The Effects of a Model-Based Physics Curriculum Program with a Physics First Approach: a Causal-Comparative Study

Ling L. Liang · Gavin W. Fulmer · David M. Majerich ·
Richard Clevenstine · Raymond Howanski

Abstract The purpose of this study is to examine the effects of a model-based introductory physics curriculum on conceptual learning in a Physics First (PF) Initiative. This is the first comparative study in physics education that applies the Rasch modeling approach to examine the effects of a model-based curriculum program combined with PF in the United States. Five teachers and 301 students (in grades 9 through 12) in two mid-Atlantic high schools participated in the study. The students' conceptual learning was measured by the Force Concept Inventory (FCI). It was found that the ninth-graders enrolled in the model-based program in a PF initiative achieved substantially greater conceptual understanding of the physics content than those 11th-/12th-graders enrolled in the conventional non-modeling, non-PF program (Honors strand). For the 11th-/12th-graders enrolled in the non-PF, non-honors strands, the modeling classes also outperformed the conventional non-modeling classes. The instructional activity reports by students indicated that the model-based approach was generally implemented in modeling classrooms. A closer examination of the field notes and the classroom observation profiles revealed that the greatest inconsistencies in model-based teaching practices observed were related to classroom interactions or discourse. Implications and recommendations for future studies are also discussed.

Keywords Physics education · Model-based physics curriculum · Introductory physics courses · Conceptual learning · Physics first · Rasch measurement

L. L. Liang (✉)
Department of Education, La Salle University, Philadelphia, PA 19141, USA
e-mail: liang@lasalle.edu

G. W. Fulmer
National Science Foundation, Arlington, VA 22230, USA

D. M. Majerich
Temple University, Philadelphia, PA 19122, USA

R. Clevenstine · R. Howanski
Ridley School District, Folsom, PA 19033, USA

Introduction

A central theme of the recent Standards-based reform movements has always been on what, how, and when to teach science to school children. For instance, according to the *National Science Education Standards* and the new framework of scientific proficiency, it is recommended that science instruction should focus on the most fundamental scientific ideas and provide students with opportunities to engage in exploration, knowledge generation, explanation, evaluation and modification, and participation in practices and discourse of science (National Research Council [NRC] 1996, 2000, 2007). Furthermore, to improve coherence and connection between the science courses in high schools, some leading American scientists and educators have also been advocating that the traditional sequence of core sciences (i.e., the biology-chemistry-physics order) be replaced with an alternative, three-year core curriculum sequence of "physics-chemistry-biology" in US high schools (also known as the "Physics First" (PF) movement; Lederman 2001; Pasero and Fermilab Education Office 2003). This is because the Physics First approach corresponds to the historical changes in science content knowledge over the development of modern science: (1) comprehending chemistry based on atomic or particulate models relies on an understanding of physical

principles and physics concepts, and (2) modern biology requires understanding of both physics and the chemical functions of molecules such as DNA and proteins.

A report based on the 2005 Nationwide Survey of High School Physics Teachers estimated that some variant of "Physics First" had been adopted at about 8% of private schools and 3% of public schools in the United States (Neuschatz et al. 2008). The survey results indicated that the PF approach did have the expected, large impact on student enrollment in high school physics courses. Proponents of the physics-chemistry-biology sequence argue that freshmen who study physics first will have more opportunities to interact with concrete and familiar phenomena and improve their ability to conduct scientific experiments. In addition, it is argued that the PF students will be able to practice their algebra skills that they learned either concurrently or in previous years (Ewald et al. 2005). Other effects of the Physics First sequence and some positive outcomes reported in conference proceedings or doctoral dissertations include increased student enrollments in challenging science courses, improved scores on standardized tests of science and mathematics, enhanced abilities in scientific reasoning and problem solving, and increased student interest in science (Ewald et al. 2005; Mountz 2006; Schuchardt et al. 2008). However, considering the wide variety in the nature and content of ninth-grade physics courses throughout the United States, the empirical studies in this area have been especially limited in number and exhibit little variation in data analysis methods. In this paper, we apply a Rasch modeling approach to re-examine student conceptual learning resulting from the implementation of a model-based physics curriculum (Wells et al. 1995) as part of the PF initiative in a public high school in the United States. The findings in this study will contribute to the research literature on evaluation of model-based physics curriculum programs combined with PF and advancement of alternative data analysis methods in physics education research.

Model-Based Physics Instruction and Its Association with PF

In the following sections, we provide a brief review of theories and applications most relevant to a model-based approach in physics education and its association with the Physics First initiatives. In the reform-based documents, models have consistently been recognized as one of the major unifying ideas that transcend disciplinary boundaries and pervade all scientific, mathematical, and technological fields (AAAS 1989; NRC 1996). A scientific model can be defined as a representation of structure that abstracts and simplifies a system to allow one to make explanations and predictions (Schwarz et al. 2009; Wells et al. 1995). Models may include physical objects, analogies, diagrams, graphs, computer programs, and mathematical relationships. In science, models are often created by scientists to describe and explain observed phenomena and make predictions, and then revised, refined, or changed through ongoing testing and discourse within the scientific community. In other words, model development, validation, deployment, revision, and discourse are fundamental aspects of scientific practice (NRC 1996, 2000).

In physics education, Hestenes and his colleagues have been advocating a model-based instructional approach (or modeling instruction) for more than 20 years (Hestenes 1987; Wells et al. 1995). In their model-based introductory physics curriculum on mechanics, for instance, course content is organized around a small set of basic mathematical models evolving with progressively increased empirical and rational complexity (e.g., Free Particle Model—Particle with Constant Velocity; Uniformly Accelerated Particle Model in One Dimension; Particle Models in Two Dimensions; Central Force Model; and Impulsive Force Model). In alignment with the practice of physics, instruction is organized into *modeling cycles* which move students systematically through all phases of model development, evaluation, and application in concrete situations. During the instruction, students conduct investigations in small cooperative learning groups, and constantly model physical objects and processes using verbal, diagrammatic, graphical, and algebraic representations. Modeling teachers are also informed by students' preconceptions of physics, and guide student discourse through scaffolding and asking probing questions (e.g., Socratic dialog) to elucidate models (Hake 1992; Wells et al. 1995). In this paper, our literature review focuses on the theories and applications most relevant to the model-based introductory physics curriculum developed by Hestenes and his physics education team.

As a particular case of inquiry-oriented pedagogy, the development and implementation of model-based physics curriculum in this study are in line with multiple theoretical perspectives and their applications in teaching such as constructivism (i.e., learners construct their understandings through interactions with the physical and/or social environment; Piaget 1970; Vygotsky 1978), conceptual change theories drawn on the history and philosophy of science and cognitive psychology (e.g., Kuhn 1970; Giere 1988; Posner et al. 1982; Thagard 1992), learning cycle inquiry model (Karplus 1977), and model-based learning and instructional theories (Clement 1989; Hestenes 1987; Lesh and Doerr 2003; Wells et al. 1995). The organization of the modeling physics curriculum and instruction is also aligned with research on learning, instruction, and the structure of scientific knowledge. From a cognitive perspective,

concepts are organized in conceptual domains that are often represented in terms of a hierarchy or taxonomy (Hamilton and Ghatala 1994). Rosch and her associates (Rosch 1978, cited in Hamilton And Ghatala 1994, p. 162) further differentiated three levels of a conceptual hierarchy or taxonomy: superordinate level (such as the concept of "vehicle"), basic level (or middle level, such as the concepts of "car," "bus," or "truck"), and subordinate (such as more specific types of cars, "sports car," or "four-door sedan"). The basic level concepts were suggested to be used in organizing knowledge about the world as the concepts at this level capture the most important aspects of that class of object without being too abstract or too detailed. In other words, the basic-level categories in the middle of a conceptual hierarchy constitute the most accessible, efficient and reliable building blocks in the process of knowledge construction and development.

In a scientific paradigm with a scientific conceptual system shared by the members of a particular scientific community, models are in the "middle" level (or basic level) of conceptual hierarchy, between scientific theories and specific concepts. For instance, the Newtonian theory consists of a set of particle models: free particle model, uniformly accelerated particle model, particles in uniform circular motion, and so on. To further study each model, many concepts such as "velocity" and "acceleration" are defined or created to describe the model. The model-centered structure of scientific knowledge ensures theory coherence and consistency from an epistemological perspective, and facilitates the knowledge development from a cognitive perspective (Giere 1988; Halloun 2004).

From a pedagogical perspective, a model-based approach with a focus on the development of coherent explanations and arguments through *modeling cycles* is well aligned with the frequent calls for "teaching science as practice" (NRC 1996, 2000, 2007). It is expected that the modeling instructional approach should help students develop deeper scientific understanding with improved inquiry skills, facilitate transfer, and promote more accurate and productive epistemologies of science. Indeed, compared to traditional teaching-by-telling methods, the modeling instruction has resulted in significantly greater student conceptual understanding and problem-solving at both secondary school and college/university levels (e.g., Brewe et al. 2010; Hestenes et al. 1992; Wells et al. 1995; Vesenka et al. 2002). It was also reported that the students experiencing the modeling physics instruction developed more expert-like knowledge structure and problem-solving skills associated with greater metacognitive awareness (Malone 2008).

Positive results of model-based instruction on physics learning from other variations of the modeling cycles have also been reported (Clement 1989, 2008; White and Frederiksen 2000). For instance, Clement and his colleagues

developed a model-centered instructional sequence focusing on "model generation, evaluation, and modification cycles" or GEM cycles, through the coordinated use of multiple analogies (bridging analogies), discrepant events, and discourse. The GEM approach has been successfully applied in physics instruction and other science subjects (Clement 1989, 2008). Another effective example of model-based curriculum in physics instruction was called "Model-Enhanced ThinkerTools (METT)" (Schwarz and White 2005), a middle school curriculum that enable students to create computer models and engage in discussions about models and the process of modeling. The METT curriculum was built on its predecessor created in the initial ThinkerTools Inquiry Project (White and Frederiksen 2000), by adding a *metamodeling knowledge* (i.e., knowledge about modeling) component to the inquiry cycle (i.e., "question $\Longrightarrow$ hypothesize $\Longrightarrow$ investigate $\Longrightarrow$ analyze$\Longrightarrow$ model $\Longrightarrow$ evaluate $\Longrightarrow$ question $\Longrightarrow$"). The research findings indicated benefits and importance of developing students' metacognitive knowledge through "scaffolded inquiry" in a model-based curriculum (White and Frederiksen 2000). In addition, it was suggested that an emphasis on model-based inquiry, accompanied by the development of metamodeling knowledge, can facilitate learning content knowledge and developing skills as well as understanding of the scientific enterprise (Schwarz and White 2005).

Connecting the PF initiative with the model-based physics program is a more recent development in some high schools in the United States. For instance, in a recent study on the effectiveness of ninth-grade modeling physics approach on student conceptual understanding by O'Brien and Thompson (2009), it was found that ninth-graders are more sensitive to the instructional method used. The model-based program appeared to be more effective than the non-modeling approaches for the ninth-graders (non-honors classes), as measured by a Mechanics Concept Survey. The study also found that the number of weeks spent on mechanics did not have a statistically significant effect on student conceptual learning of mechanics.

In the Physics First movement, various curriculum programs such as Hewitt's *Conceptual Physics*, *Active Physics* (Eisenkraft 2010), and *Modeling Physics* (Wells et al. 1995) have been adopted in schools nationwide. Many advocates of PF support a more conceptual ninth-grade physics course as most students do not learn some essential mathematics content such as trigonometry until 10th grade. However, such an approach diminished the value of those PF proponents who consider physics to be a mathematical science (Goodman and Etkina 2008). Given the consistent emphasis of modeling in both mathematics and science education standards documents (Common Core State Standards Initiative 2010; NRC 1996; National

Council of Teachers of Mathematics [NCTM 2000), the implementation of a ninth-grade modeling physics program focusing on developing mathematical models would facilitate a natural integration and coordination of physical science and mathematics (algebra). For instance, the basic mathematical models such as "constant rate" and "constant change in rate" are consistently used in the study of physics and connected with direct experience rather than abstract representations. In other words, model-based physics curriculum combined with PF provides an ideal context for students to learn and master algebra and the unifying modeling theme for the ninth graders. However, there are no empirical research studies that have systematically examined the effects of such coordination and integration between mathematics and physics (and other science subjects).

In summary, more empirical research about PF is needed. While there is some general consensus that model-based, inquiry-centered science curriculum programs are more effective on fostering students' conceptual understanding of both contents and/or the nature of scientific knowledge development, many earlier studies did not analyze the curriculum implementation data. In addition, most physics education research has used raw data (e.g., total scores, total percent scores, gain scores, normalized gains, etc.) in analysis. Such treatment of data in analysis may have led to inaccurate results or conclusions (Liu and Boone 2006). In response to the advancement in the measurements in human sciences, there is also an apparent need to reexamine the efficacy of the widely used assessment tools (such as the Force Concept Inventory) with alternative or updated measurement models.

Research Questions

As part of a larger research project, the present study focuses on the examination of the effects of a model-based introductory physics curriculum combined with PF on students' conceptual learning. Specifically, the following research questions were investigated: When compared to a conventional physics program, does the introductory modeling physics curriculum combined with PF result in the students' greater understanding of physics concepts? If so, what specific teaching practices associated with the model-based approach might have played a significant role in the students' conceptual learning?

Methodology

Given the empirical nature of our inquiry that explores differential effects of two science programs within their real-life contexts, a causal-comparative research design was adopted for this study (Gay et al. 2006). The impact of the model-based science program upon the development of student conceptual learning was measured by the conceptual test. The level of classroom implementation of the model-based instructional strategies was determined based on the classroom activity survey completed by students. In addition, classroom observations were also conducted in selected model-based lessons.

Participants and Setting

Before the study was launched, we identified and contacted multiple Mid-Atlantic high schools (including modeling and non-modeling, PF and non-PF schools) with carefully matched student demographics and the statewide, standardized reading and mathematics scores in our existing research database. Two schools responded to our invitation and volunteered to participate, each representing a different case: *modeling with PF* versus *non-modeling with non-PF*. There were five teachers and 301 students (in grades 9 through 12) involved in the study. The students in both schools were predominantly Whites (91–95%) from middle income households as defined by the state ($37,501–$57,000). The science courses studied in this paper were all one semester in length, for 85–90 min a day following a block schedule.

In the school adopting the model-based physics program, all participating physics teachers completed a three-week long summer professional development course on modeling instruction at Arizona State University and then implemented the model-based physics program in their introductory physics courses. The course content is organized around a small set of basic models, including units such as "Free Particle Model—Particle with Constant Velocity," "Uniformly Accelerated Particle Model (one-dimension)," "Particle Models in Two Dimensions," and "Central Force Model—Uniform Circular Motion." Instruction is organized into *modeling cycles* which move students systematically through all phases of model development, evaluation, and application in concrete situations. Throughout the course, students model physical situations with multiple representational tools including verbal descriptions, diagrams/motion maps, graphs, and equations.

The use of model-based physics and chemistry curriculum programs is mandated by the administration in the *modeling* school. For the non-honors class strand, students follow a "biology $\Longrightarrow$ modeling chemistry $\Longrightarrow$ modeling physics" sequence (or a non-PF approach). There are two groups of honors students: Academy or non-Academy. Within the honors class strand, the students enrolled in the Academy take a full year of physics covering topics in mechanics, electricity, waves, and optics. Honors students outside of the

Academy take a semester-long physics course emphasizing mechanics and following a block schedule. Only data from the non-Academy, honors group were included for this paper because they provide a closer match with the comparison school in terms of the course schedule and duration. In the modeling school, all honors students are required to take the modeling physics course in the ninth grade. Then, the students are channeled into Modeling Chemistry and Biology (BSCS Molecular Version) when they enter tenth grade. In grade 10, the students are highly encouraged to take AP Calculus and AP Physics B in grade 11. With coordinated supports from both the school administration and the parents, the twelfth graders are counseled to take AP Physics C or AP Biology or AP Chemistry in addition to AP Statistics. Since the inception of the reformed PF with modeling instruction programs in fall 2004, the student enrollment in the advanced placement science courses in the *modeling* school has dramatically increased from *21* students in 2005 to *123* students in 2007.

In the comparison school, students are required to take four science courses for graduation. All ninth graders take Earth and Space Science and tenth graders take Biology. Chemistry and physics are offered as electives at the eleventh and twelfth grades. Students may enroll in the introductory chemistry course with or without previously taking any physics course. The introductory physics courses involved in this study address topics in mechanics, waves, optics, and electricity. Students were also expected to conduct various experiments and develop skills in laboratory performance and reporting procedures.

Instruments

Force Concept Inventory

The FCI was designed to assess students' conceptual understanding of Newtonian mechanics (Hestenes et al. 1992/1995). The most recent version of the FCI consists of 30 multiple-choice items and focuses on the fundamental Newtonian concepts in kinematics, Newton's three laws, and types of forces, etc. The FCI has been used extensively in physics education research and is considered as a reliable and useful assessment tool to evaluate the effectiveness of instruction in introductory physics courses within and outside of the United States.

Instructional Activity Survey

To examine the classroom practices, IAS was created based on the key features of model-based approach (Wells et al. 1995), the Fundamental Abilities of Inquiry emphasized in the National Science Education Standards (NRC 1996), and instructional survey items released from the Trends in International Math and Science Study. At the end of each course, students rated how often they completed various activities in class, such as "Develop conceptual models using scientific evidence," and "Ask scientifically oriented questions." All items in the IAS use a four-point Likert-type scale (1 = Never or almost never; 2 = Sometimes; 3 = About half of the lessons; and 4 = Most of the lessons).

Reformed Teacher Observation Protocol

The RTOP, developed by the Evaluation Facilitation Group of the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT), is a 25-item observation instrument to assess the reform-based and inquiry-oriented classroom practices. According to ACEPT, the RTOP has been found to be reliable, valid, and a good predictor of student learning (Sawada et al. 2002). Using this protocol, the observer responds to twenty-five statements, clustered around three main categories: (1) *lesson design and implementation* (e.g., the teacher begins the lesson by acknowledging and respecting students' preconceptions and the students are engaged in exploration before formal presentation of concepts or definitions); (2) *content*, including both propositional and procedural knowledge, which emphasizes the development of coherent conceptual understanding, inquiry skills and metacognitive awareness; and (3) *classroom culture*, featuring decentralized communicative interactions and teacher/student relations that are more egalitarian with teachers supporting initiatives coming from students. The rater uses a response format that is based on a five-point Likert scale from 0 ("never occurred") to 4 ("very descriptive"). The sum of all the sub-scores produces an overall score (0–100), which, in turn, determines the degree for which the classroom practices observed compare with the statements about reform-based classroom practices. In addition to the statements that are rated by the observer, the RTOP also requires a narrative account of the lesson from the observer.

In this study, RTOP was used to examine the level of implementation of model-based science instruction in line with the reform-based and inquiry-oriented classroom practices. For instance, in the RTOP content section on propositional knowledge, Question 6 stated that "the lesson involved fundamental concepts of the subject." In the modeling classes, we examined whether basic physics and mathematics concepts were present and anchored in the broader ideas of the modeling unit.

Data Collection and Analysis

The data were collected during the spring and the fall semesters in two consecutive academic years. At the

beginning of the spring or the fall semester of 2007, the teachers in the *modeling* school started to implement a twenty-week physics course (ninth-grade honors strands and twelfth-grade non-honors strands), while the teachers in the comparison school taught their traditional, non-modeling, introductory physics courses to the eleventh- and twelfth-graders. The FCI conceptual test was administered at the beginning and end of the courses (as pre- and post-test assessments).

Rasch Scaling Analysis

To answer our research question on students' conceptual understanding of physics, the FCI data were first scaled using the Rasch scaling analysis, and then analyzed using analyses of covariance. The findings revealed that there was acceptable fit of the items along a single dimension, with 3 items showing acceptable fit (standardized Z fit scores lower than 2) and 27 items having a good fit (standardized Z fit scores lower than 1). Likewise, the students' ability measures all had Z fit scores within 2 standard deviations, and most lying within 1 standard deviation. This supports the use of the FCI as a single test, since the data demonstrate that students' responses fit a unidimensional measurement model.

The pretest data was scaled so that the Rasch measures had a range of points from 0 to 50 points, with a mean of 26.3 points and a standard deviation of 8.4 points. To allow accurate comparison of the students' scores across the pre- and posttests, the item loadings from the pre-test were also used for the post-test to anchor the measures. This resulted in a mean FCI posttest score of 42.3, with standard deviation of 7.0 points. All Rasch scaling was conducted using WINSTEPS (Linacre 2007).

Analyses of Covariance

Following the Rasch scaling, analyses of covariance were conducted. Differences in the schools' policies affected the distribution of students by grade level, treatment condition, and class strands of physics courses. Because of the unbalanced distribution in the sample, two analyses were conducted to compare: (A) Modeling with PF students and Comparison students (Honors strands); and (B) Modeling and Comparison groups of the students in non-PF (non-Honors strands). Table 1 summarizes the samples and variables included in the two analyses. All analyses of covariance were conducted using the *R* statistical environment (Ihaka and Gentleman 1996). Effect sizes were calculated as Hedges' (1981) *g*, which uses the same mean difference as Cohen's *d* but with a more conservative pooled standard deviation.

To examine the classroom practices, each student was asked to report the class activities by completing the Instructional Activity Survey, as described above. The responses from the IAS were examined to determine if there were significant differences in treatment and comparison group students' responses to the IAS items.

In addition, two modeling classes were each observed on three occasions: at the beginning, middle, and end of a selected modeling unit in the middle of the semester. Each unit generally required about 2–3 weeks to complete. Prior to the study, both the first and the third authors completed the training as suggested by the creators of the RTOP instrument and then observed two modeling classes together. Following the discussions on the use of the instrument in a model-based inquiry learning environment, the third author completed all classroom observations using the RTOP.

Results

Curriculum Impact on Students' Learning of Physics Concepts

As outlined in Table 1, students' conceptual learning measured by FCI scores is examined using two comparisons. The analyses were conducted separately because of imbalance in the data set. Table 2 presents the means and standard deviations for the Rasch-scaled pretest and

Table 1 Variables in each analysis of covariance

Comparison	Students in comparison	Dependent variable	Control and independent variables
A			
Modeling with PF vs. non-modeling with non-PF	Honors	Rasch-scaled FCI posttest	Rasch-scaled FCI pretest Modeling with PF
B			
Modeling with non-PF vs. non-modeling with non-PF	Non-honors, upper-class	Rasch-scaled FCI posttest	Rasch-scaled FCI pretest Modeling

For Comparison A, the modeling program was confounded with the Physics First (PF) program

Table 2 Descriptive statistics for Rasch-scaled FCI pre- and posttest, by comparison

Source	Pretest		Posttest	
	Mean	SD	Mean	SD
Comparison A (honors)				
Modeling with PF	24.80	7.37	47.50	6.57
Non-modeling with non-PF	28.69	9.18	41.95	6.57
Comparison B (non-honors)				
Modeling with non-PF	25.18	9.46	47.33	7.65
Non-modeling with non-PF	26.28	8.02	38.47	4.36

Table 3 Analyses of covariance: comparison A

Source	df	Sum sq	Mean sq	F value
FCI rasch pretest	1	560.2	560.2	15.836***
Treatment (modeling with PF)	1	1,415.1	1,415.1	40.004***
Error	123	4,351.0	35.4	

Dependent variable is Rasch-scaled FCI posttest score. *** $p < .001$; Effect size Hedges' $g = 2.45$

posttest scores on the FCI for the groups in each comparison. The results for each comparison are described below.

Comparison A

The purpose of Comparison A was to determine whether there was a difference between the modeling with PF students and the non-modeling with non-PF students (Honors strands). Recall that all Honors students in the modeling group were lowerclassmen (i.e., ninth-graders, in PF), while all Honors students in the comparison group were upperclassmen (i.e., 11th-/12th-graders, in non-PF). There were 62 students in the non-modeling group, and 64 in the modeling group. After controlling for the pretest score, there was a statistically significant difference between the modeling and non-modeling group students. This difference had an effect size of 2.45 (Hedges' g). (Table 3)

Comparison B

The purpose of Comparison B was to determine whether there was a difference between the modeling and the non-modeling students among the upperclassmen in the non-PF, non-Honors programs. There were 136 students in the non-modeling group, and 39 in the modeling group. There was no significant difference between the pretest scores ($p > .05$; not shown in table). After accounting for the pretest FCI score, there was a significant difference between the modeling and non-modeling groups, with modeling students scoring higher, on average (Table 4). This difference had an effect size of 2.62 (Hedges' g).

Table 4 Analyses of covariance: comparison B

Source	df	Sum sq	Mean sq	F value
FCI rasch pretest	1	382.7	382.7	15.339***
Modeling	1	2,494.7	2,494.7	99.981***
Error	172	4,291.7	25.0	

Dependent variable is Rasch-scaled FCI posttest score. *** $p < .001$; Effect size Hedges' $g = 2.62$

Implementation of the Modeling Approach in the Classrooms

Students' responses to the items related to inquiry and modeling instruction are presented in Table 5. As shown in Table 5, there were significant differences in how modeling and comparison group students reported the instruction that they experienced in the course.

It is apparent that the modeling group students reported higher scores on all of the statistically significant items with the exception of item 11 ("Do a science activity outside the classroom or conventional science laboratory."). No statistically significant differences were found between the two groups' responses on three items (IAS #1, 17, and 19). The treatment group differed from the comparison group students on items associated with inquiry, modeling, communicating, and reflecting aspects of classroom instruction. Further examination reveals that the following items demonstrated the most difference between the modeling and comparison groups (with effect sizes greater than 1): designing and conducting investigations; working together in small groups on experiments; writing explanations about what was observed and why it happened; writing about science in a report/paper on science topics; making a presentation to the class on the data, analysis, or interpretation; and critically reviewing other peers' work or presentations (Table 5).

In addition to the classroom activity surveys completed by students, two modeling teachers were observed three separate times, at the beginning, middle, and end of one unit or module during the semester. The RTOP scores for each of the classroom observations by teacher were presented in Table 6.

The teachers' total RTOP scores are between 77 and 95 out of 100, indicating that both teachers' classroom practices were generally in line with the model-based, inquiry-centered science instruction. There were teaching practices that were consistently observed across the classes: e.g., involving fundamental concepts of the subject in lessons, using a variety of means (models, drawings, graphs, manipulatives, etc.) to represent phenomena, and engaging students in groups to conduct inquiry. Some differences in teaching practices were also observed between the two

Table 5 Student ratings on selected instructional activities by treatment group

Item	Non-modeling		Modeling		Sig.	ES (d)
	M	SD	M	SD		
(1) Ask scientifically oriented questions	2.36	0.92	2.61	0.83		0.281
(2) Formulate our own hypotheses or predictions to be tested in an experiment or investigation	2.10	0.99	2.76	0.98	***	0.669
(3) Design and conduct experiments or investigations	2.01	0.85	3.04	1.08	***	1.102
(4) Write explanations about what was observed and why it happened	2.37	0.90	3.53	0.79	***	1.343
(5) Recognize and analyze alternative explanations by weighing evidence and examining reasons	2.22	0.91	2.76	0.96	***	0.582
(6) Develop conceptual models using scientific evidence	2.05	0.95	2.90	0.88	***	0.917
(7) Revise models and explanations based on evidence	2.12	1.02	3.07	0.85	***	0.984
(8) Write about science in a report/paper on science topics	2.24	0.91	3.16	0.80	***	1.053
(9) Make a presentation to the class on the data, analysis, or interpretation	1.70	0.87	2.97	1.15	***	1.303
(10) Critically review other peers' work or presentations	1.85	0.85	2.90	1.08	***	1.123
(11) Do a science activity outside the classroom or conventional science laboratory (for example, field trips or research)	1.54	0.87	1.18	0.51	**	−0.470
(12) Work together in small groups on experiments or investigations	2.99	0.92	3.89	0.40	***	1.150
(13) Work together in small groups to discuss our ideas	2.93	0.91	3.61	0.72	***	0.800
(14) Use multiple means (diagrams, graphs, symbols, equations, models, concrete materials, etc.) to represent the same situation	2.53	0.93	3.27	0.86	***	0.816
(15) Organize concepts (events or objects) in certain order and provide a reason for the organization	2.29	0.93	2.86	0.94	***	0.611
(16) Use mathematics to solve problems	3.71	0.68	3.94	0.29	**	0.398
(17) Study the impact of technology on society	2.46	0.91	2.39	0.96		−0.075
(18) Use computers, calculators or other educational technology for data collection, analysis, and presentation	3.28	0.86	3.67	0.63	**	0.494
(19) Relate what we are learning in science to our daily lives	2.87	0.94	2.85	0.93		−0.021
(20) Reflect on our own thinking and learning (e.g., how I know what I know; what are effective/ineffective learning strategies; etc.)	2.44	0.96	3.04	0.86	***	0.647

1 Never or almost never; *2* sometimes; *3* about half of the lessons; *4* most of the lessons. * $p < .05$; ** $p < .01$; *** $p < .001$

Table 6 Summary of mean and standard deviation for RTOP category scores for teachers during modeling instruction

RTOP category	RTOP scores by teacher*					
	Teacher A			Teacher B		
	Ob1	Ob2	Ob3	Ob1	Ob2	Ob3
Lesson design and implementation	15	12	14	19	18	19
Content						
Propositional knowledge	15	16	15	20	19	20
Procedural knowledge	18	15	17	19	18	19
Classroom culture						
Communicative interactions	15	16	16	17	16	17
Student teacher/relationships	15	18	16	20	19	20
Total score	78	77	78	95	90	95

*Each teacher was observed three times, at the beginning (Ob1), middle (Ob2), and end (Ob3) of a unit

teachers in teachers' knowledge, in lesson design and implementation (respect for students' prior knowledge, exploration preceding formal presentation), and in challenging of ideas and intellectual rigor. Overall, the greatest inconsistencies in teaching practices were found to be related to the lesson implementation and classroom interactions, such as the use of discourse, amount of student talk, and the degree to which the focus of the lesson was determined by student ideas and questions.

Discussion and Implications

This study is among the first to apply the Rasch modeling approach to the analysis of FCI and evaluation of the modeling physics program combined with PF on student conceptual learning in the United States. Our results indicate that the items on the FCI were unidimensional in their functioning. That is, the FCI measured one underlying

construct, and was not simply a collection of varying items on force. This finding is consistent with the results reported in a most recent study with Rasch analysis of FCI scores of a Croatian student sample (Planinic et al. 2010).

In the study by O'Brien and Thompson (2009), it was reported that the modeling physics instruction appeared to have a significant impact on the non-honors ninth-graders' conceptual understanding while no difference in gain scores between the two honors ninth-grade groups (modeling vs. traditional) were identified. The honors groups outperformed the non-honors groups, regardless of the type of instruction. In this study, however, we found that the model-based approach appeared much more effective than a traditional lecture-lab type instruction on student conceptual learning (with gain scores of about 10 Rasch-scaled points higher), regardless of the grade level (ninth grade vs. eleventh and twelfth grade) and the course strand (Honors and non-Honors). Although starting with somewhat lower pretest FCI scores, the Honors ninth-graders in the modeling classes achieved statistically significantly higher post-FCI scores than the honors eleventh- and twelfth-graders in the non-modeling classes did. For the non-Honors eleventh and twelfth grade classes, the modeling students also outperformed peers in the non-modeling group as measured by the FCI scores. Our results are consistent with previous findings about the general efficacy of modeling instruction (Wells et al. 1995).

According to the report of the 2006 Programme for International Student Assessment (PISA), sponsored by the Organisation for Economic Cooperation and Development (OECD 2007), in the United States, 92% of 15-year-olds are taking some kind of science course, whether compulsory or optional (compared to the OECD average of 87%). However, there were only 17% of students taking compulsory physics courses (OECD average 61%) and 11% optional courses (OECD average 15%). Physics is often perceived as a challenging subject and taken by select American students at the 11th or the 12th grade following biology and chemistry. Given the role of physics education as a foundation for advancement in all science disciplines and technology, having all 14- or15-year-old students enrolled in a PF curriculum would most likely result in "physics for all" and help achieve the goal of Scientific Literacy for All Americans as envisioned in the science education reform documents (AAAS 1989). Furthermore, since modeling has been identified as a unifying theme for both science and mathematics education (NRC 1996; NCTM 2000; Common Core State Standards Initiative 2010), a model-based physics curriculum combined with PF would provide an ideal context for students to learn and master algebra and the unifying modeling theme for the ninth graders. Further studies in this research line are highly recommended as such successful integration would

play a critical role in achieving both scientific and mathematical literacy for all high school students.

Our classroom activity survey and observation data indicate that the model-based, inquiry-oriented approach was implemented in all modeling classes. The IAS reports revealed statistically significant differences in inquiry, modeling, communicating, and reflecting aspects of instruction between the modeling and non-modeling classrooms. It appears that the following classroom teaching practices contributed most to the students' enhanced conceptual learning: working in small groups to design and conduct experiments or investigations; writing explanations about what was observed and why it happened; writing about science in a report/paper on science topics; making presentations to the class on their investigations; and critically reviewing other peers' work. A closer examination of the field notes and the RTOP profiles indicated that the greatest inconsistencies in model-based teaching practices observed were related to the classroom interactions such as the use of scaffolding and probing questions, amount of student talk, and the degree to which the lesson was built on or determined by student pre-conceptions, ideas and questions.

The emphasis on scientific inquiry in the standards documents requires learning and teaching science not only as "exploration and experiment" but also as "argument and explanation" (NRC 1996, p. 113). Such a requirement may be difficult for both teachers and students who have little prior experiences with scientific discourse. In our study, when using a model-based approach to teaching physics, scaffolded scientific discourse plays a critical role in the development of scientific understandings among students. Given the difficulties for some modeling teachers to guide students to conduct quality discourse in their model-based classes, on the one hand, we suggest that the development of teachers' expertise in guiding scientific discourse should be a focus for extended time periods beyond a weeks-long summer modeling institute for teachers. On the other hand, the curriculum developers and master modeling teachers might help create carefully-written scaffolds that systematically embed reasoning and argumentation in the modeling instructional materials, as a support for students to construct stronger arguments and improve the quality of classroom discourse.

Whereas the 2005 national survey found no evidence that students were more likely to enroll in advanced physics or other science courses after completing the ninth-grade physics at PF schools (Neuschatz et al. 2008), student enrollment in the advanced placement science courses had increased dramatically at the PF school in our study—from *21* students in 2005 at the beginning of the reform efforts to *123* students in 2007. In this particular PF school, in addition to the establishment of the model-based inquiry

science programs, the teachers and administrators had made extra efforts to set the academic stage by soliciting parental support and setting high expectations of students. We think that it is the combination of the academic program and the school cultural factors that may have led to such positive outcomes.

The present study had some limitations that indicate directions for future research. First, our findings are limited in scope as we only examined the effects of the model-based curriculum with PF on student conceptual learning of mechanics in this study. Since there was a coordinated mathematics component in the physics first program in our study, we need to find out whether the PF group score higher on a physics test involving more quantitative problem-solving items, even though prior research has indicated a positive correlation between students' FCI scores and their quantitative problem-solving performances (Hake 1998). Other assessment tools should also be used to further examine student learning outcomes closely related to model-based curriculum program, such as their understanding of the nature of scientific models and development, their scientific reasoning levels, and their ability to construct models in various contexts to solve problems or generate new questions for further scientific exploration. Second, the lack of randomization and inability to manipulate the independent variable are always the major weaknesses in any causal-comparative research. In this study, we made every effort to reduce the threats to the internal validity by matching the experimental and the comparison schools based on the students' demographics and the statewide, standardized reading and mathematics scores. While the FCI pre-test scores were used as a control in the Analyses of Covariance, the non-statistical differences in the FCI pretest scores further confirmed the comparability between the two groups of students. Yet, there was still a variety of factors outside the control of the study. For instance, one may argue that the modeling classes covered fewer physics topics than in the non-modeling classes involved in this study, therefore, the modeling students would have developed deeper understanding of the force concept due to the time spent and the depth and width of the contents covered in the course rather than the differences in instructional approaches. This might be true and it could also lead to different research questions such as: What content and skills should be emphasized in a ninth-grade level (or eleventh-/twelfth-grade level) introductory physics course from a learning progression perspective? In what depth and width? Nonetheless, physics education research at both secondary and college levels has repeatedly indicated that the conventional lecture-type instructional approach is not effective in promoting student conceptual understanding, which is independent of teachers and the number of years of physics education (e.g., Hake

1998; O'Brien and Thompson 2009; Wells et al. 1995). Our findings in this study are generally consistent with those reported in the existing research literature on modeling instruction.

Despite the above limitations, this study is significant in two ways: First, it used the Rasch modeling method in data analysis to further validate the widely used FCI research instrument and associated findings, whereas in most US-based physics education research literature involving concept tests, raw data (e.g., total scores, total percent scores, gain scores, normalized gains, etc.) were used and treated as if they were interval-level data. Second, the current study examines the effect of a model-based program combined with the increasingly popular Physics First approach using carefully-matched samples supported by both student concept scores and curriculum implementation data. The results warrant further research by means of experimental designs. More research on the sequence and coordination of science and mathematics courses should be conducted to inform both policies and practices in science education, particularly in the United States.

References

American Association for the Advancement of Science (AAAS) (1989) Science for all Americans. Oxford University Press, New York

Brewe E, Sawtelle V, Kramer LH, O'Brien GE, Rodriguez I, Pamelá P (2010) Toward equity through participation in Modeling Instruction in introductory university physics. Phys Rev Spec Top Phys Educ Res 6: 010106. Retrieved August 16, 2010, from http://prst-per.aps.org/pdf/PRSTPER/v6/i1/e010106

Clement J (1989) Learning via model construction and criticism. In: Glover G, Ronning R, Reynolds C (eds) Handbook of creativity: assessment, theory and research. Plenum, New York, pp 341–381

Clement J (2008) Six levels of organization for curriculum design and teaching. In: Clement J, Rea-Ramirez MA (eds) Model based learning and instruction in science. Springer, New York, pp 255–272

Common Core State Standards Initiative (2010) Common core state standards for mathematics. Retrieved August 10, 2010, from http://corestandards.org/the-standards

Eisenkraft A (2010) Active physics, 3rd edn. It's About Time, Armonk

Ewald G, Hickman J, Hickman P, Myers F (2005) Physics first: the right-side-up science sequence. Phys Teach 43:319–320

Gay L, Mills G, Airasian P (2006) Educational research: competencies for analysis and application, 8th edn. Prentice-Hall, Upper Saddle River

Giere RN (1988) Explaining science: a cognitive approach. University of Chicago Press, Chicago, IL

Goodman R, Etkina E (2008) Squaring the circle: a mathematically rigorous physics first. Phys Teach 46(4):222–227

Hake RR (1992) Socratic pedagogy in the introductory physics laboratory. Phys Teach 30(12):546–552

Hake RR (1998) Interactive engagement vs. traditional methods: a six-thousand student survey of mechanics test data for introductory physics courses. Am J Phys 66(1):64–74

Halloun IA (2004) Modeling theory in science education. Kluwer Academic Publishers, Dordrecht

Hamilton R, Ghatala E (1994) Learning and instruction. McGraw-Hill, New York

Hedges LV (1981) Distribution theory for Glass's estimator of effect size and related estimators. J Educ Stat 6(2):107–128

Hestenes D (1987) Toward a modeling theory of physics instruction. Am J Phys 55(5):440–454

Hestenes D, Wells M, Swackhamer G (1992/1995) Force concept inventory. Phys Teach 30:141–158

Ihaka R, Gentleman R (1996) R: a language for data analysis and graphics. J Comput Graph Stat 5:299–314

Karplus R (1977) Science teaching and the development of reasoning. J Res Sci Teach 14:169–175

Kuhn TS (1970) The structure of scientific revolutions. University of Chicago Press, Chicago

Lederman L (2001) Revolution in science education: put physics first!. Phys Today 54(9):11–12

Lesh R, Doerr HM (2003) Beyond constructivism. Math Think Learn 5(2&3):211–233

Linacre JM (2007) Winsteps (Version 3.61.2) [Computer Software]. Winsteps.com, Chicago

Liu X, Boone WJ (2006) Introduction to Rasch measurement in science education. In: Liu X, Boone WJ (eds) Applications of Rasch measurement in science education. JAM Press, Publisher, MN, pp 1–22

Malone K (2008) Correlation between knowledge structures, force concept inventory, and problem-solving behaviors. Phys Rev Spec Top Phys Educ Res. Available from http://prst-per.aps.org/abstract/PRSTPER/v4/i2/e020107

Mountz DK (2006) The effect of science core sequence reform on students' attitudes toward science. Doctoral dissertation, Immaculata University, Malvern, PA. Retrieved October 11, 2006, from http://modeling.asu.edu/Projects-Resources.html

National Council of Teachers of Mathematics (2000) Principles and standards for school mathematics, Reston, VA. Retrieved August 12, 2010 from: http://standards.nctm.org

National Research Council (1996) National science education standards. National Academy Press, Washington

National Research Council (2000) Inquiry and the national science education standards: a guide for teaching and learning. National Academy Press, Washington

National Research Council (2007) Taking science to school. The National Academies Press, Washington

Neuschatz M, McFarling M, White S (2008) Reaching the critical mass: the twenty year surge in high school physics. Findings from the 2005 nationwide survey of high school physics teachers. AIP Report number R-442, ED502250. Retrieved from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/3e/8a/eb.pdf

O'Brien MJ, Thompson JR (2009) Effectiveness of ninth-grade physics in maine: conceptual understanding. Phys Teach 47(4):234–239

OECD (2007) PISA 2006: science competencies for tomorrow's world. Retrieved July 8, 2008, from http://www.pisa.oecd.org/document/2/0,3343,en_32252351_32236191_39718850_1_1_1_1,00.html

Pasero S, Fermilab Education Office (2003) The state of physics-first programs: a report for Project ARISE. FERMILAB-Pub-01/206

Piaget J (1970) Piaget's theory. In: Mussen PH (ed) Carmichael's manual of child psychology, 3rd edn. Wiley, New York, pp 703–732

Planinic M, Ivanjek L, Susac A (2010) Rasch model based analysis of the Force Concept Inventory. Phys Rev Spec Top Phys Educ Res 6–010103:1–11

Posner G, Strike K, Hewson P, Gertzog W (1982) Accommodation of a scientific conception: toward a theory of conceptual change. Sci Educ 66:211–227

Sawada D, Piburn MD, Judson E, Falconer K, Turley J, Falconer K, Benford R, Bloom I (2002) Measuring reform practices in math and science classrooms: the reformed teaching observation protocol. Sch Sci Math 102(6):245–254

Schuchardt A, Malone K, Diehl W, Harless K, McGinnis R, Parr TD (2008) A case study of student performance following a switch to a modeling-based physics first course sequence. Paper presented at the National Association for Research in Science Teaching annual conference, Baltimore, MD

Schwarz C, White B (2005) Meta-modeling knowledge: developing students' understanding of scientific modeling. Cognit Instr 23(2):165–205

Schwarz CV, Reiser BJ, Davis EA, Kenyon L, Acher A, Fortus D, Shwartz Y, Hug B, Krajcik J (2009) Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. J Res Sci Teach 46(6):632–654

Thagard P (1992) Conceptual revolution. Princeton University Press, Princeton

Vesenka J, Beach P, Munoz G, Judd F, Key R (2002) A comparison between traditional and "modeling" approaches to undergraduate physics instruction at two universities with implications for improving physics teacher preparation. J Phys Teach Educ Online 1(1): 3–7. Retrieved March 2, 2007, from http://phy.ilstu.edu:16080/jpteo/issues/june2002.html

Vygotsky LS (1978) Mind in society: the development of higher psychological processes. Harvard University Press, Cambridge, MA

Wells M, Hestenes D, Swackhamer G (1995) A modeling method for high school physics instruction. Am J Phys 63:606–619

White BY, Frederiksen JR (2000) Metacognitive facilitation: an approach to making scientific inquiry accessible to all. In: Minstrell J, van Zee EH (eds) Inquiring into inquiry learning and teaching in science. American Association for the Advancement of Science, Washington, DC, pp 331–370