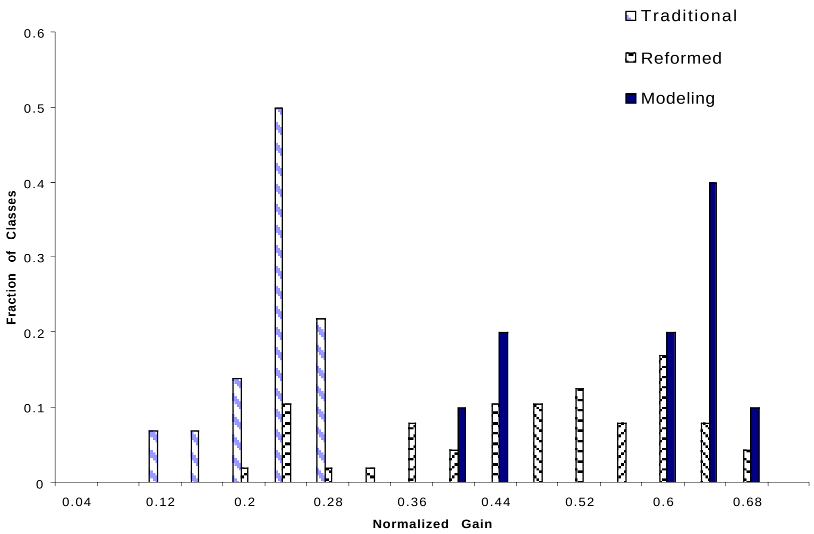


Figure 8. Comparison of physics classes using normalized gains

The FCI was chosen as one measurement tool because of its wide acceptance as a measure of the effectiveness of mechanics instruction. The thirty question 1995 version of the FCI was administered as a pre- and post-test to all courses involved in this investigation. The FCI measures student understanding of the Newtonian force concept. While the courses involved covered many other topics, suitable instruments for measuring the effectiveness of instruction in those areas have not been developed. False negatives are possible on the FCI but other research shows that they are rare (Hake, 1998). Since the FCI has been used for several years, attempts have been made to keep the instrument from being compromised. Research has shown that even when the instrument has been compromised, little change in student performance is noticed (Hake, 1998).

FCI reliabilities have been reported in the literature of 0.8 (Hake, 1998). An individual outside the ASU PER project found this reliability. Internal investigations into the reliability of the FCI yield results of about 0.9 (S. Osborne-Popp, personal communication, Fall 2000). Construct validity of the FCI has been demonstrated by a principal factor analysis. The factor analysis reveals only one factor, as determined by looking at the skree plot (see **Figure 9**). This result supports the view that the FCI measures student understanding of a unitary force concept. This factor explains 75% of the variance (S. Osborne-Popp, personal communication, Fall 2000). Validity of the FCI is also shown by the wide acceptance among the PER community (Hake, 1998, Adams & Chiappetta 1998, Yamamoto, 1998, Mazur, 1997, Cummings et al. 1999, Francis et al. 1998, Saperstein, 1995). The FCI also has strong face validity to the physics community.

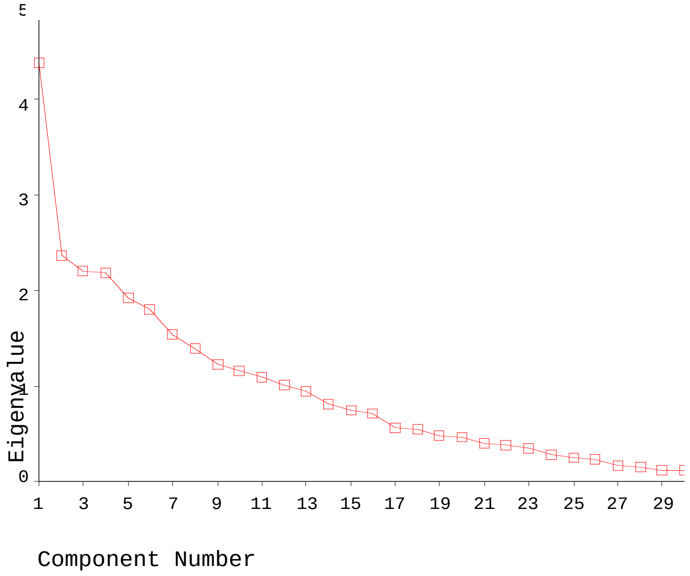


Figure 9. FCI Skree Plot

MBT

The Mechanics Baseline Test was chosen as a secondary measure because of its link to problem skills. The MBT is a 26 question multiple-choice instrument that covers all of a traditional first semester university physics course. Face validity of the MBT is high among the physics community, which can be seen by its wide use and acceptance. The calculated reliability using MBT data from this dissertation was high (r squared =0.85). The MBT was only given as a post-test because prior testing had shown that useful information was not found in pre-test results. The MBT has been given to a large number of university physics students and thus there is a large pool of existing data to compare with current results. Because the MBT requires a greater understanding of physics than the FCI, it can also be used to compare classes where high post-test FCI scores do not separate the classes. Lastly it was also chosen to evaluate student problem solving in mechanics (Hestenes & Wells, 1992).

RTOP

Observations using the RTOP were done on each of the courses in the fall of 2000. RTOP evaluations were also performed on the fall 1999 honors course and fall 1999 regular ASU physics course. The RTOP score is a numerical score between 0 and 100. The RTOP was not designed by or for modeling instruction, but was created as a means to quantify how reformed a particular classroom style is. Strong correlations between FCI normalized gains and RTOP scores have been reported (Piburn & Swada, 2000). Reliabilities for the RTOP in math/science courses have been found to be quite high (r squared = 0.954) (Piburn & Swada, 2000). Piburn and Swada found face and construct validity quite good. This paper reported that the more reformed the course, the greater the improvement on FCI normalized gains. While developing modeling discourse management the author of this dissertation never saw the RTOP evaluation to ensure that modeling discourse management was not designed to create an artificially high score on this observation protocol. Sawada and Piburn, as part of the Arizona Collaboration for Excellence in the Preparation of Teachers (ACEPT), created the RTOP (Piburn & Swada, 2000). Individuals were trained on the use of the protocol and many classes in many different disciplines have been observed using the protocol. As previously mentioned, RTOP has large correlations with FCI normalized gain scores.

MOP

The modeling observation protocol (MOP) was designed specifically for this dissertation and can be seen in **Figure 10**. The MOP is included here since it was the only instrument I created for this research. MOP is a work in progress and should be viewed as such. The MOP is not meant to quantify the use of modeling discourse management in a class. However, it was designed to help in making a subjective judgment as to the consistency with which modeling discourse management is used in a classroom. The final result of observations using MOP will be to assign a score to an observation on the lickert-scale that can be seen at the bottom of **Figure 10**. Any occurrence of an item in MOP will result in a “y” being marked. Like the RTOP, the MOP does not determine how often a technique is used, only if it is used. The MOP will only be used as a means to compare modeling use to RTOP and FCI scores, not as a quantitative analysis tool. The MOP was designed around the description of modeling discourse management given in Chapter 4. The MOP was designed to be as simple as possible while yielding an overall feel of modeling in the classroom. The observer looks for examples of the key points of modeling discourse that should be seen on a regular basis. Multiple observations are made to ensure the observer sees a wide range of activities. Since there are multiple observations, even by the same observer, inter-rater reliability is an issue. To ensure that ratings are consistent, each class was observed and taped at the same time. The tape was then used to evaluate the course 3 months later. Both evaluations yielded nearly identical results. Thus, inter-rater reliability is not a factor in this case. The instrument was tested and revised several times before being used for this dissertation by observing a course known to use modeling discourse management and one that was not. Modifications to the instrument were performed to help make sure these courses were seen as dichotomous. The final version, while not perfect, allows for a simple judgment to be made.

Class Overview:

- Evidence of small group work (y or n)
- Evidence of multi-group discussions (y or n).
- Do students make presentations to each other (y or n)
- What tools are used in presenting?
- Is consensus reached?
- Evidence of student use of models (y or n)

  If “yes” is there evidence of student understanding of why they are used (y or n)?

Seeding:

- Does the instructor interact with the small groups (y or n)?
- Are groups given hints or answers to their questions?
- Are groups given new ideas to bring to the larger discussion (y or n)?

Class atmosphere:

- Do the students help question other groups presentations (y or n)?
- Do students work out differences in ideas (y or n)?
- Are some items left for future discussion (y or n)?
- Student-Student discussion vs. Teacher Student?
- When teacher questions discussion do they address group or individual?

Discussions:

- During student discussion is the teacher taking notes (y or n)?
- Do notes involve future activities/discussion ideas (y or n)?
- Who do students speak to during discussions (teacher or student)?
- If students speak to teacher what is teacher’s response?
- Do seeded ideas come up during discussions (y or n)?
- Evidence of extended lectures (y or n)?

Overall Class Evaluation

1	2	3	4	5
no use	minimal use	some use	somewhat inconsistent use	consistent use

Figure 10. Modeling Observation Protocol

VASS

The VASS was chosen as an evaluation instrument to examine changes in student attitudes towards science as a result of modeling discourse management. Previous versions of the VASS found no positive changes in student attitudes as a result of a modeling course, traditional course, or other reformed course (Halloun & Hestenes, 1996). The results of the modified version of the VASS given in this study match those for previous versions, but the modified version more appropriately separates groups of students (Halloun, 2001). VASS results separate students into four profiles, folk, low transitional, high transitional, and expert. The VASS score is numerical; and pre-defined cutoffs for the categories mentioned have been determined. The author of the VASS argues that due to the nature and structure of the VASS, traditional measures of reliability and validity are not appropriate (Halloun, 2001). However, reliabilities for the VASS have been found to be relatively low (r squared = 0.68). It should be noted that low statistical reliability is typical for instruments of this kind. Thus, no statistical analysis will be done on the VASS other than making comparisons between distributions of students in the four categories and distributions of students in an existing pool of data in the same categories. Because of the low reliability, individual student responses and scores were not examined and compared. Since good correlations between the versions of VASS exist, comparing data from this version with data from previous versions is acceptable.

Research Design

I designed this research to utilize both existing data and data collected specifically for this dissertation. Two levels of university physics courses are included in order to investigate the effects of modeling discourse management on two different populations of students. The two levels are honors students and regular university physics courses. Three treatments are included in the honors university physics group and six treatments are included in the regular university physics. Each treatment consists of an attempt to teach students mechanics. The various modes of instruction were discussed earlier in this chapter and thus comprise a brief description of each treatment. No control group was utilized since all groups received the treatment of physics instruction. Thus only comparison groups are included in this work.

All courses covered essentially the same material, but they presented the material in different fashion and order. **Table 4** contains relevant information about instructional styles and curriculum. **Table 6** is a research design matrix for this research that shows the comparisons made. An FCI in a cell means that FCI scores (pre-test, post-test, and gain) were compared for those classes found at the beginning of the corresponding row and column. An MBT in a cell would similarly mean that the course at the beginning of that row and the top of the column would be compared using the MBT. **Table 6** shows all possible comparisons that could be made based on the data collected. However, when doing the analysis usually the honors courses were compared only to honors courses and regular courses to regular courses. This is because these are two very different populations of students. Thus, the 5 by 5 upper left matrix corresponds to the possible regular university physics course comparisons and the lower right (ignoring the pool data column) 3 by 3 matrix show the possible honors course comparisons.

Table 6. Research Design Matrix

	ASU FC 1999-2001	ASU Lecture 2000-2001	CGCC 2000-2001	CGCC 1998-1999	NCSU 2000-2001	ASU Honors 2000-2001	ASU Honors 1999-2000	ASU Honors 1996-1997	Pool Data
ASU FC 2000-2001	FCI MBT MOP	FCI RTOP MOP	FCI MBT RTOP	FCI MBT	FCI	FCI RTOP MOP	FCI MBT RTOP MOP	FCI MBT	MBT VASS
ASU FC 1999-2001		FCI MOP	FCI MBT	FCI MBT	FCI	FCI MOP	FCI MBT MOP	FCI MBT	MBT
ASU Lecture 2000-2001			FCI RTOP	FCI	FCI	FCI RTOP MOP	FCI RTOP MOP	FCI	
CGCC 2000-2001				FCI MBT	FCI	FCI RTOP	FCI MBT RTOP	FCI MBT	MBT VASS
CGCC 1998-1999					FCI	FCI RTOP	FCI MBT RTOP	FCI MBT	MBT
ASU Honors 2000-2001							FCI RTOP MOP	FCI	
ASU Honors 1999-2000								FCI MBT	MBT
ASU Honors 1996-1997									MBT

Students were not randomly assigned to the various treatment groups. Random assignment would be impossible since some of the data used in this research was collected before this research began. Therefore, this work can be regarded as quasi-experimental in nature. The research design allows for ruling out instructor effects on student conceptual improvement. As an example, the same instructor with and without using modeling discourse management taught both the ASU Honors 1999-2000 and ASU Honors 1996-1997 courses. This design allows for measuring the effectiveness of modeling discourse management while controlling for the instructor and curriculum. The design, as previously mentioned, allows the developer to be a part of the research and only measure his effectiveness in using modeling discourse management.

Collection Procedures

This section details the data collection for each of the evaluation instruments. Information on timing and administration is included.

FCI

As previously stated, some of the data for this dissertation consists of FCI pre and post scores. In all cases the pre-test was given during the first week of classes before any instruction had occurred on forces. The post-test was given during the last week of the semester. The FCI was always taken during normal class time. In all cases, the students knew their score (either pre or post) was not considered in determining their grade. However, with the exception of the honors classes, students were given a small amount of extra credit for improvement pre to post-test. Student IDs were collected with their answers to enable a match between students' pre and post scores. No other information was collected from the students. The data for the FCI and other measures was entered into Microsoft Excel by a third party and double-checked by the author.

RTOP and MOP

Observations of treatment groups were conducted using two different observation protocols. In order to judge the amount of modeling discourse management used in each course, the author performed observations of classes using MOP.   Each fall 2000 course was observed twice to ensure a reasonable range of activities was observed in each classroom. After each observation, the class for that day was given a subjective score on the Likert-scale seen in **Figure 10**. The score from the observations were averaged and that result was used as the class score for the MOP observations. Discussions with instructors from the other years were used to help determine if any aspects of modeling discourse management were used in those classes. These courses were given a score based on the information provided by the instructor using the same Likert-scale.

However, since much of this data is self-reported there exists the possibility of problems in the MOP scores.

In addition to the MOP, the RTOP was used for all courses from the fall of 1999 and 2000. The RTOP evaluation was performed by a third party to measure the amount of reformed teaching being used in each of these courses. Two different observers administered the RTOP; and the RTOP score used is the average of the two observations. Comparisons of the RTOP and MOP scores will give a better picture of the courses. Correlations between the scores will suggest that the use of modeling discourse management ensures a reformed teaching method. All three sets of data will be compared to each other to give a clearer picture of the results of using modeling discourse management in the classroom.

MBT

In all courses where the MBT was used it was part of the first semester final exam. Thus, the measure was also used as part of the student grade for the course. The students knew that the MBT was for a grade. Scoring of the multiple-choice MBT was done as right or wrong. For this study only the student responses and total score were used. No other information was obtained from the MBT.

VASS

The VASS was given the second week of March to the two classes listed in **Table 3**. Since the administration was during the second semester, only students who also took university physics from the same instructor the prior semester were counted. This instructor issue was only important for the CGCC course. Also, since the VASS was administered during the second semester there was a selection effect to be considered. Some students chose not to take the second semester or failed the course or pre-requisite for the second semester. Thus there was a decline in the enrollment for both courses using the VASS. Therefore, it could be argued that the students who did not come back would be those that make the VASS profiles more traditional looking. However, the small number of students not returning (18 at ASU and 3 at CGCC) makes this argument unlikely. As will be shown, this number of students will not have a large effect on the VASS compared to the pool data.

Analysis Procedures

The type of statistical analysis performed on each data set depends on which of the five smaller research questions defined in Chapter 1 is being investigated. Because of the data and experimental design, a multivariate analysis is not feasible and thus simple descriptive and comparative statistics were chosen (M. Thompson, personal communication, December 4, 2000). Numerous methods of visualizing the data are used, including box plots, histograms, scatter-plots and bar charts. After collecting the FCI, MBT, and VASS data, each student's paper was scored with the aid of Microsoft Excel. For the FCI, the normalized gain, or Hake gain (Hake, 1998), was calculated for each student using the pre- and post-test score. The definition of the Hake gain (g) is given in **Equation 1**.

$$g = \frac{(\text{Posttest\%}) - (\text{Pretest\%})}{100\% - (\text{Pretest\%})} \tag{1}$$

Means for each class on the FCI pre, FCI post, Hake gain, MBT and VASS were calculated. Comparisons of means and various representations presented were all done using Excel. Methods of visualizing the data were determined in consultation with the dissertation committee. Visualizations were chosen to best demonstrate various aspects of the data and to clarify differences in the various data sets. The research design matrix, **Table 6**, clearly shows all comparisons to be made. However, the reader should remember that the analysis was done separating the honors and regular university physics courses. Comparisons between honors and regular courses were not regularly performed, as these two populations were not equal at the beginning of their respective courses. Graphical comparisons of all groups were done in order to see trends in the data. The research questions and analysis specific to each are found below.

Question One

Does the inclusion of modeling discourse management into the classroom enhance student conceptual understanding of forces as measured by the FCI? I compared several first semester classes in university physics (the introductory calculus-based physics course) in an attempt to answer that question. An ANOVA was performed on all of the FCI data including the Hake gain factor (except the NCSU data, where only results are known) to look for differences. ANOVA's were performed on the pre-test scores, post-test scores, and Hake gain scores (Hake 1998). Follow-up post-hoc tests were done if differences were found to exist. A brief description of the Scheffe post-hoc test is included in **Appendix C**. The NCSU data were compared qualitatively to the FC classes as a target for a large-scale PER-based university physics course. The FCI results provided strong evidence as to the effectiveness of modeling discourse management.

Question Two

Does modeling discourse management improve student views about science? The VASS was administered to the FC 2000-2001 class and the CGCC 2000-2001 class. VASS data from these classes were compared to existing pool data to determine if modeling discourse management improved students' views of the process of science.

Only comparisons of graphical representations of distribution of students from the two modeling courses to the pool data in the four VASS categories (expert, high transitional, low transitional, and folk) were done. No statistical procedures will be performed on the VASS data. Due to several factors, this question will provide weak evidence on the effectiveness of modeling discourse management. Conclusions on this question were based on visual differences in the distributions.

Question Three

Is a class taught using modeling discourse management course viewed as highly reformed by educational researchers outside the modeling research project? Third party observers evaluated the level of reform for the six courses indicated in **Table 3** using the Reformed Teaching Observation Protocol (RTOP) (Piburn & Swada, 2000). This information helped determine the level of reform of a modeling discourse management class as seen through the eyes of trained observers. No statistical analysis of this data was performed. Conclusions on this question were based only on the differences in RTOP scores provided by third party observers. This question provides weak evidence as the effectiveness of modeling discourse management.

Question Four

Does a modeling discourse course develop problem-solving skills to the same extent as other physics courses? An ANOVA was performed on the MBT data collected for this dissertation along with baseline data from the Modeling Research Group. Follow-up t-tests were done to examine differences among pairs of classes. Results of this question were based on the outcomes of this ANOVA and post-hoc tests. This is the one question where comparisons between honors and regular courses included a statistical analysis. This question provides strong evidence as to the effectiveness of modeling discourse management.

Question Five

Is the modeling discourse management technique transferable to other instructors? This question was investigated using MOP and FCI data from the FC 00-01 course. No statistical analysis was done on the MOP data other than simple comparisons of scores to determine if possible differences exist. Since the MOP is still in the development stage, results from this question will be weak. However, it will help determine if modeling discourse management can be disseminated. The FC 00-01 course was taught by three instructors who always worked with the same subset of the class. Thus, by comparing the FCI scores for the students led by each instructor a clearer picture of how well modeling discourse management can be disseminated was found.

The results from these five questions were used to address the following question: Does the inclusion of modeling discourse management in a modeling course improve student learning and understanding? The power of the larger conclusion is tempered by the shortcomings mentioned in the analysis of each of the five smaller questions.

Table 7 lists the courses used for each specific research question.

Table 7. Table of courses with links to research questions

Institution	Year	RQ[1] 1	RQ 2	RQ 3	RQ 4	RQ 5
ASU FC	2000-2001	X	X	X	X	X
ASU FC	1999-2000	X			X	X
ASU Honors	2000-2001	X		X		X
ASU Honors	1999-2000	X		X	X	X
ASU Honors	1996-1997	X			X	
ASU Lecture	2000-2001	X		X		X
NCSU PER	2000-2001	X				
CGCC	2000-2001	X	X	X	X	
CGCC	1998-1999	X		X	X	

[1] RQ = Research Question

Chapter 4
Modeling Discourse Management

Introduction

This chapter is broken into two major sections. The first describes the impetus for creating modeling discourse management. The second section is a detailed description of modeling discourse management. Included in the last section is a narrative of a modeling activity from beginning to end.

Why Was Modeling Discourse Management Developed?

Beginning in the fall of 1995, interviews were conducted with students in an honors university physics course using one version of the modeling method. Three major observations developed from these interviews. First, students felt that the representational tools were a burden, and they only included them on exams and homework if they were specifically asked to. Second, they felt that physics homework and exams should look like those found in a math class. Third, students felt that conceptual questions were unfair and that understanding was shown by the manipulation of equations. Students held onto these beliefs even after being shown holes in their understanding of a physics concept that they had used in a traditional physics problem. Because of the large emphasis placed on representations of structure in the modeling classroom, student resistance to alternative representations was disturbing.

The instructors noted a very interesting point on an exam (Politano, 1998). One question on the exam did not specify what representational tools to be used. However, the students who used a specific representational tool rather than just equations did significantly better than those students who used only equations. Thus, to improve performance students would need to use alternative representations without instructor prompting. Therefore, a classroom environment needed to be created where students developed or adopted and used the tools.

The second observation from the interviews was that students did not see the usefulness of the models. Students felt like the models were imposed rather than being a natural tool of science. Every student interviewed felt that having to describe what model they were using and why they could use it was a waste of their time. If they got the right answer then that was all that mattered. From the interviews, it became apparent that most students felt that physics was absolute and proven and that Newton’s laws were proven facts of the universe that could be applied without discrimination. However, physics does not look upon physical laws in this manner. All physical laws have a limited range of applicability. Thus, for modeling to be successful students would have to see the models themselves as the crucial element of knowledge. No amount of prodding by the instructor would work (that was already being done without success as seen by the interviews).

In order to change the students' views of models and representations of those models, a new classroom management style was needed. The physical models are well defined in the curriculum and physics education researchers have developed many good activities for students to address misconceptions (for example Van Heuvelen, 1991b, Thornton & Sokoloff, 1990, Laws 1991b, Hestenes 1996). Therefore, the problem solution was not to change the class activities, but rather change how the class was managed. Students needed to develop the models themselves. The epistemology of science needed to be explicit. Use of shared representational tools needed to be developed collaboratively. The class needed to be a community working together like scientists through peer-peer interaction. Modeling discourse management was developed to meet these goals.

Modeling Discourse Management

Creating a classroom setting that meets the goals from the previous section required significant changes from previous modeling classroom management styles. I based on the aforementioned interviews, literature reviews, classroom learning, and personal reflection created modeling discourse management. Modeling discourse management would have to be multi-faceted and address issues from several different perspectives. There are seven major components of modeling discourse management, as seen in **Figure 11**. The names of the components and explanations that follow are my own creations based on consultation with my advisor and others in the PER community. Some of the terms used (such as Learning Community) have been used in education for an extended period of time. However, I am giving my definitions of these terms and acknowledge that others exist. Each of the aspects of modeling discourse management plays an important role in overcoming the short-comings of previous classroom management styles. The reader should note that the similarities between modeling discourse management's components and the section headings in the literature review under discourse management. The literature cited in those sections was either instrumental in the development of modeling discourse or provided evidence that that aspect would enhance a classroom management style.

- Deliberate creation of a cooperative learning community
- Explicit need for the creation of models in science (epistemology)
- Creation of shared inter-individual meaning
- Seeding
- Intentional lack of closure
- Inter-student discussion
- Formative evaluation

Figure 11. Components of Modeling Discourse Management

Creation of a Learning Community

At the beginning of the semester, the instructor must organize the class into a learning community. Before physics content is addressed, the classroom community must exist to foster student learning. The community is shaped by activities designed to encourage students to interact in a noncompetitive manner. To create this atmosphere without the pressure of "learning physics" at the same time is critical to encouraging the greatest number of students to be both involved in the discourse and prepared to be contributing members of the class (Beane, 1995). Student-student interaction is continually encouraged throughout the semester with an emphasis on cooperation.

The course begins with a community building activity. One such activity is to have small groups of students create instructions on how to make a paper airplane. The activity begins with organizing students into groups of three or four. Each group is told to create instructions on how to make a paper airplane on the paper provided. Immediately questions arise, such as:

- Can we use pictures?
- Does the plane have to fly?
- Does create mean write?

All questions are answered the same way, by repeating the charge to create instructions on how to make a paper airplane. Students are also told to decide themselves what that means and act accordingly. After the groups complete the activity, the instructions are collected and redistributed to other groups who are told to use them to construct an airplane. As the papers are being passed out to other groups, the class is told to follow the instructions exactly. Where the instructions are unclear the group must interpret as best they can. Typically students do not give this last comment much credence. As the groups begin to work the instructor passes among them asking questions. The instructor is looking for certain words or phrases in the instructions. As an example, students will often write "fold the paper lengthwise." The instructor might ask the group following that instruction what edge of the paper is the length and which is the width? Students typically respond that the long side is the length and the narrow the width. The instructor responds, “Why? What if you print on a printer in landscape mode?” The students soon see that every term on the page can be interpreted many ways. Then they gleefully create paper “airplanes” that in no way resemble what was intended by the creator of the instructions. Finally the class is then brought together for a discussion of the activity.

Discussion is best in groups of 20-30 students. If the class is larger than 30, it should be broken into multiple groups. This breakdown was done in the FC courses at ASU. This first discussion establishes a pattern for all future discussions. The students are brought into a circle with nothing inside it. The instructor explains that this will be the standard mode for class discussions. The instructor remains outside the circle during discussion. A typical student circle white board discussion can be seen in **Figure 12**. The instructor occasionally interjects a question but typically remains outside the discussion. To join the discussion the instructor must take a position in the circle. This way the instructor is seen as part of the circle and not the leader of the discussion.

Figure 12. Typical circle white board presentation

For the paper airplane exercise the students are then asked to share the difficulties in making the paper airplanes. What terms were ambiguous? What assumptions did they have to make? This portion of the discussion is typically very short. The instructor then asks the group why they might have been given this activity. What is to be learned? What role does this activity play for the rest of the classroom discussions? The discussion of these questions has been very dynamic and positive. Students quickly comment that terms need to be defined and agreed upon by the class and that pictures are often better than words. The students reach agreement on these questions quickly and without much conflict. When the discussion is winding down the instructor steps in and reviews what has been agreed upon. The instructor emphasizes the shared definition of terms and the positive tone of the discussion. Homework is given that will allow for then next class to begin with a discussion that does not seem to involve physics. In this way, students practice discussion without the pressure of “learning physics”. The homework given will be discussed later in the next section of this chapter.

The entire activity described above is completed on the first day of class to get the learning community started. The reader should note that several other modeling discourse management techniques were introduced, including seeding, creation of shared meaning, and inter-student discussion. A critical component of the modeling discourse management style is to *lay the foundation of a learning community early and continue to build the community throughout the semester*. Students are reminded of basic rules throughout the semester. The most common reminders are that only one person should talk at a time and that evaluation of other student work must be done in a positive manner.

Explicit Need for the Creation of Models in Science

Before beginning physics instruction, the instructor establishes a need for the creation of scientific models. At the end of the first day the students are given the following questions to ponder and answer with their own ideas:

What is reality?
What is science (or physics)?
Is science reality?

The second-class period begins with students working in small collaborative groups to create a white board summarizing their answers to the homework questions. Each student brings different background, experiences, and views of the world to the discussion. While the students work on the white boards the instructor seeds ideas. These ideas include the notion that science is both incomplete and in a constant state of change and evolution. Once the small groups are done, the class comes together for a discussion. As before, the discussion is done in a circle. Students are reminded to hold their white boards so that other groups can see them at all times. **Figure 12** clearly shows students holding white boards in this manner. The instructor emphasizes to the students that a goal of discussion is to reach consensus. One group is asked to present their ideas on the first question and let the discussion flow from that point.

During this discussion the instructor will often have to intercede and refocus it. The discussion is often intense and many different points of view are presented. In the end, the instructor’s goal for the discussion is to have the students come to the conclusion that no explanation is complete. A secondary goal is to have students realize that how they describe something depends on their own experiences. The description of an event will be dependent on the observer and what aspects the observer focuses on. At the end of the discussion, the instructor summarizes the agreed upon ideas and introduces the idea of a model. The instructor discusses how a model has similarities to the object it represents but certain details are missing. The discussion ends with the idea that science continuously creates models because no one model is ever complete. Scientific models do not explain reality – they only represent reproducible patterns that anyone can observe.

By the end of the second class the students have began to create a learning community and have developed the idea that science is about creating models, rather than discovering absolutes. Because of their participation in creating it, the students are developing ownership of the idea that scientific models are the basis of science.

Like the paper airplane activity the discussion of the first homework assigned was referred to often in the class. Throughout the semester students are reminded of their agreement that they are creating models and that models have limitations. Students are also reminded that it is important to communicate those limits so that others understand the model was used appropriately. Students are reminded that there is no “right answer”, as the answer you develop depends on the model used.

Creation of shared inter-individual meaning

For a discussion to be meaningful the participants must speak the same language. The whole community must agree upon and understand definitions of commonly used terms. Therefore, modeling discourse management aims to have students realize the importance of shared meaning. The paper airplane activity is designed to bring this to their attention in a memorable way.

Besides definitions, students also need to agree on concepts and scientific models. Throughout the semester discussions are focused on building such agreement. The agreed upon terms, concepts, or models can then be used freely in discussions. Terms that have not been agreed upon cannot be used in the discussions. Guidlines for discussion must also be agreed upon. One such rule is that no term can be used until a definition has been agreed upon.

Another important aspect of shared meaning is to understanding the role of communication tools. The whiteboards used is a physical tool to aid the discussion. Other tools are more essential to physics. One common tool in physics is mathematics. Equations are abstract and often difficult for students to reason with. Modeling theory uses a variety of more intuitive representational tools. Students should discuss the utility and scope of the tools. **Figure 5** shows several such tools, including system schema and force diagrams.

The use of such tools is not unique to modeling instruction. However, modeling discourse management differs from other classroom techniques in involving the students in selecting the tools and assessing their utility. Intentional lack of closure in discourse is often used as an impetus for introducing new tools. The tools are seeded to small collaborative groups that introduce them into class discussion. Since students introduce the tools, their peers are more likely to question and evaluate new tools. Ideas introduced by an authority figure, like an instructor, are often accepted without adequate thought. The shared understanding of the tools is built during whole class discussion. The tools are then available for subsequent class use.

Seeding

Seeding is the primary technique for introducing new ideas into modeling discourse. The instructor seeds a small collaborative group with a question or a hint. This group may be struggling and need some extra help or be further along and need a challenge. The seeding process is illustrated in **Figure 1** in Chapter 1.

By seeding questions, concepts, and ideas the instructor need not introduce an idea to the whole class because the small collaborative group introduces the seeded idea into the class discussion. Seeding is done during small collaborative group time so the group has time to work out details and gain ownership of the seeded idea. Thus, peers, instead of an authority figure, introduce the ideas to the whole class.

What and when to seed are the toughest questions for a modeling discourse instructor. At the start of each activity the instructor should have an agenda and goals for the activity. This much is similar to previous modeling classroom management styles. While the small groups are working the instructor looks and listens for key words or pictures from the small groups. The instructor must then formulate a question or hint. After asking a leading question, the instructor should not necessarily wait for an answer. It is usually better to leave the group to contemplate an answer of its own. A well designed seed should be direct and induce the group to move forward. Seeding should be done early in the small collaborative group work so that the students have time to work out the details of the seeded idea and gain ownership.

Seeding also stimulates broader participation in class discussion. If a group is reticent, the instructor might seed that group with a simple idea that is important but easy to introduce, or with a question that the group can ask of the whole class. If one group tends to dominate discussions the instructor should suggest to that group individually that they let other groups explain the results. The instructor can also ask certain groups to present their white boards first and ensure the most vocal groups do not start out the discussion.

Students often come up with a technique or idea that the instructor had not thought of. Therefore the instructor must be flexible and prepared to develop ideas for seeding in real time. Occasionally a seeded idea will lead to a discussion tangent to the planned activity or discussion. As long as the result is desirable, the tangent should be explored. Thus, instructor flexibility is required.

Intentional lack of closure

Previous sections stated that at the end of the whole class discussion the instructor reminds the class of what had been agreed upon and unresolved observations or problems. The instructor does not resolve the issues but merely keeps them alive. Without closure, students continue to wrestle with the issues outside class and return with new ideas to share. Thus, lack of closure can foster student thinking about the class activity between classes and keep the discussion lively. An unexpected benefit of not resolving issues before the end of class is an increase in office hour attendance.

Adequate follow-up is essential to reap benefits from lack of closure. Lack of closure is of three common kinds. First, at the end of a class the class may not have agreed upon some issues or definitions. The next activity or assignment should then be designed to help students resolve the remaining issues. The students are then given the opportunity to discuss the results with the whole class. The second kind of lack of closure comes from an incomplete class activity. For example, students working in small groups may be asked to think of at least five questions that their group has not resolved. The students then white board their questions for whole class discussion. At the end of the discussion a master list of unresolved questions from the various groups are developed. The instructor gives no closure at this time other than stating that these questions need to be answered. By the next the instructor has developed activities or homework that help students address those questions. However, by that time the students may have arrived at answers themselves.

The third kind of intentional lack of closure is more radical. Students are occasionally given a homework problem that they do not have the tools or knowledge to complete. This kind of problem should be given early in the semester. Students are typically not pleased they cannot solve the problem. Most are worried that since they did not finish the problem, their grade will suffer. However, the instructor should explain that this problem will only be graded on effort and that the class will discuss the solution. The problem should also require application of the next topic that the class is scheduled to cover. The students are prepared for a new idea because of their inability to solve the problem. While small collaborative groups are working on the problem the instructor seeds the new idea to a few groups. Follow-up activities or homework that use the new idea are given.

This last technique is ineffective unless a working learning community has been developed. A level of trust among the students and instructor must exist. Otherwise students worry more about their grade than the problem at hand. Also, the students need to see the class as built on cooperation and not competition. Otherwise, students will not share seeded ideas.

There is a one possible negative to this lack of closure. Some students use this kind of problem as an excuse for not putting sufficient effort into homework. Some students assume that any problem that requires extra effort will be discussed in class. Therefore it is important that the follow up problems to an impossible problem require thought and work. The follow up problems cannot be simple problems.

Inter-student discussion

All aspects of modeling discourse management aim to foster student-student dialogue. Inter-student discussions allow for the free flow of ideas. Real discussions and real cognitive dissonance occur more frequently when students do not feel the pressure of an authority figure questioning their ideas. A fundamental goal of modeling discourse management is for inter-student discussion to be the dominant mode of discussion. The components of modeling discourse management previously discussed contain techniques and methods to help foster inter-student interactions. However, these techniques alone are not sufficient to foster quality student-student interactions.

Two more critical issues need to be addressed: the physical layout of the classroom and the role of the instructor during the class discussion? These two issues are discussed below.

To facilitate modeling discourse a classroom should have two major components. First, it should have tables on which students can perform experiments, have small group discussions and create white boards. **Figure 13** shows what this portion of the room might look like. Second, the room should have an area for whole class discussions that is free from obstructions. When in the circle for whole class discussion there should be no tables or chairs between the students. This can clearly be seen in **Figure 12**. Having the students in this arrangement allows for students to see all white boards at the same time. Then, similar items on the white boards are quickly noted and attention can turn to differences. Since common items are not repeated, the discussion moves more quickly and deeper issues can be addressed. Also, the avoidance of unnecessary repetition keeps the students from becoming bored and allows for more questions to be addressed. Note that this overall approach is a large departure from traditional modeling white boarding.

Figure 13. Workspace for small groups in a modeling classroom

Some of the modeling discourse classrooms used in this research had less than optimal designs. However, the classrooms were physically arranged to be as close to optimal as possible. The CGCC classrooms were very well designed for modeling discourse management. A diagram of the CGCC classroom can be seen in **Appendix D**.

During the all class discussions, the instructor's actions are critical. First, the instructor should not be central to the discussion if involved at all. If the instructor is the focus, student-student interactions are unlikely to occur. Therefore, the instructor should observe discussions from an unobtrusive position. The instructor should be able to easily hear the discussion, but in a location that is difficult for students to see. The instructor should typically only intervene to enforce the agreed upon rules of discussion, refocus the discussion with a question or comment, or summarize the agreed upon ideas at the end of the discussion. Major ideas that the instructor wants addressed in discussion should be seeded during the small group work. A modeling instructor should expect silent time during the discussion and should not talk during the dead time. The instructor must resist the urge to “explain the idea” that is not being understood. This seemingly passive role is difficult for many instructors. However, the next section of this chapter discusses how the instructor is not passive.

Formative Evaluation

The whole class discussion is a critical time for a modeling discourse instructor. Though the instructor should appear passive to the students, the instructor is actively evaluating the student discussion for understanding, misconceptions, and conceptual holes. The instructor should pay careful attention and take copious notes on the discussion. However, the students should not be aware of the amount of notes the instructor is writing. The notes serve two purposes. First the notes serve as a reminder of agreed upon concepts or ideas and of questions to ask the class at the end of the discussion. Second the notes serve as information for the instructor to use in post-class reflection. Critical reflection by the instructor is essential for the success of a modeling discourse class.

After each class the instructor must evaluate the days classroom activities. The instructor must decide if the goals were met, if progress was made, if misconceptions surfaced and if correct unplanned ideas were discussed. Without adequate notes this reflection is very difficult. During reflection the instructor develops future activities that address the shortcomings of student understanding. The instructor can also modify the pace of the class to meet the students’ level of understanding or change the order of activities to better suit student needs. Reflection on the notes after each class can be time consuming. After each class the instructors using modeling discourse management spent 1-2 hours on reflection. Some of this time was in self-reflection. Self-reflection would include personal evaluation on what the instructor seeded and how well it worked. It would also include reflection on the management of the classroom discussion. Common questions reflected upon include the following:

- Did I intervene too much during discussion?
- Did I let the discussion progress too far/not far enough?
- Was more time needed for the activity?
- What ideas surfaced that I was not prepared for?

Not only is instructor reflection important, but so is student reflection. The students should be given activities and homework that requires them to evaluate their own understanding and the classes' understanding. The instructor’s job of evaluating class progress is simplified when students reflect on their own understanding and bring their unresolved questions to the class for discussion. The instructor designs activities that address the questions students have identified.

In addition to instructor and student reflection there is another critical element to formative evaluation in modeling discourse management. The instructor should keep a journal for each class taught using modeling discourse. This journal is based on the classroom notes and the reflection upon those notes. The journal allows for both reflection on the whole semester and tracking of student and instructor progress. The instructor can perform self-evaluations of discourse management. The instructor also tracks what activities have an impact on students for a specific course. Students from different courses overcome misconceptions at different rates and through different activities. The journal helps the instructor recall the types of activities or homework that has been the most effective for a particular course.

The journal also serves as a starting point for future courses. Future courses are built using activities and homework that were effective; and less effective activities are replaced with new ones. The journal helps the instructor pace the future course. The journal provides reference for future semesters on difficulties and common places misconceptions arise. Previous semester’s journals are a valuable tool for formative evaluation of current classes.

Narrative Example of modeling discourse: The ball bounce

This section follows a modeling activity from introduction to conclusion. This particular activity was chosen because it has been video taped and viewed by many people in physics education. Included with this dissertation is a Quicktime™ movie on CD-ROM of this activity that was created by Mangela Joshua and used here with permission of the developer (M. Joshua, personal communication, December, 2001). The activity was one class period long. Homework related to the activity was given and discussed during the next class. However, this narrative will only include the one class period.

Prior to this activity students had developed and deployed particle models of objects undergoing constant acceleration. Causes of motion had not been discussed – only descriptions of motion. Representational tools developed to this point included only graphs and motion diagrams. This activity is the first of the semester where students are confronted with a situation in which their existing models fail.

The activity began with a brief introduction by the instructor. The instructor showed the ball that would be used and demonstrated how to drop the ball. The instructor also instructed the students to hold the motion detector above the ball. Having described what motion the class would investigate the instructor moved on to explain expectations for the activity. The students were asked to first predict what graphs of position, velocity, and acceleration would look like for the ball. Students were allowed to drop the ball to help in making their predictions but were instructed not to take any data. Students were told to put their predictions on a white board and show them to the instructor before taking data. After showing the instructor the predictions the students were to take data and get a data set that contained at least three bounces. Students were to then compare their predictions to the actual data. Students were told to compare and contrast the data and their predictions. Lastly, the students were told to identify places in the data where the models they developed would not be appropriate. The instructor ended the instructions by stating that after all groups completed white boards a whole class discussion would occur.

The groups of students began by getting a ball and all the equipment needed for taking data. The equipment included a ball, a computer, a motion detector, and a computer lab interface. After setting up the computer the students began to make their predictions. While the students were making their predictions the instructor moved around the classroom looking at the predictions without making any comments. Groups of students started to call the instructor over to get their predictions approved so they could begin to take data. The first group was told to redo their predictions based on the fact that the motion detector was above the ball. That was the only hint given before the instructor moved on to the next group. The instructor asked the next group to identify events on the different graphs that occurred at the same time. The remaining four groups were told to make the same clarification. The instructor never commented on the actual predictions but only clarified times and location of the motion detector. Shortly thereafter, all groups' predictions were approved and students began to take data.

The instructor kept a low profile sitting at the front of the class observing while students began taking data. Students quickly realized that taking data required the group to work together. A few minutes after students began taking data the instructor started to move among the groups offering small pieces of advice on taking data. For example, the instructor pointed out objects that might interfere with the motion detector. After a few practical pieces of advice all groups began to collect data. Several groups focused on getting more bounces while others focused on getting as big of bounces as possible. Soon, all groups had data they were satisfied with and began to sketch their data on top of their predictions.

At this point the instructor began to seed the major ideas for this activity to the students. The instructor seeded three major ideas. Each idea was seeded to two groups in slightly different manners. Two groups were seeded with the question: why does the ball not bounce as high each time? Another two groups were shown that the ball was still moving downward even after making contact with the floor. The last two groups were told to focus on the data that looked like the data from previous experiments.

The seeding was done in a similar manner to the following for all groups. The initial idea or question was given to the students and then the instructor left the group. The students were left alone to struggle with the question or idea. All groups had difficulty deciding what the question or idea meant and how to use the information supplied by the instructor. However, after a brief period of frustration each group started to use the information or develop other questions for the instructor. The resulting interactions for each seeded idea will be discussed below.

The groups seeded with focusing on the data that looked familiar progressed the quickest. The students quickly identified the parts of the data that behaved like previous experiments. These results were quickly added to the predictions on the white board. However, the students brought back the instructor to ask a follow up question. The students wanted to know what to do with the portions of the graphs that did not look like previous experiments. The instructor suggested that the students try to idealize those portions of the graphs. The two groups used this information to develop similar conclusions. These students modeled the collision with the ground as occurring instantaneously and made graphs appropriate to this model.

These two groups were seeded this information for different reasons. One group was chosen because they tended to be quiet during a whole class discussion. By having a simplified model they were an appropriate group to go first during the all class discussion. The other group did not easily accept new ideas and would get actively involved in a discussion even if they never presented their white board. Therefore, they were chosen not to be presenters but questioners of the other groups new and unique ideas.

The two groups seeded with the idea that the ball was still moving downward even after contacting the floor quickly realized that the particle model of objects would not be appropriate. Both groups asked a similar follow-up question of the instructor – how do we model the ball? The instructor guided each group in developing a model that was simple, yet would behave similar to the data. After developing a model of the ball the instructor left, suggesting that the groups now decide what might be an appropriate model for the floor. The white boards from these groups focused less on the data and more on the models of the ball and floor. One group went further and focused on when each model needed to be used based on the data collected.

These two groups were chosen for this seeding because several of the students in the groups were very creative. They liked challenges and looking at more complex situations. However, these groups could also easily go astray and thus were given very specific objects and models to focus on. The instructor hoped one of these groups would present their ideas after the initial group went.

The last two groups were seeded the most complex idea. They were seeded the question of why does the ball not bounce as high each time? They were chosen because these groups tended to work faster than the other groups and needed a challenge to ensure they finished about the same time as the other groups. One of these two groups was good at making clear, understandable white boards. The other group was very good at explaining ideas and synthesizing ideas from several groups.

After being asked the initial seeded question by the instructor, both of these groups began brainstorming for ideas. The instructor left the groups so that this brainstorming could occur. Soon both groups were out of ideas and asked the instructor for a little more guidance. The instructor asked each group about what ideas they had considered. In both cases the idea of energy loss had been discussed. The instructor built on that idea and other ideas the students had about energy. The students all stated a belief that energy was conserved and something moving had energy. The instructor introduced to the groups a new tool, the energy pie chart. The groups were shown how to use the tool and told to incorporate it into their white boards. Both groups quickly synthesized the ideas into their own and developed appropriate white boards.

During this period of small group work the instructor went from group to group seeding ideas and answering questions, normally with one or more questions. The whole activity to this point had taken about one hour and fifteen minutes. The whole class was now ready for discussion. The instructor circled the class and picked a group to present their ideas first. The group picked was one of the groups seeded with the idea of ignoring the data the existing models did not explain.

The discussion was lively, friendly and engaging. Only one group was asked to present their ideas but all groups contributed something from their white boards. Nearly every student made at least one comment and those that made no comments were actively following the discussion. During the discussion the instructor only intervened for the following two reasons: to ensure only one person talked at a time and to remind students that terminology must be agreed upon before being used in discussion. The discussion lasted for an hour and fifteen minutes.

During the discussion the students presented all of the seeded ideas. Agreement was reached on many ideas. First the students came to consensus on when the models developed before this class period were appropriate to use. Second they agreed that their existing models were insufficient to explain the data. They agreed upon new models that helped explain the data. Even though the instructor seeded many of these ideas the students presented them as their own and explained them in their own words and not the instructor's. The students worked towards consensus without being competitive.

While the discussion was occurring the instructor sat behind the class with a white board. The instructor took notes on the white board. These notes included ideas to be clarified, terms to be agreed upon, and questions to ask the class. As items were clarified by the discussion the instructor removed these items from the white board. Not all items of clarification were removed from the white board by the end of the discussion. At the end of the discussion the instructor recapped what the students had agreed upon. The instructor did not address the issues for clarification on his white board. Those items were left for another class. The class ended with the instructor handing out homework using the new tool of energy pie charts. The students put away their equipment and left for the day. The total class time was two hours and forty minutes.

Conclusions

This narrative was not meant to be a complete description of the ball bounce activity. It does, however, show a typical modeling activity from beginning to end. The narrative also shows how several of the major components of modeling discourse management are used in the classroom. Other components, such as development of a learning community, were only implied. The modeling discourse class used this pattern for activities. The activities were not always lab based but the management of the classroom was always the same.

Modeling discourse management is a work in progress. The formative evaluation is not only used to guide the classroom activities but also to improve modeling discourse management. The work of development continues and new activities are continually being tried and evaluated. An outline of activities for the 1998-1999 CGCC class is located in **Appendix E**. Other modeling discourse classes used in this study used nearly the same activities.

Chapter 5

Analysis and Results

Introduction

This chapter is divided into five sections. Each section reports on the results of the analysis of one of the smaller research questions outlined in Chapter 3. Analysis and visual representations of the data are given for each of the research questions as appropriate. Results for each question are discussed in the appropriate section of this chapter. The conclusion of this chapter answers the larger research question from Chapter 1. The results and strength of the results from the five research questions are used in answering the larger research question. Course descriptions can be found in Chapter 3.

Each section begins with a summary of the data for that research question. Following the summary is a description of factors affecting the results of the analysis. In all cases group comparisons are done using $\propto < 0.05$ for rejection of the null hypothesis. Results from the testing and conclusions drawn for each research question are given. Lastly, each section ends with an answer to the appropriate research question. **Table 8** is provided as a reminder about the acronyms involved and **Table 9** serves as reminder of the various classes.

Table 8. Names and Acronyms

Full Name	Acronym
North Carolina State University	NCSU
Chandler-Gilbert Community College	CGCC
Arizona State University	ASU
Force Concept Inventory	FCI
Mechanics Baseline Test	MBT
Views About Science Survey	VASS
Reformed Teaching Observation Protocol	RTOP
Modeling Observation Protocol	MOP
Foundation Coalition	FC

Table 9. Course Information Reminder

Institution	Year	Level of Class	Modeling Curriculum	Instructional Style
NCSU	2000-2001	Regular	No	Interactive/Socratic
CGCC	1998-1999	Regular	Yes	Modeling Discourse
CGCC	2000-2001	Regular	Yes	Modeling Discourse
ASU	2000-2001	Regular	No	Lecture/Lab
ASU	1999-2000	FC[1]	Yes	Modeling Discourse Inconsistently
ASU	2000-2001	FC[1]	Yes	Modeling Discourse
ASU	1996-1997	Honors	Yes	Interactive/Socratic
ASU	1999-2000	Honors	Yes	Modeling Discourse
ASU	2000-2001	Honors	No	Lecture/Lab

[1] FC is a regular university physics course.

Research Question One

Research Question One examined FCI data to determine if modeling discourse management improved student understanding. FCI pre and post data were collected and the Hake gain was calculated for each course. The summary of this data is given in **Table 10**. The three honors courses are at the beginning of Table 3 and the remaining regular courses follow. For the Hake gain column it should be noted that many of the courses had extremely high standard deviations compared the actual hake gain score. In these cases this was because one or two students did not seem to take the post-test seriously. **Figure 14** and **Figure 15** show the same information in bar chart form. **Figure 14** shows bar charts of FCI pre and posttest scores. **Figure 15** shows Hake gains for the various courses.

Table 10. FCI Percentages and Standard Deviations for All Classes Used in Research Question One

Institution	Year	Number of Students	FCI pre-test (Standard Deviation)	FCI post-test (Standard Deviation)	Hake Gain (Standard Deviation)
ASU Honors	1996-1997	13	71.0%	86.0%	0.36
			(18.5)	(11.1)	(0.56)
ASU Honors	2000-2001	12	63.3%	80.0%	0.39
			(19.9)	(20.8)	(0.33)
ASU Honors	1999-2000	19	63.0%	82.0%	0.56
			(20.0)	(14.9)	(0.30)
ASU FC	2000-2001	60	49.3%	79.0%	0.63
			(19.9)	(15.1)	(0.24)
ASU Lecture	2000-2001	116	45.4%	60.1%	0.25
			(20.2)	(22.1)	(0.35)
ASU FC	1999-2000	44	45.3%	71.0%	0.49
			(17.1)	(15.2)	(0.24)
NCSU PER	2000-2001	55	44.1%	73.0%	0.52
			(20.7)	(19.3)	(0.25)
CGCC	2000-2001	20	41.7%	82.7%	0.73
			(21.6)	(11.8)	(0.19)
CGCC	1998-1999	15	38.0%	89.0%	0.82
			(14.7)	(11.7)	(0.17)

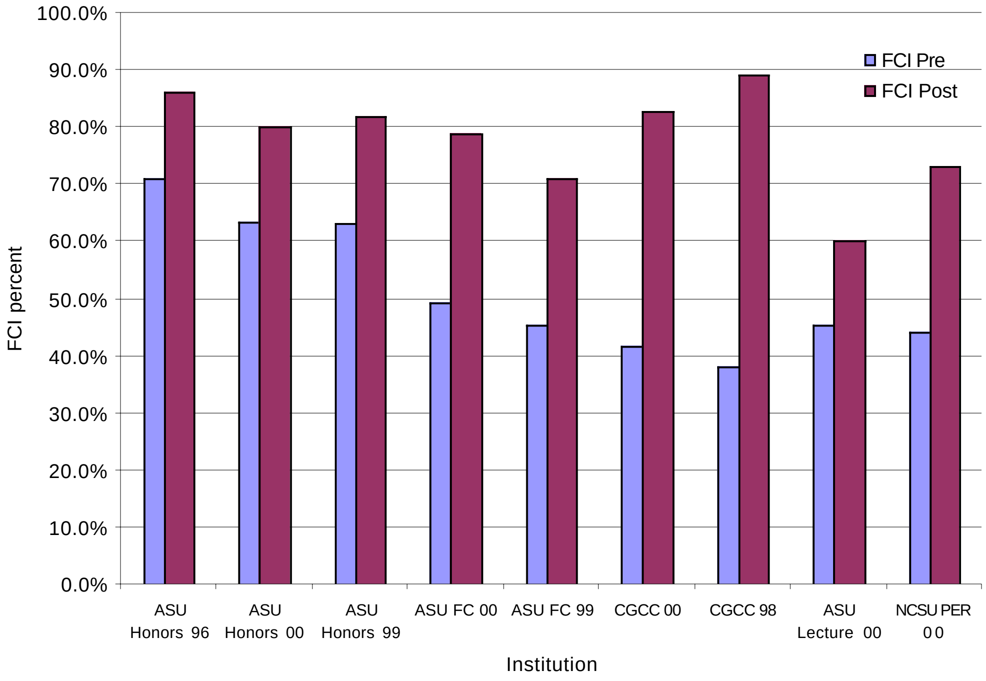


Figure 14. FCI percentages for all Courses

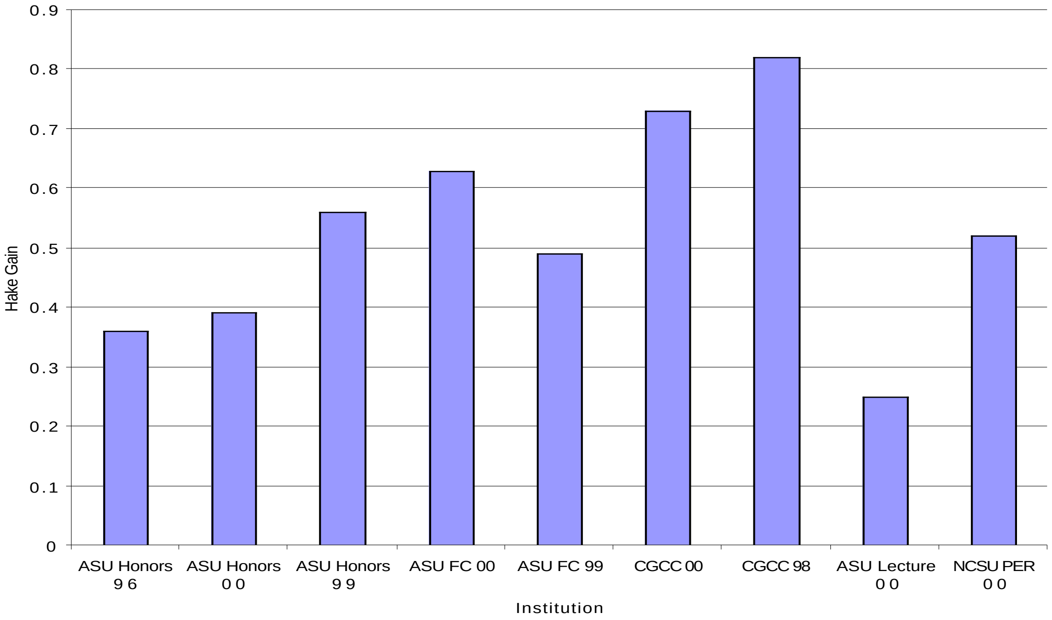


Figure 15. Hake Gains

Figure 14 and **Figure 15** show important details about the FCI scores for the honors classes. **Figure 14** makes it apparent that the honors classes begin the instruction with much higher FCI pre-test scores than regular classes. This affirms the assumption that the regular and honors courses are two different populations that should be analyzed separately. **Figure 14** also shows that the honors courses' scores began and ended at nearly the same percentages. However, **Figure 15** clearly shows that there is a large difference among the honors classes' Hake Gains. The honors 1999-2000 course clearly outperformed the others. It should be noted that this is the only honors course to use modeling discourse management.

For the regular university physics courses **Figure 14** and **Figure 15** clearly show that all courses outperformed the traditional ASU lecture course. **Figure 14** and **Figure 15** both show that the courses that used modeling discourse consistently (FC 2000, CGCC 2000, and CGCC 1998) outperformed the other courses on both FCI post-test score and Hake Gain. It should also be noted that the course that used modeling discourse inconsistently (FC 1999) was on par with the NCSU PER course using interactive/Socratic methods.

Therefore, inclusion of modeling discourse management seems to improve student understand of forces. To bolster this claim more detailed analysis was done. **Figure 16** contains a box plot for the FCI post-test scores. The threshold for understanding Newton's laws occurs at 60% with mastery occurring at 85% (Hestenes et al., 1992). **Figure 16** clearly shows that a greater percentage of students from modeling discourse classes are above the 60% threshold compared to the NCSU and ASU lecture courses. **Figure 16** also shows the CGCC courses performed as well as the honors classes post-test even though they had lower FCI pre-test scores. Therefore, the CGCC courses eliminated the gap that existed between the populations before instruction.

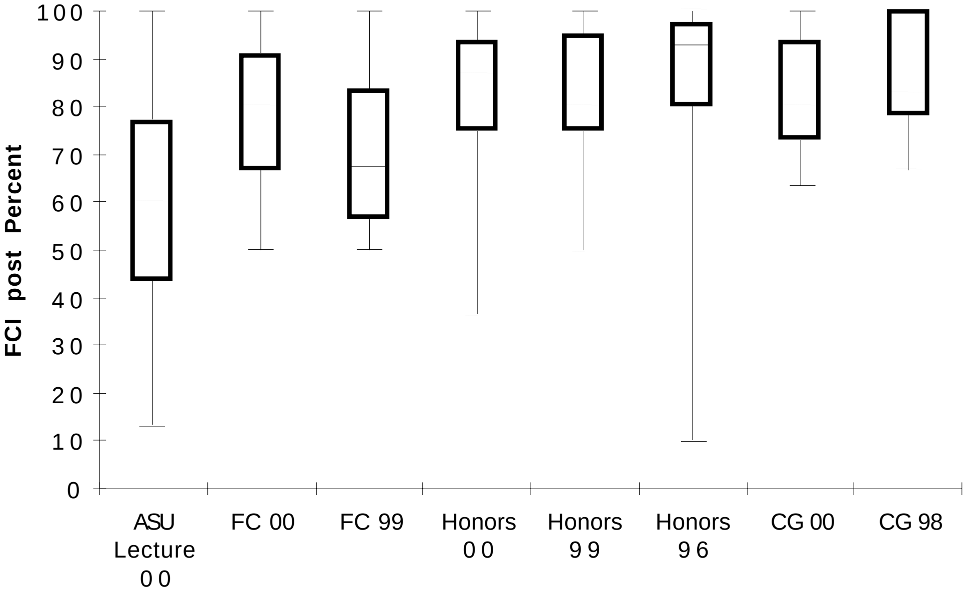


Figure 16. Box Plot of FCI Post-Test Scores

Table 11 shows effect sizes for the FCI data. **Table 11** clearly shows that the effect size is the greatest for the courses using modeling discourse management consistently. It also shows that the FC modeling discourse courses have effect sizes slightly larger than those of NCSU. Thus, while all the courses improved student scores on the FCI, modeling discourse management had a greater effect than other treatments. The effect sizes for the modeling discourse courses are large and give power to the claim that modeling discourse management helps improve student understanding of forces. **Figure 17** is a bar chart for effect size. This representation of the data from **Table 11** makes the differences between modeling discourse classes and the others more apparent.

Table 11. Effect Sizes for FCI scores

School and year	Raw Gain (post-pre)	Effect Sizes for FCI
Honors 00-01	16.7	0.80
Honors 99-00	19	1.28
Honors 96-97	15	1.35
ASU lecture 00-01	14.7	0.67
CGCC 00-01	41	3.47
CGCC 98-99	51	4.36
FC 00-01	29.7	1.97
FC 99-00	25.7	1.69
NCSU 00-01	28.9	1.50

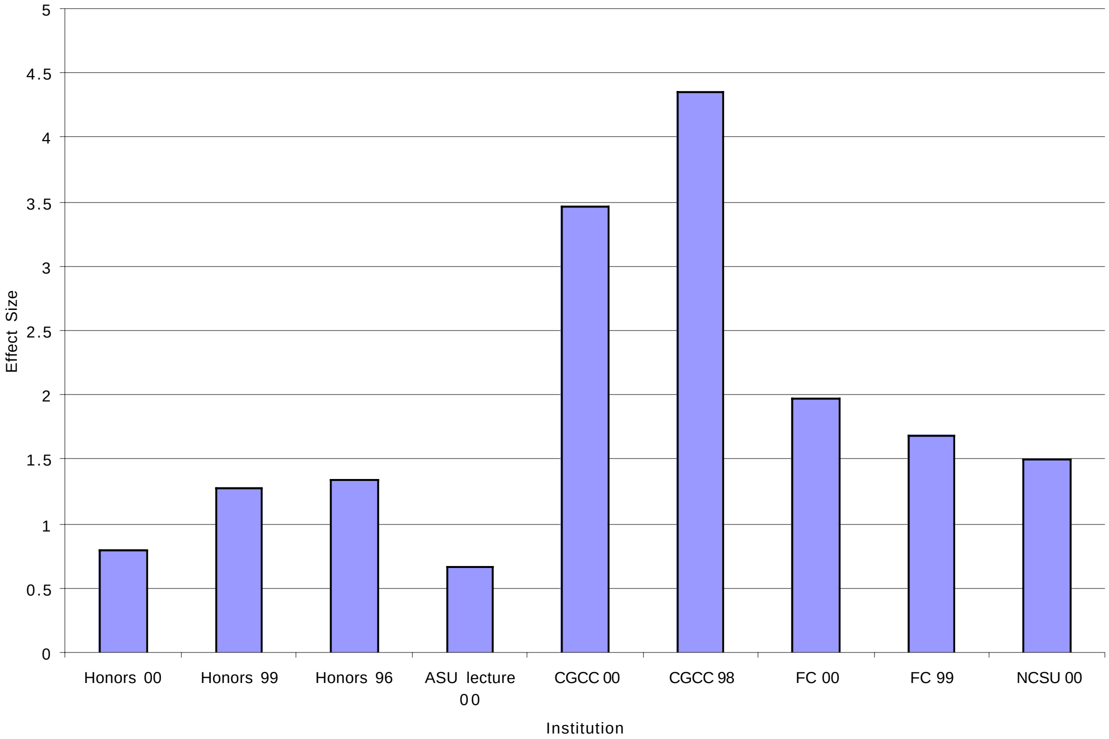


Figure 17. Effect Size Bar Chart

All of the FCI analysis to this point has involved investigation of descriptive statistics for the class scores. To determine if the differences seen in the figures and tables above were statistically significant a more rigorous analysis was done. Since I was only provided summary data for the NCSU course it will not be included in any of the remaining analysis. All ANOVAs were performed on the actual FCI score (0-30 possible) rather than a rounded percentage.

<u>Pre-test analysis</u>

Table 12 contains an ANOVA for the regular university physics courses. It shows there are no differences among the groups. Therefore, one can safely assume that the other four classes are drawn from the same population. **Table 13** contains an ANOVA comparing the honors classes. No differences exist among the honors classes, and therefore they can be assumed to come from the same population.

Table 12. Analysis of Variance for FCI Pre-Test Scores for the Regular University Physics Courses (ANOVA)

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	*F critical*
Between Groups	173.268858	4	43.3172145	1.26909959	0.28265316	2.40775222
Within Groups	8533.06055	250	34.1322422			
Total	8706.32941	254				

Table 13. Analysis of Variance for FCI Pre-Test Scores for Only the Honors University Physics Courses

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	*F critical*
Between Groups	54.5347814	2	27.2673907	0.7933385	0.4586953	3.20927995
Within Groups	1512.29921	44	34.3704366			
Total	1566.83399	46				

Post-test analysis

To determine if statistically meaningful differences exist among the FCI post-test scores an ANOVA was performed for both the regular courses and the honors courses. Table 14 contains the ANOVA for the regular courses while Table 15 contains the ANOVA for the honors courses.

Table 14. Analysis of Variance for FCI Post-Test Scores for Regular University Physics Courses

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	*F critical*
Between Groups	2163.88793	4	540.971983	17.7951808	7.1902E-13	2.40775222
Within Groups	7599.97874	250	30.3999149			
Total	9763.86667	254				

Table 15. Analysis of Variance for FCI Post-Test Scores for the Honors University Physics Courses

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	*F critical*
Between Groups	15.5677639	2	7.78388193	0.2356147	0.79107544	3.20927995
Within Groups	1453.60543	44	33.036487			
Total	1469.17319	46				

Table 14 shows that among the regular university physics courses there are statistical differences among the post-test scores. Recall that there were no differences among the pre-test scores for the same students. **Table 15** shows there are no statistical differences among the honors university physics courses. This lack of statistical difference is not surprising because of the small number of students who completed the course with both pre- and post-test scores, especially for the 2000-2001 course.

To better understand the differences that exist among the data, a post-hoc Scheffe test was performed on the regular university physics data. These comparisons are appropriate since the ANOVA has shown differences among the groups as seen in **Table 14**. The results from the Scheffe test can be seen in **Table 16**. Comparisons among the honors classes were not done since **Table 15** shows that no differences among those groups exist.

Table 16. Results of FCI Post-Test Scheffe Test

(I) CLASS	(J) CLASS	Mean Differenc (I-J)	Std. Erro	Sig.	95% Confidence Interva	
					Lower Bound	Upper Bound
ASU Lecture	FC 00	-5.6322*	.87677	.000	-8.3532	-2.9112
	FC 99	-3.2155*	.97620	.031	-6.2451	-.1860
	CGCC 00	-6.7655*	1.33494	.000	-10.9084	-2.6227
	CGCC 98	-8.2989*	1.51285	.000	-12.9938	-3.6039
FC 00	ASU Lecture	5.6322*	.87677	.000	2.9112	8.3532
	FC 99	2.4167	1.09434	.303	-.9795	5.8128
	CGCC 00	-1.1333	1.42361	.959	-5.5513	3.2847
	CGCC 98	-2.6667	1.59164	.591	-7.6061	2.2728
FC 99	ASU Lecture	3.2155*	.97620	.031	.1860	6.2451
	FC 00	-2.4167	1.09434	.303	-5.8128	.9795
	CGCC 00	-3.5500	1.48691	.226	-8.1645	1.0645
	CGCC 98	-5.0833	1.64850	.053	-10.1993	.0326
CGCC 00	ASU Lecture	6.7655*	1.33494	.000	2.6227	10.9084
	FC 00	1.1333	1.42361	.959	-3.2847	5.5513
	FC 99	3.5500	1.48691	.226	-1.0645	8.1645
	CGCC 98	-1.5333	1.88326	.956	-7.3778	4.3111
CGCC 98	ASU Lecture	8.2989*	1.51285	.000	3.6039	12.9938
	FC 00	2.6667	1.59164	.591	-2.2728	7.6061
	FC 99	5.0833	1.64850	.053	-.0326	10.1993
	CGCC 00	1.5333	1.88326	.956	-4.3111	7.3778

*. The mean difference is significant at the .05 level.

Table 16 clearly shows that all courses were statistically different compared to the ASU lecture course. Using the data from Figure 1 and Table 3 it can be concluded that all courses outperformed the ASU lecture course. Since all the other courses used modeling discourse management at some level, it can be concluded that modeling discourse is an improvement over a traditional course. **Table 16** also shows that none of the courses were statistically different from each other. This is also not surprising based on the data in **Figure 14**.

FCI gains analysis

Table 17 shows the results of an ANVOA on the gains for the regular university physics courses involved in this study. As can be seen in **Table 17,** statistical differences do exist. Therefore post-hoc Scheffe test results can be seen in **Table 18**. **Table 19** is the results of an ANOVA on the honors courses. No differences were found among the honors courses. This is not surprising due to the small number of students. While **Table**

19 shows no statistical difference among the honors courses the honors 1999 course that used modeling discourse management did have a greater effect mean gain (as seen in **Figure 15**).

Table 17. FCI Gains ANOVA for Regular University Physics courses

Source of Variation	*SS*	*df*	*MS*	*F*	*P*	*F critical*
Between Groups	8.95492452	4	2.23873113	30.3585775	1.3462E-20	2.40775222
Within Groups	18.435738	250	0.07374295			
Total	27.3906625	254				

Table 18. Results for Hake Gain Scheffe Test for Regular University Physics Courses

(I) CLASS	(J) CLASS	Mean Difference (I-J)	Std. Erro	Sig.	95% Confidence Interva	
					Lower Bound	Upper Bound
ASU Lecture	FC 00	-.3520*	.04320	.000	-.4861	-.2179
	FC 99	-.2155*	.04809	.001	-.3647	-.0662
	CGCC 00	-.4490*	.06577	.000	-.6531	-.2449
	CGCC 98	-.5392*	.07453	.000	-.7705	-.3079
FC 00	ASU Lecture	.3520*	.04320	.000	.2179	.4861
	FC 99	.1365	.05391	.174	-.0308	.3039
	CGCC 00	-.0970	.07014	.752	-.3147	.1207
	CGCC 98	-.1872	.07842	.226	-.4305	.0562
FC 99	ASU Lecture	.2155*	.04809	.001	.0662	.3647
	FC 00	-.1365	.05391	.174	-.3039	.0308
	CGCC 00	-.2335*	.07326	.040	-.4609	-.0062
	CGCC 98	-.3237*	.08122	.004	-.5758	-.0717
CGCC 00	ASU Lecture	.4490*	.06577	.000	.2449	.6531
	FC 00	.0970	.07014	.752	-.1207	.3147
	FC 99	.2335*	.07326	.040	.0062	.4609
	CGCC 98	-.0902	.09278	.918	-.3781	.1978
CGCC 98	ASU Lecture	.5392*	.07453	.000	.3079	.7705
	FC 00	.1872	.07842	.226	-.0562	.4305
	FC 99	.3237*	.08122	.004	.0717	.5758
	CGCC 00	.0902	.09278	.918	-.1978	.3781

*. The mean difference is significant at the .05 level.

Table 19. FCI Gains ANOVA for Honors Courses

Source of Variation	*SS*	*df*	*MS*	*F*	*P*	*F critical*
Between Groups	0.25939136	2	0.12969568	0.55911317	0.57572418	3.20927995
Within Groups	10.2065383	44	0.23196678			
Total	10.4659297	46				

Table 18 clearly shows that the ASU lecture course was statistically different from the other courses. The only other differences were between the CGCC courses and the FC 1999 course. The FC 1999 course did not use modeling discourse consistently and I (the modeling discourse management developer) taught the CGCC courses. Thus, it is not suprising that the CGCC courses outperformed the FC 1999 course. However, it is important to remember that any regular university physics course that used modeling discourse had better gains than a traditional course.

Conclusions

The FCI data analysis shows that using the modeling method is effective in improving student understanding of forces. It also shows that modeling discourse improves student understanding of forces even more effectively than other methods. Modeling discourse had both higher Hake gains and larger effect sizes than the other courses. Therefore, modeling discourse management is more effective in improving student understanding of forces than the other treatments.

Research Question Two

Research Question Two uses the VASS to investigate the effect of modeling discourse management on student views about science. This question cannot be completely answered based on the data collected for this dissertation, however; comparing data for students who have taken a modeling discourse management course to those who have not, still yields interesting, if not altogether convincing, data.

Problems with VASS analysis

There are several problematic features with the VASS data. First, the VASS was given at the start of the second semester. Therefore, students who either chose not to continue in a modeling discourse course, or could not because of academic difficulties, were not included in the data. Second, the version of the VASS given to the two modeling discourse classes was new; therefore, not much baseline data exists. However, results from the new VASS were calibrated to match the distributions of profiles from the older versions (Halloun, 2001). The calibration was done by picking categories cut off values such that the same percentage of students fell into the VASS categories. The cut off values were then tested against another set of data and found to give consistent percentages in the profiles. Lastly, the CGCC 00-01 course had a small sample and therefore also weakens the VASS results.

The VASS categorizes students into one of four profiles: expert, high transitional, low transitional and folk. An expert profile matches that of a scientist. A folk profile is the opposite. The two transitional profiles are exactly as they state, transitions from folk to expert.

The baseline data used for comparison was obtained from college students who had completed one semester of physics. This semester was either university physics or an algebra-based physics course. Therefore, another potential problem exists in comparing the data. However, previous studies have shown there is little difference between VASS results for university physics and for algebra-based physics (Halloun, 1997 & 2001).

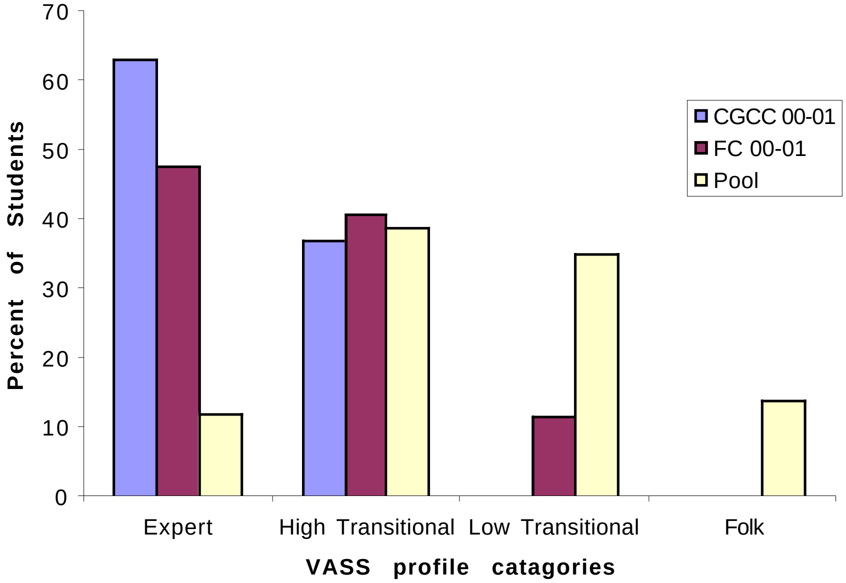


Figure 18. VASS Profile Distributions

<u>Data and Results</u>

The data for the VASS can be seen in **Figure 18**. The number of students for the CGCC class was 19 and for the FC class, 44. These numbers are lower than for the FCI for the reasons mentioned in the previous section. The number of students used in the pool was 251 students (Halloun, 2001). Looking at **Figure 18** one quickly recognizes a dramatic difference in the modeling discourse classes and the pool data. The modeling discourse classes are heavily weighted to the expert profile. The pool data is centered about the two transitional profiles. Also, the two modeling discourse courses have very similar distributions.

Therefore, it can be concluded that modeling discourse students seem to have a different view of science than traditional students after one semester. However, this claim is tempered by the comments made in the previous section. More work would be needed to strengthen this data. However, a foundation for future research has been laid.

Research Question Three

Outside observers were brought into several of the courses used in this study to observe the classes using the RTOP. The RTOP was designed as a research tool to help assess the level of reform in a particular course. This study used the RTOP to check if modeling discourse courses were seen as reformed by an outside observer.

Table 20 contains the average score on the RTOP for the six courses observed. The scores for traditional courses are in the twenties while the modeling courses are in the eighties to low nineties. RTOP scores range from 0-100, with 100 being the most reformed and 0 being the least. Therefore, differences of sixty on the RTOP are a large difference. The RTOP scores are the average of multiple visits by multiple observers. Therefore, another observation using the RTOP would most likely yield very similar results.

Table 20. RTOP Scores for Various University Physics Courses

Institution	Year	RTOP Average
ASU FC	2000-2001	83
ASU Honors	2000-2001	21
ASU Honors	1999-2000	82.5
ASU Lecture	2000-2001	26
CGCC	2000-2001	92
CGCC	1998-1998	98

Modeling discourse management is highly reformed as measured by outside observers using the RTOP. The RTOP does not measure how effective instruction is, rather the level of reform in the classroom. However, the conclusion that modeling discourse management results in a reformed class is important. Modeling discourse management was created to improve upon an existing reform. Therefore, modeling discourse should be seen as highly reformed. The fact that modeling discourse was judged to be highly reformed gives strength to the claim that modeling discourse is a reformed teaching technique. Thus, the RTOP scores give little evidence as to the effectiveness of modeling discourse management but provide strong evidence that modeling discourse is highly reformed.

Research Question Four

This question uses the MBT to help determine if modeling discourse management aids in the development of problem-solving skills. The MBT was given to six modeling classes. Four of the courses used modeling discourse management consistently, one used it inconsistently and the last used a Socratic discourse management style. The course descriptions are found in Chapter Three. Also, included in the analysis is a sampling of MBT pool data (N = 170) for various university physics courses around the country. This set of data was provided voluntarily to the modeling research group at ASU. Comparing the modeling discourse classes with pool data helped to determine if modeling discourse improved student problem-solving skills compared to more traditional university physics courses.

MBT data analysis

The MBT is not typically given as a pre-test. The material contained in the test assumes the completion of a mechanics course and therefore a pre-test yields little useful information. Therefore, all analysis was done on MBT using only post-test data. **Table**

21 gives summary data for the MBT. **Figure 19** presents the same information as **Table 21** but helps to show the dramatic differences among the MBT scores.

Figure 19 clearly shows that all courses outperformed the pool data. The FC 1996 course is the only course other than the pool data to not use modeling discourse management. The two CGCC courses are on par with the honors courses, which is quite impressive since it is likely the CGCC students started out knowing less physics (for example see FCI pre-test results). It should also be noted that the FC 99 course had the second lowest average and was the only non honors course to use modeling discourse management inconsistently (for how this was determined see page 118).

Table 21. MBT Averages and Standard Deviations

Institution	Year	MBT Percent (Standard Deviation)	Number of Students
ASU Honors	1996-1997	76.9% (7.6)	13
ASU Honors	1999-2000	68.9% (13.8)	19
ASU FC	2000-2001	66.8% (11.3)	60
ASU FC	1999-2000	58% (14.3)	44
CGCC	2000-2001	72.1% (11.5)	20
CGCC	1998-1999	79.2% (13.6)	15
Pool Data	Various	49.1% (16.6)	170

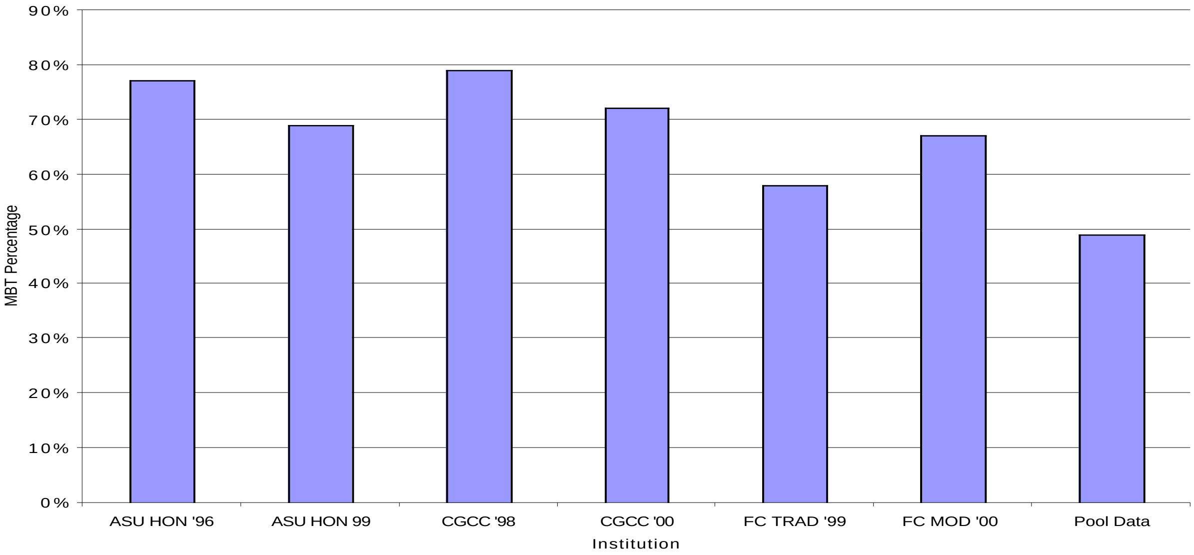


Figure 19. MBT Average Bar Chart

To determine if any differences existed among the groups on the MBT an ANOVA was performed. The results of the ANOVA are in **Table 22**. **Table 22** shows that there are differences among the groups. Therefore, a follow-up post-hoc Scheffe test was performed to determine where those differences existed. The results of the Scheffe test can be found in **Table 23**.

Table 22. ANOVA Table for MBT Scores

Source of Variation	*SS*	*df*	*MS*	*F*	*p*	*F critical*
Between Groups	2526.42913	6	421.071521	29.1786436	2.9298E-28	2.12472884
Within Groups	5007.49178	347	14.430812			
Total	7533.9209	353				

Table 23. Results of MBT Post-Hoc Scheffe Test

(I) CLASS	(J) CLASS	Mean Difference (I-J)	Std. Erro	Sig.	95% Confidence Interva	
					Lower Bound	Upper Bound
FC 00	CGCC 00	-1.3849	.97499	.918	-4.8661	2.0963
	FC 99	2.2800	.73219	.142	-.3343	4.8942
	CGCC 98	-3.2349	1.09138	.190	-7.1317	.6618
	Honors 99	-.5480	.92547	.999	-3.8523	2.7564
	Honors 96	-2.6349	1.06348	.410	-6.4321	1.1622
	Pool	4.5945*	.56031	.000	2.5939	6.5951
CGCC 00	FC 00	1.3849	.97499	.918	-2.0963	4.8661
	FC 99	3.6649*	1.01419	.045	.0437	7.2860
	CGCC 98	-1.8500	1.29753	.916	-6.4828	2.7828
	Honors 99	.8370	1.16145	.998	-3.3100	4.9839
	Honors 96	-1.2500	1.27415	.987	-5.7993	3.2993
	Pool	5.9794*	.89801	.000	2.7731	9.1858
FC 99	FC 00	-2.2800	.73219	.142	-4.8942	.3343
	CGCC 00	-3.6649*	1.01419	.045	-7.2860	-.0437
	CGCC 98	-5.5149*	1.12654	.001	-9.5372	-1.4926
	Honors 99	-2.8279	.96668	.204	-6.2794	.6236
	Honors 96	-4.9149*	1.09953	.003	-8.8407	-.9890
	Pool	2.3145*	.62604	.036	.0793	4.5498
CGCC 98	FC 00	3.2349	1.09138	.190	-.6618	7.1317
	CGCC 00	1.8500	1.29753	.916	-2.7828	6.4828
	FC 99	5.5149*	1.12654	.001	1.4926	9.5372
	Honors 99	2.6870	1.26075	.604	-1.8145	7.1884
	Honors 96	.6000	1.36528	1.000	-4.2747	5.4747
	Pool	7.8294*	1.02320	.000	4.1761	11.4827
Honors 99	FC 00	.5480	.92547	.999	-2.7564	3.8523
	CGCC 00	-.8370	1.16145	.998	-4.9839	3.3100
	FC 99	2.8279	.96668	.204	-.6236	6.2794
	CGCC 98	-2.6870	1.26075	.604	-7.1884	1.8145
	Honors 96	-2.0870	1.23667	.827	-6.5025	2.3286
	Pool	5.1425*	.84399	.000	2.1290	8.1559
Honors 96	FC 00	2.6349	1.06348	.410	-1.1622	6.4321
	CGCC 00	1.2500	1.27415	.987	-3.2993	5.7993
	FC 99	4.9149*	1.09953	.003	.9890	8.8407
	CGCC 98	-.6000	1.36528	1.000	-5.4747	4.2747
	Honors 99	2.0870	1.23667	.827	-2.3286	6.5025
	Pool	7.2294*	.99338	.000	3.6825	10.7763
Pool	FC 00	-4.5945*	.56031	.000	-6.5951	-2.5939
	CGCC 00	-5.9794*	.89801	.000	-9.1858	-2.7731
	FC 99	-2.3145*	.62604	.036	-4.5498	-.0793
	CGCC 98	-7.8294*	1.02320	.000	-11.4827	-4.1761
	Honors 99	-5.1425*	.84399	.000	-8.1559	-2.1290
	Honors 96	-7.2294*	.99338	.000	-10.7763	-3.6825

*. The mean difference is significant at the .05 level.

Conclusions

From **Table 23** and **Figure 19** several conclusions can be made. First, all of the courses using modeling discourse management at any level outperformed classes from the pool. The difference between the pool data and the modeling courses is quite dramatic as demonstrated by **Table 23** and the average scores from **Table 21**. **Table 23** also indicates that courses using modeling discourse management consistently tended to outperform the FC 1999 course that used modeling discourse inconsistently. Note that several of the regular university physics courses performed equal to honors courses. This result is unexpected since honors students tend to have a much greater understanding of physics entering a university physics course (for example, see the analysis of FCI pre-test scores earlier in this chapter). These regular university physics courses used modeling discourse consistently.

Modeling discourse developed student problem-solving skills more than traditional instruction. Also, courses that consistently used modeling discourse tended to show better results on the MBT than those that used modeling discourse inconsistently.

Research Question Five

This research question looked at the transferability of modeling discourse management from the developer to other instructors. This question will be answered using two different methods. First the MOP is being developed for use in observing and evaluating the quality of modeling discourse in a classroom. Data from the MOP was collected for several of the classes. This data does not provide strong evidence of the transferability of modeling discourse, but does provide some evidence of the transferability. Work is continuing on the MOP.

Second the FC 2000-2001 class was structured so that each of the three instructors always worked with same set of students. One of the instructors was the developer of modeling discourse management. Therefore, comparing the students of the other two instructors to the developer will allow for a judgment on the transferability of modeling discourse. The FCI data will be used for this comparison.

This section is broken into three parts. The first looks at the MOP data. The second section looks at the FC 2000-2001 data. The last section provides a conclusion for the transferability of modeling discourse based on the evidence in the previous two sections.

MOP data

Table 24 contains the average MOP score for the observed university physics courses. A few details about the data should be highlighted. First, when observing the FC 1999-2000 and FC 2000-2001 classes, I was both an observer and an instructor in the course. Therefore, the observations are based on observing the other two instructors and not myself. Each MOP score is based on a minimum of three observations, with score reported being the average of all observations.

To allow for comparisons between myself (the developer of modeling discourse management) and others, videotapes of my class were scored using MOP by a third party. The third party had attended a modeling discourse management workshop and had been applying what was learned in their classroom. To ensure consistency among the scorers we both observed a different videotape and compared scores. The scores were consistent. Therefore, **Table 24** shows a score for CGCC 2000 based on the videotaped observations.

The MOP scores range from 0-10. The higher the score the more modeling discourse management was observed. **Table 24** clearly shows a distinction between the courses that used modeling discourse and those that were traditional in nature. A large score gap exists between courses using modeling discourse (consistently or inconsistently) and those using traditional methods. Courses that used modeling discourse inconsistently (such as FC 1999-2000) had slightly lower MOP scores when compared with those using modeling discourse consistently. Therefore, MOP is able to distinguish between those courses using traditional methods, those using modeling discourse inconsistently, and those using modeling discourse consistently. Scores of classes using modeling discourse consistently were similar to my CGCC scores.

Table 24. MOP Average Scores

Institution	Year	MOP Average
ASU FC	2000-2001	8.3
ASU FC	1999-2000	6.5
ASU Honors	2000-2001	2.5
ASU Honors	1999-2000	7.5
ASU Lecture	2000-2001	2.0
CGCC	2000-2001	9.1

FCI analysis for FC 2000-2001

As a further test of the transferability of modeling discourse the FC 2000-2001 course was structured such that each instructor always worked with the same group of students. One of the instructors for the FC 2000-2001 course was the developer of modeling discourse (thus the expert). Therefore, because of the class design it was possible to compare the expert with two other instructors learning modeling discourse. The other two instructors each had one year of experience using modeling discourse management.

ANOVAs were performed on the FCI pre-test scores, FCI post-test scores, and FCI gains. In all three cases the ANOVAs found no statistical difference among the three groups. Therefore, the novices were able to aid student understanding using modeling discourse management as well as the expert and developer.

Conclusions

Both the MOP data and FCI data suggest that it is possible for instructors to learn modeling discourse management. The FCI data provides the stronger evidence for this claim. The fact that the two novice modeling discourse instructors were experienced and familiar with the modeling theory of physics should be highlighted; however, using modeling discourse was a large departure from the classroom management style used by both instructors previously. Thus, there is reasonable evidence that modeling discourse can be disseminated.

Answering the larger question

The five smaller research questions all indicate that modeling discourse management is a successful enhancement of a modeling course. Inclusion of modeling discourse management improved FCI post-test scores and gains. Evidence was presented that modeling discourse improved student understanding of science as measured by the VASS. Outside evaluators viewed a modeling discourse management class as highly reformed. Scores on the MBT were better in courses using modeling discourse management indicating better development of problem-solving skills. Lastly, evidence was presented indicating modeling discourse management can be disseminated.

Thus, based on the positive answers to all five of the smaller research questions, it can be concluded that including modeling discourse management in a modeling course does improve student learning and understanding. As mentioned previously some of the smaller questions provided stronger evidence than others in coming to this conclusion.

However, all evidences collected for this dissertation point to modeling discourse management being a success at improving student learning. Future research to strengthen the evidence for modeling discourse management can found in Chapter 6.

Chapter 6

Conclusions and Future Research

The results given in Chapter Five indicate that modeling discourse management does enhance a modeling course. However, there are some shortcomings with the methods and data used in this dissertation. Also, several of the smaller research questions need more investigation. Therefore, this chapter will detail future research that should be performed to strengthen the claims about modeling discourse management. This chapter also gives final conclusions and summarizes the research work.

Future Research

Several of the smaller research questions, while displaying positive results, were not strong indicators of the effectiveness of modeling discourse. This section discusses future research that needs to be performed to look further at these questions.

VASS research

First, more research needs to be done on the effect of modeling discourse on student views towards science. The decision to have students take the VASS was made at the end of the first semester. To better ensure modeling discourse makes changes in student views, a study needs to be performed in which students are given the VASS pre- and post-instruction. Several courses should be used, some of which do not utilize modeling discourse management. Then a clearer picture of the effect of modeling discourse on student views could be obtained.

Also, the VASS needs further research to continue to improve its reliability. **Table 5** (Chapter Three) clearly shows that the reliability of the VASS is too low to be used as a research instrument for publication. However, it should be noted that VASS reliability research is being performed. The current version of VASS makes both comparisons among classes and scoring simpler than previous versions (Halloun, 2001). Once the VASS is improved, a study as described in the previous paragraph could yield quality information on the effect modeling discourse has on student views.

MOP research

The MOP continues to develop and evolve. Many more observations need to be made using MOP to determine its shortcomings. These observations need to be performed in a variety of university physics courses so that a range of classroom management styles can be observed. The MOP will need to be revised based on these observations and once again field-tested.

Many observations with the revised MOP should be done so that reliability for the MOP can be determined. Other observers should be trained in using the MOP, so that a training protocol can be created that helps ensure inter-rater reliability. After these improvements have been created the MOP should be used in a research study that investigates levels of modeling discourse use in various physics courses. The courses chosen should include classes using modeling discourse, modeling without modeling discourse, traditional methods, and other PER-based methods. Comparisons with RTOP scores for the same classes would allow for determination of where the MOP and RTOP are similar and different. Such a comparison would allow for the use of MOP and RTOP in assessing classes to have more clearly defined roles.

Problem-solving research

This dissertation used the MBT to help determine that modeling discourse improved student problem-solving skills. However, a more systematic approach to investigating problem solving should be done. This systematic approach is the focus of **Error! Bookmark not defined.**'s dissertation (E. Brewe, personal communication, 2001). Brewe's dissertation will include comparing problem solutions from modeling discourse classes with other PER based and traditional courses. Also, future work should include investigating the effectiveness of the various modeling tools in improving student problem solving skills.

Research on modeling discourse components

Research needs to be performed on the various components of modeling discourse management. The following questions should be investigated:

1. Which of the parts of modeling discourse management described in Chapter Four are the most critical?
2. What role does each play in improving student understanding?
3. Can a course not using modeling be successful using modeling discourse management?

These three important questions need to be the focus of future research.

Research should be designed to answer all of the questions from the previous paragraph. In addition, the modeling discourse management success should be investigated at levels other than university physics. The following questions should be asked: Would modeling discourse be appropriate for a high school physics course? If not, would some of the more critical components enhance a high school physics course? To answer the previous question, research must have already been performed determining the critical element of modeling discourse.

The research on modeling discourse management is not complete, but rather only beginning. Modeling discourse needs to be dissected and the impact of each of the components investigated. Research needs to determine why modeling discourse is effective. In short, this research has shown modeling discourse to be effective, now the question must be why was it effective.

Limits on this research study

While this study has shown that modeling discourse is effective at improving student understanding, this study has limitations that should be mentioned. First, this study dealt only with students enrolled in a university physics course. All of the students were enrolled at either a community college or a large state university. Therefore, the results may not generalize to the broader population of physics courses and students.

Second, several of the courses involved in this study consisted of low numbers of students who finished the course. Therefore many of the statistical analyses may be skewed due to the small number of students in courses. This study also only dealt with students who completed the first semester of university physics. Therefore, there exists a population of students who started university physics but were not included because they did not complete the course. This fact is especially true for the honors 2000-2001 course. Therefore, the results might have been different had these students remained in the course. One highlight for modeling discourse is that courses that consistently used modeling discourse had very high retention rates, both during the first semester and into second semester.

Third, this study was geographically isolated to the Phoenix metropolitan area. Therefore, generalizing to other urban areas is reasonable. However, this does not mean the results would generalize to rural areas. Also, the results are from a southwestern city in the United States. Because of differences in students these results might not generalize to other parts of the U.S. or the world.

Final Conclusions

Modeling discourse was effective in improving student learning and understanding. This conclusion was reached by evaluation of the five smaller research questions. Each of the five research questions demonstrated that modeling discourse is improvement over traditional teaching methods and over traditional modeling classroom management. Modeling discourse increased FCI and MBT scores. Modeling discourse was seen as reformed by outside observers. Modeling discourse was transferable and seemed to improve students' views about science.

Thus, the larger research question was not answered directly but was answered by looking at several smaller questions. Doing research in this manner limits the strength of the results, however; the intent of this research was to perform a broad-based evaluation of modeling discourse management. Future research will look more in-depth at individual components of modeling discourse management and its effects on student learning. This research, however, was successful in showing modeling discourse was successful when investigated broadly.

Future research has been discussed that will not only help improve modeling discourse but also pinpoint what aspects of modeling discourse are the most effective. Future research will also help to make the conclusions of this work generalize to a larger group of physics students and physics classes. The initial development and evaluation of modeling discourse is complete. However, the formative evaluation of modeling discourse continues, as does the research on the effectiveness of modeling discourse.

REFERENCES

Adams, A., & Chiappetta, E. (1998). Students' Beliefs, Attitudes, and Conceptual Change in Traditional High School Physics Classroom. Retrieved from ERIC database 2000.

Aikenhead, G. (1987). High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge. Science Education, 71, 145-161.

Arons, A. (1981). Thinking, reasoning and understanding in introductory physics courses. Physics Teacher, 19, 166-168.

Baker, D., & Piburn, M. (1991). Process skills acquisition, cognitive growth, and attitude change of ninth grade students in a scientific literacy course. Journal of Research in Science Teaching, 28, 423-436.

Barone, T. (1993). Breaking the Mold: The New American Student as Strong Poet. Theory into Practice, 32, 236-243.

Beane, J. (1995). Curriculum Integration and the Discipline of Knowledge. Phi Delta Kappan, 76, 616-622.

Beichner, R. (1994). Testing student interpretation of kinematics graphs. American Journal of Physics, 62, 750-755.

Beichner, R. & Saul J. (2000). Student-Centered Activities for Large Enrollment University Physics (SCALE-UP). Announcer, 30, 53.

Ben-Ari, R., & Kedem-Friedrich, P. (2000). Restructuring heterogeneous classes for cognitive development: social interactive perspective. Instructional Science, 28, 153-167.

Boller, B. (1998-1999). Non-Traditional Teaching Styles in Physics. Issues of Education at Community Colleges: Essays by Fellows in the Mid-Career Fellowship Program at Princeton University, Princeton, N.J.: Princeton Press.

Bower, G., & Morrow, D. (1990). Mental Models in Narrative Comprehension. Science, 247, 44-48.

Brewe, E. (2000). Vectors vs. Coordinates. Submitted paper Summer AAPT, University of Guelph, Guelph, Canada.

Brown, J., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. Educational Researcher, 19, 32-41.

Browne, N. (1986-1987). Preconditions for Encouraging Critical Thinking on Campus. International Journal of Social Education, 1, 18-27.

Buckmaster, L. (1994). Effects of Activities That Promote Cooperation among Seventh Graders in a Future Problem Solving Classroom. Elementary School Journal, 95, 9-62.

Champagne, A., Klopfer, L., & Anderson, J. (1980). Factors Influencing the Learning of Classical Mechanics. American Journal of Physics, 48, 1074-1079.

Champagne, A., Klopfer, L., & Gunstone, R. (1981). Students Beliefs about Gravity and Motion. Problem Solving, 3, 12-14.

Clark, C. (1998). Hello learners: living social constructivism. Teaching Education, 10, 89-110.

Clement, J. (1977). Catalogue of Students' Conceptual Models in Physics, Unpublished manuscript. Amherst: University of Massachusetts.

Clement, J. (1981). Students Preconceptions in Physics and Galileo's Discussion of Falling Bodies. Problem Solving, 3, 3-5.

Clement, J. (1982). Students Preconceptions in Introductory Mechanics. American Journal of Physics, 50, 66-71.

Clement, J., Lockhead, J., & Monk, G. (1981). Translational Difficulties in Learning Mathematics. American Mathematical Monthly, 88, 28-290.

Confrey, J. (1995). How compatible are Radical Constructivism, Sociocultural Approaches, and Social Constructivism? Constructivism in Education, Editors L. Steffe and J Gale, Hillsdale, N.J.: Erlbaum.

Cummings, K., Marx, J. & Thornton, R. (1999). Evaluating innovation in studio physics. American Journal of Physics Supplement 1, 67, S38-S44.

D'Alessandris, P. (2000). What is Spiral Physics? Retrieved August 1, 2000, from http:://members.localnet.com/~daletz.

Day, T. (1995). New class on the block. The Science Teacher, 62, 28-30.

DeDobbelaere, J., Leaf, K., & Ziegler, U. (1998). Improving Behavior through Social Skills Instruction. Retrieved 2000, from ERIC database.

DeVries, R., & Zan, B. (1995). Creating a constructivist classroom atmosphere. Young Children, 51, 4-13.

Driscoll, M. (1994). Psychology of learning for instruction. Boston, MA: Allyn and Bacon.

Driver, R., Asoko, H., Leach, J., Mortimer, E., & Scott, P. (1994). Constructing Scientific Knowledge in the Classroom. Educational Researcher, 24, 5-12.

Duit, R., Roth, W., Komorek, M., & Wilbers, J. (1998). Conceptual Change cum Discourse Analysis to Understand Cognition in a Unit on Chaotic Systems: Towards an

Integrative Perspective on Learning in Science. International Journal of Science Education, 20, 1059-1073.

Dykstra, D. F., Boyle, C. F., & Monarch, I. A. (1992). Studying conceptual change is learning physics. Science Education, 76, 615-652.

Elby, A., & Hammer, D. (2000). On the substance of a sophisticated epistemology. Retrieved August 2000, from http://ww2.physics.umd.edu/~elby.

Ennis, R. (1985). A Logical Basis for Measuring Critical Thinking Skills. Educational Leadership, 27, 44-48.

Ernest, P. (1995). The One and the Many. Constructivism in Education. Editors L. Steffe and J. Gale, Hillsdale, N.J.: Erlbaum,

Escalada, L. (1997). Physics for All. Science Teacher, 64, 26-29.

Falconer, K., MacIsaac, D. & Sawada, D. (2000). Evaluating Reforms in the Physics Classroom by RTOP. Submitted paper Summer AAPT, University of Guelph, Guelph, Canada.

Francis, G., Adams, J., & Noonan, E. (1998). Do they stay fixed? High Force Concept Inventory gain scores persist long after instruction. Physics Teacher, 36, 488-490.

General Criteria For Accrediting Programs in Engineering In The United States, Accreditation Board for Engineering and Technology, Inc. Baltimore, MD (2000).

Gray, P., & Chanoff, D. (1984). When Play Is Learning: A School Designed for Self-Directed Education. Phi Delta Kappan, 65, 608-611.

Green, T. (2000). Responding and Sharing: Techniques for Energizing Classroom Discussions. Clearing House, 73, 331-334.

Guzzetti, B., Williams, W., & Skeels, S. (1997). Influence of text structure on learning counterintuitive physics concepts. Journal of Research in Science Teaching**,** 34, 701-719.

Hake, R. (1992). Socratic Pedagogy in the Introductory Physics Laboratory. Physics Teacher, 30, 546-552.

Hake, R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66, 64-74.

Halloun, I. (1984). The Use of Models in Teaching Newtonian Mechanics. Unpublished dissertation, Arizona State University.

Halloun, I. (1997). Views about science and physics achievement. The VASS story. Proceeding of the International Conference on Undergraduate Physics Education, AIP press.

Halloun, I. (1998a). Interactive Model-Based Education: An Alternative to Outcomes-Based Education in Physics. South African Journal of Science, 94, 112-125.

Halloun, I. (1998b). Schematic Concepts for Schematic Models of the Real World: The Newtonian Concept of Force. Science Education, 82, 239-263.

Halloun, I. (2001). Student Views About Science: A Comparative Survey. Unpublished manuscript.

Halloun, I., & Hestenes, D. (1985a). Common sense concepts about motion. American Journal of Physics, 53, 1056-1065.

Halloun, I., & Hestenes, D. (1985b). The initial knowledge state of college physics students. American Journal of Physics, 53, 1043-1055.

Halloun, I., & Hestenes, D. (1996). Interpreting VASS dimensions and profiles. Science & Education, 7, 553-577.

Heller, P., & Huffman, D. (1995). Interpreting the Force Concept Inventory: A Reply to Hestenes and Halloun. Physics Teacher, 33, 507-511.

Henry, F., Reed, V., & McAllister, L. (1995). Adolescents' perceptions of the relative importance of selected communication skills in their positive peer relationships. Language, Speech and Hearing Services in Schools, 26, 263-272.

Hestenes, D. (1979). Wherefore a Science of Teaching? Physics Teacher, 17, 235-242.

Hestenes, D. (1987). Towards a modeling theory of instruction. American Journal of Physics, 55, 440-454.

Hestenes, D. (1992). Modeling games in the Newtonian World. American Journal of Physics, 60, 732-748.

Hestenes, D. (1996). Modeling Method for Physics Teachers. Proceedings of the International Conference on Undergraduate Physics.

Hestenes, D. (1998). Who needs physics education research? American Journal of Physics, 66, 465-467.

Hestenes, D. (2000). Remodeling University Physics with Physics Education Research. Invited talk given to University Physics Department Chairs Conference,

Hestenes, D., Dukerich, L. & Swackhamer, G. (2000). Modeling Workshop Project: End of project report. Final report for NSF-DUE grant.

Hestenes, D., & Halloun, I. (1995). Interpreting the Force Concept Inventory: A Response to March 1995 Critique by Huffman and Heller. Physics Teacher, 33, 504-506.

Hestenes, D., & Politano, M. (1999). Restructuring University Physics. Funded NSF-DUE proposal.

Hestenes, D., & Wells, M. (1992). A Mechanics Baseline Test. Physics Teacher, 30, 159-166.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force Concept Inventory. Physics Teacher, 30, 141-158.

Hudson, H., & Liberman, D. (1982). The Combined Effect of Mathematics Skills and Formal Operational Reasoning on Student Performance in the General Physics Course. American Journal of Physics, 50, 1117-1119.

Huffman, D., & Heller, P. (1995). What Does the Force Concept Inventory Actually measure? Physics Teacher, 33, 138-143.

Jetton, T. & Alexander, P. (1997). Instructional Importance: What Teachers Value and What Students Learn. Reading Research Quarterly, 32, 290-308.

Johnson, D., Johnson, R., & Johnson, H. (1983). Cycles of Learning: Cooperation in the Classroom, Edina, Minnesota: Interaction Book Co., 1-19.

Johnston, P., Woodside, J., & Day, J. (1992). Teaching and Learning Literate Epistemologies. CELA Research Report. Retrieved from ERIC database 2000.

Karplus, R. (1977). Science Teaching and the development of reasoning. Journal of Research in Science Teaching, 14, 169-175.

Larking, J., McDermott, L., Simon, D. & Simon, H. (1980). Expert and Novice Performance in Solving Physics Problems. Science, 208, 1335-1342.

Laws, P. (1991a). Calculus-based physics without lectures. Physics Today, 44, 24-31.

Laws, P. (1991b). Workshop Physics: Learning Introductory Physics by Doing It. Change, 23, 20-27.

Laws, P., Rosborough, P., & Poodry, F. (1999). Women's responses to an activity-based introductory physics program. American Journal of Physics Supplement 1, 67, S32-S37.

Lawson, A., Drake, N., Johnson, J., Kwon, Y., & Scarpone, C. (2000). How Good Are Students at Testing Alternative Explanations of Unseen Entities? The American Biology Teacher, 62, 249-255.

Lawson, A., Lewis Jr., C., & Birk, J. (2000). Why Do Students 'Cook' Data? Journal of College Science Teaching, 29, 191-198.

Leonard, J. (1999). From Monologue to Dialogue: Facilitating Classroom Debate In Mathematics Methods Courses. Retrieved from ERIC database 2000.

Lopez-Freman, M. (1995). Multicultural science classroom. The Science Teacher, 62, 10.

Lucas, K., & Roth, W. (1996). The Nature of Scientific Knowledge and Student Learning: Two Longitudinal Case Studies. Research in Science Education, 26, 103-127.

Malcolm, I. (1991). ’All Right Then, If You Don’t Want to Do That…’: Strategy and Counter-Strategy in Classroom Discourse Management. Guidelines, 13, 17.

Maloch, B. (1999). Shifting to Student-Centered, Collaborative Classrooms: Implementing Student-Led Discussion Groups. Retrieved from ERIC database 2001.

Malone, K., & Reiland, B. (1995). Exploring Newton’s third law. Physics Teacher, 33, 410-411.

Maloney, D. (1987). Ranking Tasks: A new type of test item. Journal of College Science Teaching, 16, 510-515.

Maloney, D. (1990). Forces as interactions. Physics Teacher, 28, 386-390.

Maloney, D., O’Kuma, T., Hieggelke, C., & Van Heuvelen, A. (2001). Surveying students’ conceptual knowledge of electricity and magnetism. American Journal of Physics Supplement 1, 67, S12-S23.

Mazur, E. (1997). Peer Instruction, Upper Saddle River, N.J., Prentice Hall, Inc.

McCloskey, M., Caramazza, A. & Green, B. (1980). Currvilinear Motion in the Absence of External Forces. Science, 210, 1139-1141.

Mcdermott, L., & Shaffer, P. (1998). Tutorials in Introductory Physics. Upper Saddle River, N.J., Prentice Hall, Inc.

McGinn, M., & Roth, W. (1999). Preparing Students for Competent Scientific Practice: Implications of Recent Research in Science and Technology Studies. Educational Researcher, 28, 14-24.

Mestre, J., & Touger, J. (1989). Cognitive Research—What’s in It for Physics Teachers? Physics Teacher, 27, 447-456.

Minstrell, J. (1982). Explaining the “At Rest” Condition of an Object. Physics Teacher, 20, 10-14.

Mortimer, E., & Machado, A. (2000). Anomalies and Conflicts in Classroom Discourse. Science Education, 84, 429-444,

Musheno, B., & Lawson, A. (1999). Effects of Learning Cycle and Traditional Text on Comprehension of Science Concepts by Students at Differing Reasoning Levels. Journal of Research in Science Teaching, 36, 23-37.

Piaget, J. (1964). Part I, Cognitive development in children: Piaget development and learning. Journal of Research in Science Teaching, 2, 176-186.

Piaget, J. (1970). The Child’s Conception of Movement and Speed, Translated by Holloway and Mckenzie. Routledge and Kegan, London.

Piburn, M. & Swada, D. (2000). RTOP

Pierce, C. (1877). The Fixation of Belief. Popular Science Monthly, 12, 1-15.

Prawat, R., & Folden, R. (1994). Philosophical Perspectives on Constructivist Views of Learning. Educational Psychology, 29, 34-48.

Redish, J. (1994). Implications of cognitive studies for teaching physics. American Journal of Physics, 62, 796-803.

Redish, J., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in physics. American Journal of Physics, 66, 212-224.

Reed, V., Mcleod, K., & McAllister, L. (1999). Importance of Selected Communication Skills for Talking with Peers and Teachers: Adolescents' Opinions. Language, Speech and Hearing Services in Schools, 30, 32-49.

Reynolds, K., & Nunn, C. (2000). Engaging Student Participation and the Instructional Factors That Shape It. ASHE Annual Meeting Paper. Retrieved 2000, from ERIC database.

Roth, W. (1995). Inventors, Copycats and Everyone Else: The Emergence of Shared Resources and Practices As Defining Aspects of Classroom Communities. Science Education, 79, 475-502.

Roth, W. (1996). The Co-evolution of Situated Language and Physics Knowing. Journal of Science Education and Technology, 5, 171-191.

Roth, W. (1997a). From Everyday Science To Science Education: How Science and Technology Studies Inspired Curriculum Design and Classroom Research. Science and Education, 6, 373-396.

Roth, W. (1997b). The Local Production of Order in Traditional Science Laboratories: A Phenomenological Analysis. Learning and Instruction, 7, 107-136.

Roth, W. (1997c). Why May students Fail To Learn from Demonstrations? A Social Practice Perspective on Learning in Physics. Journal of Research in Science Teaching, 34, 509-533.

Roth, W. (1998). Teacher-As-Researcher Reform: Student Achievement and Perceptions of Learning Environment. Learning Environments Research, 1, 75-93.

Roth, W., & Bowen, M. (1995). Knowing and Interacting: A Study of Culture, Practices and Resources in a Grade 8 Open-Inquiry Science Classroom Guided by a Cognitive Apprenticeship Metaphor. Cognition and Instruction, 13, 73-128.

Roth, W., Boutonne, S., Lucas, K., & McRobbie, C. (1999). One Class, Many Worlds. International Journal of Science Education, 21, 59-75.

Roth, W., & McGinn, M. (1997). Science in Schools and Everywhere Else: What Science Educators Should Know about Science and Technology Studies. Studies in Science Education, 29, 1-43.

Roth, W., & McGinn, M. (1998). Inscriptions: Toward a Theory of Representing Social Practice. Review of Educational Research, 68, 35-39.

Roth, W., McGinn, M., & Bowen, M. (1996). Applications of Science and Technology Studies: Effecting Change in Science Education. Science, Technology and Human Values, 21, 454-484.

Roth, W. & Roychoudhury, A. (1992). The Social Construction of Scientific Concepts or the Concept Map as Conscription Device and Tool for Social Thinking in High School Science. Science Education, 76, 531-557.

Roth, W., & Roychoudhury, A. (1993). The Concept Map as a Tool for the Collaborative Construction of Knowledge: A Microanalysis of High School Physics Students. Journal of Research in Science Teaching, 30, 503-534.

Roychoudhury, A., & Roth, W. (1992). Student Involvement in Learning: Collaboration in Science for Preservice Elementary Teachers. Journal of Science Teacher Education, 3, 47-52.

Saperstein, A. (1995). Learning via study versus learning via living. Physics Teacher, 33, 26-27.

Scherr, R., Shaffer, P., & McDermott, L. (2000). The Multi-Layered Nature of Student Difficulties: An Example from Special Relativity. Summer AAPT meeting.

Scott, P. (1998). Teacher Talk and Meaning Making in Science Classrooms: A Vygotskian Analysis and Review. Studies in Science Education, 32, 45-80.

Sokoloff, D., & Thornton, R. (1997). Using interactive lecture demonstrations to create an active learning environment. Physics Teacher, 35, 340-347.

Stage, F., Muller, P., Kinzie, J., & Simmons, A. (1998). Creating Learning Centered Classrooms. What Does Learning Theory Have to Say? Retrieved 2000, from ERIC database.

Starver, J. (1998). Constructivism: sound theory for explicating the practice of science and science teaching. Journal of Research in Science Teaching, 35, 501-520.

Thornton, R. (1997). Conceptual dynamics: Following changing student views of force and motion American Institute of Physics Conference Proceedings No. 399 The Changing Role of Physics Departments in Modern Universities: Proceedings of the International Conference on Undergraduate Physics Education, editors Redish, E & Rigden J., 241-266.

Thornton, R., & Sokoloff, D. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. American Journal of Physics, 58, 858-867.

Thornton, R., & Sokoloff, D. (1998). Assessing student learning of Newton's laws: the force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula. American Journal of Physics, 66, 338-352.

Tobin, K. (1990). Research on science laboratory activities: In pursuit of better questions and answers to improve learning. School Science and Mathematics, 90, 403-418.

Tomanek, D. (1994). A Case of Dilemmas: Exploring My Assumptions about Teaching Science. Science Education, 78, 399-414.

Townsend, J. (1998). Silent Voices: What Happens to Quiet Students During Classroom Discussions? English Journal, 87, 72-80.

Trowbridge, D., & McDermott, L. (1980). Investigation of Students Understanding of the Concept of Velocity in One Dimension. American Journal of Physics, 48, 1020-1028.

United States Department of Education. (2000). Exemplary and Promising Educational Technology Programs.

Van Heuvelen, A. (1991a). Learning to think like a physicist: a review of research-based instructional strategies. American Journal of Physics, 59, 891-897.

Van Heuvelen, A. (1991b). Overview, Case Study Physics. American Journal of Physics, 59, 898-897.

Vygotsky, L. (1962). Thought and Language. Cambridge, Ma: MIT Press.

Wells, M. (1987). Modeling Instruction in High School Physics, Unpublished dissertation, Arizona State University.

Wells, M., Hestenes, D., & Swackhamer, G. (1995). A modeling method for high school physics instruction. American Journal of Physics, 63, 606-619.

Wheeler, D., & Charoenkul, N. (1998). Whole Vectors. Physics Teacher, 36, 274.

Wheeler, D., & Wheeler, C. (1998). Whole Vectors. Physics Teacher, 36, 274.

Yamamoto, T. (1998). A Comparison of the Quality of Integrated Freshman-Year Curriculum for Science, Engineering and Mathematics with a Conventional Curriculum. Retrieved 2000, from ERIC database.

Younger, M. (1999). I think I can! A first-grade teacher takes a successful football coach's philosophy into the classroom. Instructor, 108, 72.

Zollman, D. (2000). http://www.phys.ksu.edu/perg