5-5-2007

A cross-cultural, multilevel study of inquiry-based instruction effects on conceptual understanding and motivation in physics

Meiko Negishi

A CROSS-CULTURAL, MULTILEVEL STUDY OF INQUIRY-BASED INSTRUCTION EFFECTS ON CONCEPTUAL UNDERSTANDING AND MOTIVATION IN PHYSICS

By

Meiko Negishi

A Dissertation
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in Educational Psychology
in the Department of Counselor Education and Educational Psychology

Mississippi State, Mississippi

May 2007

A CROSS-CULTURAL, MULTILEVEL STUDY OF INQUIRY-BASED

INSTRUCTION EFFECTS ON CONCEPTUAL UNDERSTANDING

AND MOTIVATION IN PHYSICS

By

Meiko Negishi

Approved:

Anastasia D. Elder
Assistant Professor of Educational Psychology
(Director of Dissertation)

David T. Morse
Professor of Educational Psychology
(Committee Member)

Linda W. Morse
Professor of Educational Psychology
(Committee Member)

Vincent R. McGrath
Professor of Education
(Committee Member)

Jianzhong Xu
Associate Professor of Education
(Committee Member)

Glen R. Hendren
Professor of Counselor Education
Graduate Coordinator

Richard Blackbourn
Dean of the College of Education

Name: Meiko Negishi

Date of Degree: May 5, 2007

Institution: Mississippi State University

Major Field: Educational Psychology

Major Professor: Dr. Anastasia D. Elder

Title of Study: A CROSS-CULTURAL, MULTILEVEL STUDY OF INQUIRY-BASED INSTRUCTION EFFECTS ON CONCEPTUAL UNDERSTANDING AND MOTIVATION IN PHYSICS

Pages in Study: 103

Candidate for Degree of Doctor of Philosophy in Educational Psychology

Student achievement and motivation to learn physics is highly valued in many industrialized countries including the United States and Japan. Science education curricula in these countries emphasize the importance and encourage classroom teachers to use an inquiry approach. This dissertation investigated high school students' motivational orientations and their understanding of physics concepts in a context of inquiry-based instruction. The goals were to explore the patterns of instructional effects on motivation and learning in each country and to examine cultural differences and similarities.

Participants consisted of 108 students (55 females, 53 males) and 9 physics teachers in the United States and 616 students (203 females and 413 males) and 11 physics teachers in Japan. Students were administered (a) Force Concept Inventory measuring physics conceptual understanding and (b) Attitudes about Science

Questionnaire measuring student motivational orientations. Teachers were given a survey regarding their use of inquiry teaching practices and background information. Additionally, three teachers in each country were interviewed and observed in their classrooms.

For the data analysis, two-level hierarchical linear modeling (HLM) methods were used to examine individual student differences (i.e., learning, motivation, and gender) within each classroom (i.e., inquiry-based teaching, teaching experience, and class size) in the U.S. and Japan, separately. Descriptive statistical analyses were also conducted.

The results indicated that there was a cultural similarity in that current teaching practices had minimal influence on conceptual understanding as well as motivation of high school students between the U.S. and Japan. In contrast, cultural differences were observed in classroom structures and instructional approaches. Furthermore, this study revealed gender inequity in Japanese students' conceptual understanding and self-efficacy.

Limitations of the study, as well as implications for high school physics teachers are discussed. Future research in this line could explore students' use of cognitive strategies to overcome misconceptions in Western and Eastern cultures. Also, exploring the best practices in changing student misconceptions and promoting motivation across cultures would enrich our understanding and current teaching practices.

DEDICATION

For my parents who always provided me understanding, encouragement, and love.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the many people who assisted and supported me in writing this dissertation. First of all, sincere thanks are due to Dr. Anastasia Elder, my committee chair, for her thoughtful guidance and warm support throughout the intricacies of the doctoral program and the dissertation process. Expressed appreciation is also due to the other members of my dissertation committee, including Dr. David Morse, Dr. Linda Morse, Dr. Vincent McGrath, and Dr. Jianzhong Xu, for the invaluable aid and direction provided by them. I also would like to express my gratitude to the high school students, physics teachers, and school administrators in the U.S. and Japan, who supported my study by dedicating their time and efforts. Finally, I would like to thank my family and friends who always supported and believed in me.

TABLE OF CONTENTS

APPENDIX Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

> Cross-cultural comparison is a powerful way to unveil unnoticed but ubiquitous practices
> (Stigler, Gallimore, & Hiebert, 2000, p. 88)

In many industrialized countries, people highly value mathematics and science because the realm of science promotes industrial productivity. These values are reflected in the social and cultural expectations and educational standards. Therefore, it has been one of the major goals of the technologically advanced societies such as the United States and Japan to enhance student achievement in mathematics and science. People are interested in knowing other nations' education by comparing student learning.

Since 1995, a large scale cross-national study on student achievement in mathematics and science has been conducted on a four-year cycle by the International Association for the Evaluation of Educational Achievement (IEA). This study, the Trends in International Mathematics and Science Study (TIMSS), has not only measured student achievement but also students' attitudes toward the subjects and instructional practices.

International studies including TIMSS are helpful to understand the trends in student achievement, students' attitudes toward the subjects, and instructional practices. However, what is more valuable about these studies is interpreting the trends from cross-cultural perspectives. So far, TIMSS has focused on 4^{th} and 8^{th} grade students and their teachers. However, a growing concern in industrialized countries centers on students' competence in advanced mathematics and physics, and there is few cross-cultural study examining student achievement, students' attitudes, and instructional practices in physics and advanced mathematics.

Therefore, this current study investigated high school students' learning, motivation, and instructional practices in physics in the United States and Japan. It has been more than a decade since the National Research Council (NRC, 1996) endorsed inquiry approach in teaching science to enhance student learning and motivation. In Japan also, since 1990s observations and experiments in science curriculum have been emphasized in their Gakushu Shido Youryou (1998) national curriculum standards. It would be informative for teachers, administrators, and policy makers to know the effects of inquiry-based teaching practices on high school students' achievement and motivation in physics.

In the following section, previous studies relating to the current research will be reviewed. The first part will focus on what difficulties students face and how they construct conceptual understandings by highlighting the issues in learning in physics. The second part will examine student motivation such as goals and self-efficacy and their relations to student learning in science. The third part will explore science instruction that enhances students' motivation and promotes their learning. Finally, the fourth part will examine cultural issues related to learning, motivation, and instruction.

Misconceptions in Physics

For most students, physics is a challenging subject. The factors contributing to its difficulty are abstractness of the material and the advanced reasoning and mathematical skills required. However, researchers have pointed out that students' misconceptions in physics are additional contributing factors (Caramazza, McCloskey, & Green, 1981; Champagne, Klopfer, & Anderson, 1980; Clement, 1982; Halloun & Hestenes, 1985). Every student begins a physics course with his or her own belief system about how the physical world works derived from personal experience. Therefore, students tend to interpret what they see and hear in a physics course using their own beliefs based on their prior experience (Halloun & Hestenes, 1985). For instance, students often believe that a heavy-weight object falls faster than light-weight object even in a frictionless world. These existing beliefs based on their experience lead to students' misconceptions. Often, these misconceptions are highly stable and difficult to change.

Among the most studied area of students' misconceptions in physics are force and motion. Champagne, Klopfer, and Anderson (1980), among the pioneers finding the existence of students' misconceptions, examined students' beliefs about motion of objects during free fall. A group of 110 students in an introductory college physics course was included in this study. Of 110, four students had two or more years of high school physics, 68 students had one year of high school physics, and 32 students never took a physics course. The participants were asked to observe the motion of an object during free fall. They were then required to describe the observations, answer questions, and provide justifications for their answers. The results indicated that approximately four in five participants believed that heavier objects fall faster than lighter ones, when everything else was held constant (i.e., size and shape). The surprising realization to the researchers was not that students did not know the major concept of Newtonian mechanics. Rather, it was that each student had commonsense, intuitive ideas, which were competing with the paradigm of Newtonian mechanics.

Caramazza, McCloskey, and Green (1981) also found similar results. They examined students' beliefs about trajectories of objects, and 44 undergraduate students were included in the study. Of 44, ten had completed at least one college physics, 20 had taken high school physics, and 14 never had formal physics instruction. The students were presented with figures of a problem in which a ball was assumed to be moving in an arc. They were then asked to draw the path the ball would follow if the string were cut when the ball was at four different locations. The results indicated that 75% of the students had misconceptions about projectile motion. The students failed to consider initial velocity of the ball as well as the action of gravity or either one of them. The results also showed that the students were systematically applying incorrect beliefs about the path of moving objects. This study indicated students' misconceptions of projectile motion are resistant to change because 67% of the students who had completed a physics course at high school and/or college level still held incorrect belief systems.

In a similar study conducted by Clement (1982), written tests and videotaped problem-solving interviews were used with college students in introductory mechanics course. He administered three types of problems including a pendulum problem, a coin problem, and a rocket problem. For example, in the coin problem, students see a figure indicating that coin was tossed. They were then asked to draw and label arrows showing the direction of each force acting on the coin. The participants were 34 engineering freshman who mostly had had high school physics. Clement found that 88% of the students answered incorrectly. All of the errors they had were showing an arrow labeled as a force pointing upwards. Clement found pervasive misconceptions on all three types of problems, and he coined it as “motion implies a force” misconception, which meant continuing motion implies the presence of a force in the same direction as the motion. He further investigated students’ misconceptions after the instruction. The participants were another group of 43 students in a mechanics course from the same institution as the freshman group reported earlier. The results indicated that 75% of the students held the misconception on the coin problem. Clement pointed out that “motion implies a force” misconception was highly stable even after completed the course. This study played an important role for successive researchers in the field of students’ misconceptions in physics.

Halloun and Hestenes (1985) also investigated misconceptions including not only university students but also high school students. They extended previous studies by synthesizing isolated concepts being studied. Further, they published an instrument called the Mechanic Diagnostic Test (MDT). The test questions were initially aimed to assess students’ qualitative knowledge of Newtonian mechanics, and to identify common misconceptions which had been reported by previous researchers. Various versions of the test were given over a period of three years to more than 1000 college students in introductory physics courses. From these qualitative, written answers from students, Halloun and Hestenes developed a questionnaire with 36 multiple-choice items as the final version of the MDT. They conducted face validity and content validity by using experts such as physics professors and graduate students. They also administered a pilot study with students in introductory physics courses. They further reported the internal consistency reliability based on the Kuder-Richardson test. The range of the coefficient was between 0.86 and 0.89, which indicated that the MDT was highly reliable instrument. The researchers administered the MDT to high school students (24 honors, 25 general) in beginning physics classes. The results indicated that students, regardless of their academic levels, scored extremely low on pretest, with an average of 11 out of 36 (30%) correct. Moreover, at the end of the course their posttest scores were still less than 20 (56%) correct. From this study, the researchers concluded that misconceptions about force and motion were firmly in place. Therefore, it is important for physics teachers to be aware of the common misconceptions and also work on changing them.

Hestenes, Wells, and Swackhamer (1992) continued examining students' misconceptions about force and motion. They improved the MDT by supplying a more systematic and complete profile of the various misconceptions, and called it the "Force Concept Inventory (FCI)." The multiple-choice format of the inventory remained the same, but it had fewer items (n = 30). The Inventory aimed to find students' belief systems. The researchers cautioned that non-Newtonian concepts, commonly labeled as misconceptions should be regarded as reasonable hypotheses grounded from everyday experience. The FCI has been administered to more than 1500 high school students including regular, honor, and AP classes. This study also found similar finding to the one from the study by Halloun and Hestenes (1985). Regardless of the class levels, the students' scores were low; ranging of 48% to 57% on the posttest of FCI.

In summary, previous studies found that most students in introductory physics courses exhibited misconceptions. Further, the studies indicated that these misconceptions were stable and difficult to change by conventional instruction that does not take them into account. The challenges in designing instruction to modify student misconceptions should be examined not only from cognition but also student motivation that has been considered an important aspect of learning. Pintrich, Marx, and Boyle (1993) argued that cognitive factors of learning (i.e., learning strategies) do not fully explain students' actual cognitive engagement in classroom academic tasks. In fact, students' motivational factors contribute to their engagement in classroom tasks. Therefore, in the following section, theories and empirical studies investigating student motivation and its relationship to learning are reviewed.

Motivation Theories

Among many existing motivation models related to student academic performance, the contemporary cognitive view on motivation assumes that achievement motivation derives from an individual's conscious beliefs and values that are influenced by recent experiences and consequences such as success or failure (Stipek, 1988). Based on these experiences, students do develop beliefs about their reasons for choosing a task and about their capability to perform a task (Pintrich, Marx, & Boyle, 1993). Therefore, motivation theories including goal and self-efficacy theories are discussed and relevant studies are reviewed next.

Goal Theory

In goal theory, individuals are motivated to learn for particular goals. Some students want to learn the course material, while others are interested in pleasing their parents. All students are motivated either internal reasons (intrinsic motivation) or external reasons (external motivation). Previously, motivational theorists investigated possible constructs of students' goals for achievement. Maehr (1983) attempted to define the meaning of achievement using a questionnaire to assess individual student's motivational orientation. Using participants from more than 30 cultural-linguistic groups, Maher found that an individual's meaning of achievement involves one's projected goals in performing. These goals are defined as either intrinsic or extrinsic motivation. Within intrinsic goals there are task-oriented goals and ego-oriented goals. Students with task-oriented goals focus on understanding and experiencing novelty. In contrast, those with ego-oriented goals seek to outperform others and to look smarter than their peers. Within extrinsic goals there are social solidarity goals and rewards goals. Students with social solidarity goals aim at pleasing teachers and/or parents whereas students with rewards goals seek to get good grades.

Harter (1981) also aimed to develop a self-report scale that identifies a student's intrinsic versus extrinsic motivational orientation. Data was collected from over 3,000 third through ninth grade students. As a result, Harter defined five components, each defined by an intrinsic and an extrinsic pole: (a) preference for challenge versus preference for easy work, (b) curiosity/interest versus pleasing teacher, (c) independent mastery versus dependence on teacher, (d) independent judgment versus reliance on teacher's judgment, and (e) internal criteria versus external criteria.

Pintrich and DeGroot (1990) measured students' achievement-related motivational beliefs using a self-report questionnaire, the Motivated Strategies for Learning Questionnaire (MSLQ). The MSLQ used in this study included nine items on intrinsic value and nine items on self-efficacy. Within intrinsic value, there were task value and mastery goals. Task value refers to students' perception of the importance of course work (e.g., "it is important for me to learn what is being taught in this class"). Mastery goals refer to students' focus on understanding and challenging (e.g., "I prefer class work that is challenging so I can learn new things"). The original MSLQ includes an additional goal category, extrinsic goals. Within extrinsic goals there are rewards goals and ability goals.

Table 1 displays the summary of the constructs for goals. Harter (1981) classified intrinsic (i.e., challenging task, interest, active, independent judgment, and internal criteria) versus extrinsic (i.e., easy task, pleasing others, mastery, dependent judgment, and external criteria). Maehr (1983) categorized intrinsic (i.e., mastery-oriented and ego-oriented goals) versus extrinsic (i.e., social solidarity and rewards). Pintrich and DeGroot (1990) defined intrinsic (i.e., mastery and task value) and extrinsic (i.e., rewards and ability).

Table 1

Goal Constructs

Construct		Description	Researchers
Intrinsic	Understanding /Challenge *(Mastery Goals)*	Understanding something	Harter (1981) Maher (1983) Pintrich et al. (1990)
	Task Value (*Task Value)*	Importance of the task; course work	Pintrich et al. (1990)
	Ego	Besting others	Maher (1983)
Extrinsic	Rewards *(Extrinsic)*	Getting good grades	Maher (1983) Pintrich et al. (1990)
	Ability *(Extrinsic)*	Evaluation by others; competition	Pintrich et al. (1990)
	Social Solidarity	Pleasing others	Harter (1981) Maher (1983)

Note: MSLQ categories are included in italics for comparison purposes

Barlia (1999) used the MSLQ to examine the relationship between high school students' motivational orientations and their learning. The findings indicated that students' task value was significantly related to student learning. A study by Pintrich and DeGroot (1990) found that junior high school students' intrinsic value (i.e., task value and mastery goals) was positively correlated with their use of cognitive strategies in science. Other studies also confirmed the relationship between students' mastery goals and achievement (Pintrich, 2000; Tuan, Chin, & Shieh, 2005; Wolter, 2004). Therefore, from these previous studies, students high in task value and mastery goals are more likely to perform better than those low in motivation.

Self-Efficacy Theory

An individual student is also motivated to achieve based on one's perception of efficacy. Self-efficacy is defined as individual's personal judgment about one's capability of performance (Bandura, 1982). It is viewed as relatively situation-specific, not as a global personality trait. According to Bandura (1982), there are four sources of self-efficacy development: (a) performance accomplishments, (b) vicarious experience, (c) verbal persuasion, and (d) emotional arousal. He explained that some people have strong self-efficacy due to their successes, while others have low self-efficacy due to their failures. Also, people develop self-efficacy by observing similar others perform successfully. Further, people's self perception of efficacy is enhanced by verbal persuasion from others. Lastly, people gain self-efficacy when they were able to manage stressful situations and succeed.

Another important aspect of self-efficacy is that it mediates the relationship between knowledge and action (Bandura, 1982). Individuals with high self-efficacy are more likely to engage themselves in challenging tasks than those with low self-efficacy are. Moreover, people who have high sense of self-efficacy tend to choose and persist in the task, which would lead to greater performance.

In a study by Pintrich and DeGroot (1990), junior high school students' self-efficacy was measured by the MSLQ. The scale consisted of nine items regarding students' perceptions of competence and confidence in performance of class work (e.g., I am sure I can do an excellent job on the problems and tasks assigned for this class). The results indicated that self-efficacy was significantly related with cognitive strategies use. Students with high self-efficacy were more likely to use cognitive strategies such as rehearsal, elaboration, and organization than students with low self-efficacy. Another study by Bandura, Barbaranelli, Caprara, and Pastorelli (1996) examined the impact of students' self-efficacy beliefs on academic performance. The participants were 279 students ranging in age from 11 to 14 and they were measured on beliefs in their capabilities to master coursework in science. The results from path analysis revealed that students' academic self-efficacy beliefs were significantly linked to their academic achievement.

In summary, current research on student learning emphasizes importance of students' motivation and its relation to their cognition. It is recommended that the learning environment should consider both motivational and cognitive aspects. Physics instruction is not an exception. Therefore, instructional design needs to focus on enhancing students' sense of self-efficacy and promoting students' internal goal motivation for better learning. One of the promising approaches recommended by researchers to enhance students' motivation, overcome their misconceptions, and develop successful understanding is inquiry-based instruction (Erylimaz, 2002; Gibson & Chase, 2002; Stamp & O'Brien, 2005). Inquiry-based instruction such as questioning, experimentation, explanations, and discussion allows students to compare existing misconception and rival concepts (Sinatra & Pintrich, 2003). In the following section, how inquiry is conceptualized and how inquiry-based instruction has been implemented in classrooms practices are discussed.

Inquiry-Based Instruction

The National Science Education Standards by the National Research Council (NRC; 1996) define learning as an active process and learning science is something that students do, not something that is done to them. They also state that teaching should move away from mere presenting information and covering topics. Rather, science teaching must engage students in activities in which they ask questions and investigate phenomena. These are the essential ideas of inquiry-based instruction and learning. The definitions are given in the Standards (NRC, p.23).

> "Inquiry refers to the activities of students in which they develop knowledge and understanding of scientific ideas, as well as an understanding of how scientists study the natural world"
> "Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other resources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations and predictions; and communicating the results"

Therefore, inquiry approach is recommended to be a vital part of science instruction.

Inquiry-Based Instruction and Conceptual Understanding

Research has examined the effects of inquiry-based approach in science education, and has found it to be highly effective in improving students' conceptual understanding measured by achievement tests in science (Ertepinar & Geban, 1996; Von Secker, 2002) and reasoning ability (Gerber, Cavallo, & Marek, 2001).

Ertepinar and Geban (1996) explored the effectiveness of inquiry-based laboratory activities on student achievement in science. A total of 43 students (23 in control, 20 in experimental) from general science course in 8^{th} grade were involved in a five-week study. The control group was exposed to worksheet exercises as a supplement to classroom instruction. The worksheets included conceptual and mathematical problems, which required written responses. Upon the completion of the worksheets, the teacher collected, graded, and provided feedback. The experimental group, on the other hand, was exposed to inquiry-based laboratory activities as a supplement. In the laboratory activities, students proposed their own hypotheses individually, and presented a procedure for solving given problems. They were also required to design and carry out the experiments on their own with the materials and equipment provided, guided by the teacher if needed. Students gathered and interpreted data; and they were asked to draw conclusions and make generalizations. Both groups were administered pretest and posttest using an 18-item multiple choice test measuring student science achievement, which yielded an estimated KR reliability coefficient of .87. The experimental group students scored statistically significantly higher than the control group did on the science achievement test. Therefore, this study indicated that inquiry-based laboratory activities were effective for enhance students' conceptual understanding.

Another study by Von Secker (2002) also examined the effects of inquiry-based teaching practices on student science achievement using data from the National

Education Longitudinal Study (NELS). The sample consisted of 4,377 students in 1,406 classes who enrolled in 10th grade. The researcher analyzed student data such as science achievement scores, gender, race-ethnicity, and socioeconomic status. Student achievement was measured by a standardized test developed by the Educational Testing Service to assess understanding of fundamental concepts, mastery of basic skills, and higher order thinking skills. Teacher practices of inquiry-based approach were measured by a teacher survey asking how much emphasis they placed on (a) eliciting student interest and engagement, (b) using appropriate laboratory techniques, (c) problem solving, (d) conducting further study, and (e) scientific writing. Data analysis using hierarchical linear modeling method revealed that the inquiry-based teaching practices positively affected student achievement regardless of their background characteristics (gender, race-ethnicity, and socioeconomic status). However, the researcher pointed out that teacher practices had little impact in reducing academic inequity, especially gaps between majority and minority female students.

Gerber, Cavallo, and Marek (2001) explored the relationship between inquiry-oriented teaching and student scientific reasoning ability. The researchers suggested that an inquiry approach provides students direct experiences that promote cognitive conflict and hence encourage learners to develop new concepts. They also emphasized that the role of the teacher is to provide physical experiences and encourage student social interaction and reflection. In this study, the participants were 505 students enrolled in 7th through 10th grade science classes, and 16 science teachers. Students were administered a 12-item written test to assess students' ability to conserve weight and volume, separate variables and use propositional logic, combinatorial reasoning, and correlations. Teachers were identified either as inquiry or non-inquiry science teachers based on their participation in an inquiry-based science teaching methods course, and descriptions of teaching practices. Inquiry teachers emphasized material-rich, student-discovery activities. They also frequently asked students to question, formulate explanations of phenomena, work in groups, discuss results and conclusions, and present findings. On the other hand, non-inquiry teachers heavily relied on giving lectures, doing book reports, completing worksheets, watching videos, and doing verification laboratories. The results of the study found that students in science class taught by inquiry-based instruction scored statistically significantly higher on the scientific reasoning test compared to those in non-inquiry teaching classes. Therefore, this study informed the importance of inquiry-oriented classroom teaching practices to scientific reasoning abilities.

Inquiry-Based Instruction and Motivation

Pintrich, Marx, and Boyle (1993) argued that besides the links between motivation and cognition, the instructional characteristics of the actual classroom context may impact students' motivation and cognition. Studies have investigated the effectiveness of inquiry-based instruction in enhancing students' attitudes toward science and science learning.

Canton, Brewer, and Brown (2000) examined the impact of inquiry methods on students' attitudes toward science. Four high school teachers participated in a three-day institute that focused on wind-energy through collaboration with eight scientists. During the institute, teachers worked together to build working windmills using kits of everyday supplies and equipment. They were asked to design experiments and test their hypotheses. Participants had the opportunity to ask questions of scientists and to bring up conceptual issues into discussions. Teachers were also involved in discussions of the challenges and effective teaching strategies regarding inquiry methods. Before and after completing the institute, the teachers administered to their 230 high school students a survey regarding attitudes toward inquiry activities. The survey measured five elements of the classroom environments including student satisfaction with the class, class cohesiveness, friction among classmates, difficulty of work, and classmate competitiveness. The results found that student satisfaction with the class was statistically significantly increased after the inquiry-based methods were introduced. Also, friction among classmates was statistically significantly reduced after the inquiry-based instruction.

Another study by Tretter and Jones (2003) examined the impact of inquiry-based instructions on student participation, classroom grades, and scores on a standardized test. This was a case study of physical science classes taught by a high school teacher over four years. The first two years of classes were taught using traditional instruction with a relatively low level of inquiry-based teaching, and the last two years were taught by inquiry methods. One hundred and sixty-four non-inquiry group students followed the lab work described in the textbook without necessarily demonstrating an understanding of the underlying physical concepts. On the other hand, 94 inquiry group students were asked to devise their own investigations and to develop their own procedures to understand and explain underlying concepts. Three variables pertaining to the student (participation, grades, and scores) were examined and compared across these two instructional groups. Student participation was measured by (a) the percentage of students who took the end-of-course test, (b) the mean absence rate, and (c) the percentage of students who did not complete the course. The results found that the inquiry group students were more likely to take the end-of-course test, had higher attendance, and were less likely to give up the course. The mean grade of the inquiry group was a statistically significantly higher than that of the non-inquiry group. However, the mean standardized test score of the inquiry group was not statistically significantly different from that of the non-inquiry group. This study indicated that an inquiry-based teaching style was effective to develop student positive attitudes toward the subject matter and student understanding measured by course grades. Also, the study suggested that a standardized achievement test may be not compatible with inquiry-based instructional goals.

Gibson and Chase (2002) explored the long-term impact of inquiry-based instruction on students' attitudes and interest about science. They conducted the Summer Science Exploration Program (SSEP), a 2-week inquiry-based science camp, which intended to stimulate greater interest in science and scientific careers among middle-school students. Over a five-year time span, two surveys measuring attitudes toward science and career interests were administered to 79 SSEP participants and 35 classmates of SSEP participants as a control group. The SSEP provided participants with the opportunity to examine different biological and health related topics through inquiry-based activities. Students learned how to formulate their own questions that can be addressed through observation. They designed experiments and practiced laboratory and field techniques. Students also analyzed data and discussed the results and conclusions with their classmates. Additionally, college faculty and middle-school teachers were working collaboratively to create an exciting and comfortable atmosphere for students. The results found that both SSEP participants and control group students' attitudes toward science and interest in science careers decreased as they went from middle to high school. However, only the control group showed a statistically significant decrease on both surveys. This finding indicated that over the years, SSEP students maintained a more positive attitude toward science and a higher interest in science careers than control group students. Therefore, this study confirmed the long-term effects of inquiry-based activities on students' attitude and interest about science.

Patrick and Yoon (2004) examined 4 eighth grade (2 females and 2 males) students' motivational beliefs and their conceptual understanding during a series of inquiry-based science investigations over six weeks. The researchers used classroom observations to measure student motivation, and a test (pretest and posttest) to assess student understanding of concepts related to global warming. Those four students showing strong interests in science class were among 27 eighth graders who participated in the Global Warming project, which involved an inquiry approach to explore global climate phenomena. In the project, students worked together on the three investigations to answer the question, "Why do scientists think people are making the earth's climate warmer?" The researchers used transcriptions and descriptions of teacher and student conversations and behavior. They identified students' self-competence perceptions, goal orientations (mastery and performance), thoughtfulness, and understanding of science concepts. The results indicated that even though the students showed high levels of motivation and engagement, they differed from each other in terms of types of motivation. Among the four, a female student who gained the most on the conceptual test within the class exhibited self-competence, strong mastery goal orientation, and low performance goal orientation. Also, her statements were thoughtful and meaningful. On the contrary, a male student whose score on the conceptual test did not increase exhibited self-competence, a very low mastery goal orientation, and strong performance goal orientation. Many of his comments were superficial. The researchers concluded that different combinations of motivational beliefs have different implications for student conceptual understanding. Particularly, students having higher mastery orientation and lower performance orientation will likely increase their understanding of science concepts. The researchers also mentioned that they confirmed previous studies that indicated boys appeared to be more strongly performance oriented than girls.

Cross-Cultural Studies

Current school reform in the United States is challenging not only national standards, but also global standards. Many of the cross-cultural studies on student achievement have dealt with mathematics education. Although not in the area of science, these studies help elaborate on ways that learning and teaching differ in the classroom in different cultures. After a review of this research in mathematics, this review turns to the relevant focus in science.

Mathematics

Researchers attempted to understand the reasons for Asian (e.g., Chinese, Japanese, and Korean) students' academic success in mathematics. Stevenson and Lee (1990) examined a total of 1,440 first and fifth grades students in Taiwan, Japan, and the United States. In this study, the children were administered a mathematics achievement test, the children and their mothers were interviewed, and the children's teachers were given a questionnaire. The results found significant differences between Asian countries and the U.S. Compared to American mothers, Chinese and Japanese mothers showed higher interest and held higher standards for their child's academic achievement. Chinese and Japanese mothers stressed the importance of effort while American mothers tended to emphasize innate ability. Therefore, the researchers concluded that some of the reasons for the high academic achievement of Chinese and Japanese students are influenced by these sociocultural aspects.

Another study by Stevenson, Lee, Chen, Lummis et al. (1990) also examined the mathematics achievement of elementary school children. The participants consisted of 3,607 first and fifth grades children from Beijing, China and Chicago. This study included interviewing mothers as well as children. The results showed that American children believed they were doing well in mathematics, and they did not perceive mathematics as difficult subject. They also found that American mothers held lower standards for academic achievement. Therefore, the researchers concluded that American children's poor performance could be attributed to children's low motivation for devoting attention to mathematics and their parents' lower standards.

Lummis and Stevenson (1990) were interested in examining student achievement from developmental and gender perspectives. The participants were children in kindergarten, first grade, and fifth grade in Taiwan, Japan, and the United States. They were administered cognitive tasks and achievement tests in reading and mathematics. The researchers found gender differences on achievement tests of mathematics and reading in all 3 countries. As early as the first grade, however, boys appeared to do better in solving word problems and problems involving visual estimation of quality and distance. The researchers further examined expectations and beliefs of children's mothers. They found that mothers in all three countries similarly distinguished boys and girls. Mothers believed that girls were better readers than boys, and boys were better in mathematics than girls. The researchers concluded that these biases of mothers may be conveyed to children and affect their academic achievement.

Previous studies highlighted factors contributing to higher achievement of Asian students such as high standards, emphasis on effort, and a high value on education. Chen and Stevenson (1995) were interested in whether these factors are found in Asian-American students, and if so how they are related to their motivation and mathematics achievement comparing to that of Caucasian-American and East Asian students. The participants included 304 Asian-American, 1,958 Caucasian-American, 1,475 Chinese, and 1,120 Japanese eleventh graders. For the mathematics achievement, Asian-American students scored statistically significantly higher than did Caucasian-American students, but their scores were statistically significantly lower than those of Chinese and Japanese students. In all four groups, males scored statistically significantly higher than females. However, gender differences were smaller for the American than the East Asian students. For the standards and expectations, Japanese students had statistically significantly higher standards than other three groups. Further, Chinese and Asian-American students had statistically significantly higher standards than Caucasian-American students. Regarding students' beliefs about effort, the majority of Chinese and Japanese students rated "studying hard" as the most important factor in influencing performance in mathematics. However, the majority of both Asian-American and Caucasian-American students rated "a good teacher" as the most important factor. The researchers explained that cultural aspects of Asian-American students fell between those of East Asian and Caucasian-American students.

Cross-cultural study has become more important in understanding not only differences in student achievement but also differences in teaching practices in countries around the world. For the TIMSS 1995 Study, the researchers expanded the research methodology by incorporating a qualitative approach (video) into traditional quantitative methods (survey) in order to capture detail-specific cultural factors that influence student learning (Stigler, Gallimore, & Hiebert, 2000). From this study, the researchers found that teaching varied little within one culture. Japanese teachers give students problems to work on that they have not seen before, and expect students to struggle with solving problems before teaching how to do it. On the other hand, U.S. teachers tend to show students how to solve a problem, and then ask them to solve examples on their own. The researchers concluded that Japanese mathematic lessons focused more on development of mathematical concepts than U.S. lessons, which probably contributed to Japanese students' higher achievement. The researchers also pointed out that there was a gap between U.S. teachers' perceptions regarding current ideas about teaching and learning mathematics and their actual practices in their classrooms, which was not a new finding. The problem associated with educational reform and its research is that success is evaluated mistakenly on whether teachers are using certain approaches. The researchers concluded that genuine changes in teaching practices must be initiated by teachers themselves. As a possible alternative to U.S. reform models for improving teaching, they proposed a Japanese Lesson Study where "small group of teachers meet regularly to collaboratively plan, implement, evaluate, and revise lessons (Hiebert & Stigler, 2000, p. 10). In this way, the classroom lesson is seen as the smallest unit of teaching system that need to be reformed. Much of the cross-cultural research focused on student achievement and teaching practices in mathematics. A few studies investigated the area of science from cross-cultural perspectives. These are reviewed next.

Science

The most recent TIMSS, conducted in 2003 with 49 countries, indicated that science achievement of eighth graders in the United States has improved since 1999 and it ranked 9th internationally while that of Japanese eighth graders has been steady and ranked 5th (Martin, Mullis, Gonzalez, & Chrostowski, 2004). In both countries at the eighth grade, male students outperformed female students in science.

Interestingly, however, eighth grade students' attitudes toward science in these two countries were not consistent with their achievement rank order. The students' self-confidence in learning science was measured by four statements (e.g., I usually do well in science; I learn things quickly in science) and value of science was measured by seven statements (e.g., I enjoy learning science; I need to do well in science to get into the university of my choice). Students were asked to rate on these statements with 4-point Likert scale (1 = agree a lot; 2 = agree a little; 3 = disagree a little; 4 = disagree a lot). In the U.S., 56% of eighth graders indicated high self-confidence whereas 20% of Japanese students reported high self-confidence. Moreover, 35% of U.S. eighth graders reported that they placed high value on science whereas 19% of Japanese students indicated high value on science. According to the TIMSS 2003 report (Martin, Mullis, Gonzalez, & Chrostowski, 2004), not only Japan but also Chinese Taipei, Hong Kong, and Korea had lowest percentages of students in the high self-confidence category yet had high average science achievement. Since all of these are Asian countries, the TIMSS results indicated that there may be cultural traditions that encourage modest self-confidence.

Additionally, eighth grade science teachers were asked about their instructional practices. A TIMSS survey including 11 statements investigated how often teachers ask students to do scientific inquiry activities in science lesson. Teachers chose their response from the four options (every or almost every lesson; about half the lessons; some lessons; never) to the items (e.g., conduct experiments or investigations; relate what students are learning in science to their daily lives). In both the U.S. and Japan, approximately 50% or more of the teachers reported the use of student activities in half the lessons or more in conducting experiments, working together in small groups on experiments, and relating what students learned in science to their daily lives.

In summary, early cross-cultural studies investigated factors relating to the gap in students learning between the U.S. and Asian countries. Although they found differences in expectations in mathematics achievement and emphasis on efforts, more recent studies focused on classroom context such as teaching practices and teacher effects. In the next section, studies examined characteristics of teachers are reviewed.

Teacher Characteristics

For the last forty years, many researchers have empirically examined measurable factors affecting students' academic achievement. Researchers have examined a variety of characteristics of teachers such as class size (Fergason & Ladd, 1996; Hanushek, 1992) and teaching experiences (Ehrenberg & Brewer, 1994; Ehrenberg & Brewer, 1995). Results from these studies indicated that students achieved more when their class sizes were smaller, and when their teachers had more teaching experience. Negishi, Elder, and Mzoughi (2004) examined the effects of teacher characteristics on students' physical science and physics achievement. They found that students' achievement scores were higher in smaller classes with more experienced teachers.

Additional studies conducted in the context of physics further highlight teacher effects. Hestenes (1998) reported findings regarding teachers' competence from a study conducted with nearly 150 high school physics teachers. First, subject content knowledge was concluded to be vital to teacher effectiveness; yet "proficiency in scientific inquiry is more important than specific content knowledge" (p. 467). Second, teachers' discourse management skills such as planning and preparation as well as experience are essential to effective teaching. Lastly, teachers' constructivist viewpoints contribute to create an environment where students can actively construct their knowledge.

Geelan, Wildy, Louden, and Wallace (2004) also explored the characteristics of "expert" teachers by conducting a case study of physics teachers in a suburb, middle-class socioeconomic background, and public high school in Australia. They found that "expert" teachers tended to spend much of the class time talking and leading discussion. Small group work and student-student interaction were quite rare strategies in their classrooms. Instead, the expert teachers had an on-going conversation with the class, asking questions for particular students, and often following up for a number of interactions. Their teaching strategies placed values on deep understanding and high-level thinking skills over simple memorization and algorithmic solving of problems. Therefore, particularly in a physics context, teaching practice such as posing questions and leading discussion is more effective than traditional lectures.

Summary

Current science education faces many challenges. As part of technologically advanced nations, students in the United States and Japan are expected to develop knowledge and skills needed in promoting industrial productivity. As such, students need to be prepared for cultivating their scientific knowledge through the process of scientific investigations. Therefore, science curriculum has emphasized inquiry approach that exposes students to the natural world and encourages them to wonder how and why certain phenomena occur. In these learning environments students can actively engage in the process of investigation. Furthermore, an inquiry learning environment can promote conceptual understanding and can enhance student motivation such as mastery goals and task value.

Previous studies suggest a positive influence of inquiry-based teaching in enhancing students' motivation as well as their academic attainment. Cross-cultural study in teaching such as TIMSS has highlighted the practices that promote student learning and their motivation to learn. However, currently available information on cross-cultural perspectives in science education is limited to elementary or middle school levels. More cross-cultural research investigating student learning and motivation in advanced levels is called for. Therefore, this current study examined high school students' physics achievement, motivation, and inquiry-based instructional practices in the United States and Japan.

Research Purpose

The purpose of this study is to examine how inquiry-based instruction affects high school students' conceptual understanding and motivation in physics courses across two nations: the United States and Japan. Student understanding of physics concepts and their motivation to learn science were examined in the context of inquiry approach teaching practices using two-level hierarchical linear modeling (HLM) methods. HLM is an appropriate statistic technique recommended for nested data such as student-level variables, and teacher-level variables (Raudenbush & Bryk, 2002).

The research questions for this dissertation are:

1) To what degree do inquiry-based instructional practices explain differences in students' conceptual understanding in physics after controlling for the student-level variable of gender and teacher-level variables (teaching experience, class size) in the United States and Japan?

2) To what degree do inquiry-based instructional practices explain differences in students' motivation toward science after controlling for the student-level variable of gender and teacher-level variables (teaching experience, class size) in the United States and Japan?

3) To what extent do the HLM models differ by culture?

CHAPTER II

METHODOLOGY

Participants

Teacher participants in the United States included 9 physics teachers (8 females, 1 male). There were 7 public (co-ed) and 2 private (co-ed) schools. The average teaching experience was 12.33 years ranging from 3 to 33 years. The average class size was 13.44 students ranging from 7 to 21 students. A variety of class types included 6 Regular classes, 2 AP classes, and 1 Honor class. Student participants in the U.S. included 108 students (55 females, 53 males) who took physics courses in high schools in Mississippi during the academic year of 2004-2005 or 2005-2006. The students were either 11th or 12th graders.

Teacher participants in Japan included 11 physics teachers (1 female, 10 males). There were 6 public (co-ed), 5 private (2 girls, 2 boys, 1 co-ed) schools. The average teaching experience was 18.27 years ranging from 2 to 33 years. The average class size was 28.72 students ranging from 12 to 42 students. Class types were all regular physics classes although 8 were university bound, and 3 were junior college bound schools. Student participants in Japan included 616 high school students (203 females and 413 males) from 22 classrooms who were taking physics class during the academic year of 2005-2006. The students were all 11th graders except one 10th grade class.

Instruments

The current study examined the effects of inquiry-based instruction on students' understanding of Newtonian force concept and their motivation toward science. For students, two instruments were administered: the *Force Concept Inventory* and the *Attitudes about Science Questionnaire*. For teachers, another two instruments were given: *Teacher Survey* and *Teacher Interview*.

Force Concept Inventory

The Force Concept Inventory (FCI; see Appendix A), one of the most widely used physics concept tests in use today, can inform the effectiveness of physics instruction and it has been used as diagnostic assessment tool at every level of introductory physics instruction from high school to university.

The FCI was originally developed and published by Hestenes, Wells, and Swackhamer (1992) and revised by Halloun, Hake, Mosca, and Hestenes in 1995. The inventory consists of 30 multiple-choice format items designed to assess students' understanding of Newtonian force concept. Each question has five possible responses; one representing the Newtonian concept and four representing the most common misconceptions that students often believe. A student's FCI score is determined by summing the correct responses on the 30 items, ranging from 0 to 30.

Regarding the validity and reliability on the FCI, Hestenes, Wells, and Swackhamer (1992) did not report specific numbers when the FCI was published because the test design was so similar to the Mechanics Diagnostic Test (MDT) for which validity and reliability of scores had been established. According to Hestenes et al., about half of the FCI is the same as the MDT, and the FCI is considered as an improved version of the MDT. They interviewed 20 high school students about their written answers to the FCI questions. They found that the students repeated the answers they had marked on the written test. Further, the students had firm reasons for their choices. Hestenes et al. also interviewed 16 first-year graduate students beginning graduate mechanics. The interview was in depth on the questions the students missed on the FCI. It was found that the students' responses to the questions were consistent with their performance in a graduate mechanics class. Therefore, the FCI was considered to yield valid and reliable scores measuring knowledge of Newtonian concepts.

More recently, Hestenes and Halloun (1995) extended their justification of validity issues of the FCI. The face validity was examined by physics professors including their suggestions for improvement on wording or diagrams of the questions. The content validity was examined using interviews based on the responses by Newtonian thinkers (FCI score of 60% and above). They estimated the probability of a false negatives and false positives. An answer is a false negative if a Newtonian thinker chose a non-Newtonian response. An answer is a false positive if a Newtonian response was chosen for non-Newtonian reasons. The probability of a false negative was found to be less than ten percent. The probability of a false positive was not reported, but mentioned that the multiple choice test with five options have a 20% chance of false positive. The solution for reducing a false positive was including powerful distractors.

Attitudes about Science Questionnaire

The Attitudes about Science Questionnaire (ASQ) is a 32-item questionnaire (see Appendix B) adapted from the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich, Smith, Garcia, & McKeachie, 1991). The original MSLQ was designed to assess college students' motivational orientations and their use of learning strategies. There are two sections in the MSLQ, a motivation section and a learning strategy section, however, the ASQ adapted only the motivation section. The motivation section of the MSLQ consisted of 31 items that assess "students' goals and value beliefs for a course, their beliefs about their skill to succeed in a course, and their anxiety about tests in a course" (p.3). The validity and reliability of MSLQ scores were established from a sample of 380 Midwestern college students from 37 classrooms. The construct validity was tested by confirmatory factor analyses and confirmed six latent factors: Intrinsic Goal Orientation (4 items), Extrinsic Goal Orientation (4 items), Task Value (6 items), Control Beliefs about Learning (4 items), Self-Efficacy for Learning and Performance (8 items), and Test Anxiety (5 items). The reliability of scores for each construct based on Cronbach's alpha method ranged from 0.62 to 0.93.

The ASQ measures student motivation about science in five constructs (Appendix C): Self-Efficacy (6 items), Task Value (6 items), Mastery Goals (7 items), Ability Goals (6 items), and Extrinsic Goals (7 items). Students are asked to rate themselves on a 5 point Likert scale (5 = strongly agree, 4 = agree, 3 = slightly agree/slightly disagree, 2 = disagree, 1 = strongly disagree). A student's ASQ score is based on the mean score of each construct, ranging of 1 to 5. Table 2 displays the estimated reliability of these constructs based on the Cronbach's alpha for the U.S. students and the Japanese students.

Table 2

Reliability of the Attitudes about Science Questionnaire (ASQ)

Construct	Cronbach's alpha	Example items
Self-Efficacy		I expect to do well when we work with science.
U.S.	.69	
Japan	.75	
Task Value		It is important to me to learn about science.
U.S.	.84	
Japan	.90	
Master Goals		The main reason I do science experiments is because I can learn new things.
U.S.	.79	
Japan	.79	
Ability Goals		I want to do better on the science experiments than other students in the class.
U.S.	.77	
Japan	.76	
Extrinsic Goals		The main reason I do science experiments is because the teacher says so.
U.S.	.77	
Japan	.60	

Note: *n* = 108 in the U.S.; *n* = 616 in Japan

Teacher Survey

The teacher survey (see Appendix D) was developed by the investigator. The teacher survey included questionnaires about demographics and inquiry-based instruction practices. The demographic questions asked teachers to indicate background information such as number of years teaching high school physics, type of physics course, number of students in each class and school tuition.

The questions on teachers' inquiry-based instruction practices included 10 items partially adapted from the *National Science Education Standards* (National Research Council, 1996) and the *TIMSS 2003 International Science Report* (Martin, Mullis, Gonzalez, & Chrostowski, 2004). They measured how often teachers asked students to do inquiry-based methods according to the ten dimensions: Make observations, pose questions, examine textbooks/other resources, plan experiments, analyze data, formulate hypotheses, find solutions to real problems, share the results, work together in small groups on experiments, and discuss their ideas in class. Teachers were asked to rate themselves on a five-point Likert-type scale (5 = all the time, 4 = often, 3 = sometimes, 2 = seldom, 1 = never). The scoring is based on the composite score over the 10 items. The range of possible scores is from 10, indicating that the teacher never used any of inquiry-based instruction at physics, to 50, indicating that the teacher used all the inquiry-based instruction all the time. The reliability of the scores estimated by Cronbach's alpha was .72 for the U.S. teachers (n = 9) and .84 for the Japanese teachers (n = 11).

Teacher Interview

An interview protocol was planned as a supplement in an effort to elaborate on the cultural context and to better understand responses on the survey about inquiry teaching practices. Teacher interview questions (see Appendix E) were developed by the investigator. They aimed at obtaining qualitative information in addition to the quantitative data to facilitate further understanding of the possible differences between the two countries in teaching physics. The structured interview consisted of seven items with a combination of concrete and abstract questions. The first question (i.e., please tell me about your physics class) was general and intended to break the ice and obtain a big picture of a physics class. The second question (i.e., what are some challenges you face in teaching physics) was intended to identify some of the possible obstacles the teachers faced. The third question (i.e., please tell me about typical class activities students do) was intended to shift the focus from teachers to students' activities. The fourth question (i.e., which inquiry-based teaching practices are more important than other) was chosen to clarify teachers' responses on the teacher survey. The fifth question (i.e., what kind of advice would you give for a novice physics teacher) allowed teachers to reflect on their teaching experiences and to offer suggestions for others. The sixth question (i.e., what is the goal of teaching physics) is abstract and was intended to help understand why teachers design and use certain teaching approaches for their students. The last question (i.e., could you share any lesson plans and/or students' work) enabled the researcher to obtain relevant artifacts, if possible.

Instrument Translations

All the instruments were translated into Japanese for the Japanese students and Japanese teachers. The FCI was previously translated into Japanese by a group of Japanese physics researchers at the Tokyo University of Science. This translation was used in this study. The ASQ, the Teacher Survey, and the Teacher Interview instruments were translated by the investigator. In order to secure the original translation by the investigator, proof reading and back translations (translating Japanese version into English) were conducted by two Japanese persons holding graduate degrees from the U.S. institutions. The procedure involved a few exchanges with the translators. First, the investigator translated the original version (English) into Japanese. Second, after one Japanese person read the Japanese version, we discussed key words. In Japanese, the term for science as a curriculum in secondary education is *rika* while as a discipline including biology, chemistry etc. is *kagaku*. For the ASQ, *kagaku* was used. Third, after another Japanese person back translated the Japanese version into English, we discussed the differences between the original and the Japanese version. However, there were no significant differences. Some of the examples are:

Example 1

Original Version: It is important to me to learn about science.

Back Translation: It's important to me to study science.

Example 2

Original Version: It is okay with me to make mistakes in science if I am learning new things.

Back Translation: It's OK if I make mistakes when I study something new in science.

Procedure

Recruitment

The participants in the United States were recruited from high schools in Mississippi. They were also the participants for a summer workshop, Teacher Training in Physics (TTIP). The four-week workshop intended to reform pedagogy in physics by focusing on student centered, active, and constructive instruction. The teachers participated in laboratory work using a variety of technology, problem solving, and discussions of physics concepts and teaching strategies. For year 2004-2005, thirteen high school physics teachers were recruited; however, seven teachers actually taught physics during the following semesters. These seven teachers were included as U.S. teacher participants. For the year 2005-2006, twelve teachers were recruited including six teachers who attended in the previous year. Among the six teachers attending for the first time, two taught physics in the following semester. These two were included as U.S. participants. Therefore, a total of nine high school physics teachers in the U.S. participated in this study. Six teachers were from schools in the central region of Mississippi, two were from the northern region, and one was from the southern region of Mississippi.

The participants in Japan were recruited from the areas of Tokyo, Chiba, Kanagawa, and Hokkaido. The first contact was made to the principals at forty Tokyo Metropolitan High Schools by sending a letter describing the purpose of the study. As a result of follow-up phone calls to these principals, three physics teachers agreed to participate in the study. Due to the small number of participating teachers, the researcher contacted her former mentors and colleagues. From these contacts, six high school teachers were recruited. Also, the researcher contacted a high school physics teacher having a similar research interest, who helped in recruiting two more teachers to participate in this study. Therefore, a total of 11 high school teachers in Japan became the participants.

Data Collection

Data collection was conducted near the end of the school year in each country. The academic calendar in the two countries was different. The U.S. school year started in August and ended in May. The Japanese school year began in April and ended in March including 6-week summer vacation from July to August. Therefore, data were collected near the end of school year for the U.S. schools (April 2005/2006) and end of school year for Japanese schools (January 2006) (see Figure 1).

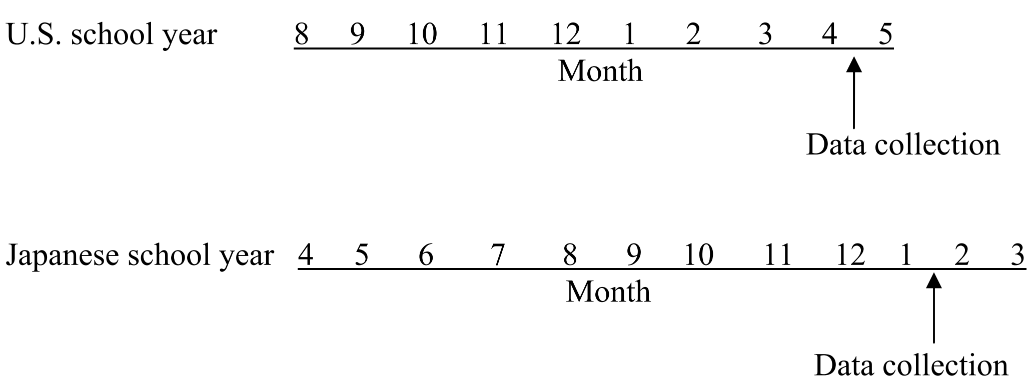

Figure 1. Point of Data Collection in the U.S. and Japanese Academic Calendars

The students were administered the FCI and the ASQ by their teachers in the U.S. and Japanese schools. All the students were asked to mark their responses on machine-readable answer sheets. They were also asked to mark their gender on the sheets.

The teachers in the U.S. completed the teacher survey as a part of the program evaluation of large project, which included a four-week workshop for high school physics teachers.

The teacher interview was conducted only with those who were willing to do so. Three Japanese teachers were interviewed in January 2006. Three U.S. teachers were interviewed in April 2006.

Data Analysis

In this analysis, two-level hierarchical linear modeling (HLM) methods were used. Much social research including educational research involves hierarchical data structures (Raudenbush & Bryk, 2002). Investigations of teacher effects are hierarchical because students are nested within teacher, school, state, and nation. Prior to the availability of multilevel methods such as HLM, there was conceptual and methodological difficulty in conducting this type of research (Lee, 2000). Traditional single-level methods, such as multiple linear regression and analysis of variance, traditionally use one unit of analysis, either as student-level or teacher-level. This approach ignores the substantial variance that exists in the dependent variable as well as the independent variables (Lee, 2000). HLM allows researchers to consider more than one unit of analysis. Therefore, this current study used two-level HLM to examine individual and unique associations of inquiry-based practices with the physics conceptual understanding and motivation toward science of high school students (Level-1) in physics classrooms (Level-2) in the U.S. and Japan, respectively. Table 3 provides students' outcome variables and predictors in each level.

Table 3

Descriptions of Student-Level and Teacher-Level Variables

Variable	Description
Student-Level (Level-1)	
Outcome Variables	
FCI	Total raw score on conceptual understanding in physics (Max = 30)
ASQ	Self-reported motivational beliefs (Average scores for each section)
Self-Efficacy	Belief in one's ability to accomplish a task (Max = 5)
Task Value	Beliefs in value of the task (Max =5)
Mastery Goals	Goals include learning and curiosity (Max = 5)
Ability Goals	Goals involve besting others (Max = 5)
Extrinsic Goals	Goals include grades and rewards (Max = 5)
Control Variable	
Gender	Students' gender (Female = 0, Male= 1)
Teacher-Level (Level-2)	
Predictor Variable	
Inquiry	Total score on self-reported frequency of using inquiry-based teaching practices (Max = 50)
Control Variables	
Experience	Years of teaching experience in physics
Class Size	Number of students in a class

CHAPTER III

RESULTS

Inquiry-based instruction has been encouraged by the standards both in the United States and Japan. Thus, the impact of inquiry-based teaching practices has important implications for how students are motivated to learn physics and how their learning is attained. Current research focused on three research questions: "Does more frequent use of inquiry method have positive impact on students' learning?" "Does more frequent use of inquiry method have positive impact on students' motivation?" and "What are the differences and similarities of the inquiry teaching in the United States and Japan?"

Descriptive Statistics

Students' Conceptual Understanding and Motivation

On the Force Concept Inventory (FCI), the average score of U.S. students (M = 9.11, SD = 4.64) was statistically significantly different from that of Japanese students (M = 11.45, SD = 5.02), t (722) = -4.52, p <.001. Table 4 presents the descriptive statistics.

On the Attitudes toward Science Questionnaire (ASQ), five motivational constructs were examined: self-efficacy, task value, mastery goals, ability goals, and extrinsic goals. A MANOVA yielded a statistically significant difference between U.S.

students and Japanese students, F (5, 718) = 28.67, $p < .001$. The univariate tests with an alpha level of .05 yielded statistically significant differences in self-efficacy, F (1, 722) = 101.27, MSE = .50, $p < .001$; task value, F (1, 722) = 22.23, MSE = .80, $p < .001$; mastery goals, F (1, 722) = 28.82, MSE = .51, $p < .001$; and extrinsic goals F (1, 722) = 4.75, MSE = .34, p = .03. By comparing the means, U.S. students had higher self-efficacy (M = 3.49, SD = 0.57), task value (M = 3.81, SD = 0.72), mastery goals (M = 3.76, SD = 0.65) and extrinsic goals (M = 2.47, SD = 0.69) than Japanese students' self-efficacy (M = 2.75, SD = 0.73), task value (M = 3.38, SD = 0.92), mastery goals (M = 3.35, SD = 0.73), and extrinsic goals (M = 2.34, SD = 0.57). The mean scores of ability goals were not statistically significantly different between the two countries, p = .26.

Table 4

Means and Standard Deviations of Students in the U.S. and Japan

Variable	U.S. *M* (*SD*) *n* = 108	Japan *M* (*SD*) *n* = 616	ES
FCI*	9.11 (4.64)	11.45 (5.02)	0.48
ASQ			
Self-Efficacy*	3.49 (0.57)	2.75 (0.73)	1.14
Task Value*	3.81 (0.72)	3.38 (0.92)	0.52
Mastery Goals*	3.76 (0.65)	3.35 (0.73)	0.59
Ability Goals	2.85 (0.77)	2.76 (0.76)	0.12
Extrinsic Goals*	2.47 (0.69)	2.34 (0.57)	0.21

* $p < .05$

Gender Differences in Conceptual Understanding and Motivation

In the U.S., there was no statistically significant gender difference in the average scores of the FCI, t (106) = -3.8, p = .71. A MANOVA at an alpha level of .05 revealed no statistically significant gender difference in the average scores of the ASQ, F (5, 102) = 1.36, p = .25. Table 5 presents means and standard deviations of U.S. students.

In Japan, on the other hand, there was a statistically significant gender difference in the average scores of conceptual understanding on the FCI, t (614) = -3.9, $p < .001$. Japanese female students (M = 10.34, SD = 4.31) scored significantly lower than did Japanese male students (M = 12.00, SD =2.25). A MANOVA at an alpha level of .05 yielded a statistically significant gender difference in the average scores of the ASQ, F (5, 610) = 6.10, $p < .001$. A follow-up, univariate tests revealed a statistically significant gender difference in self-efficacy scale, F (1, 614) = 13.46, MSE = .52, $p < .001$. Japanese female students (M = 2.59, SD = 0.68) indicated significantly lower self-efficacy than did Japanese male students (M = 2.82, SD = 0.74). No other constructs were statistically significant. Table 6 displays means and standard deviations of Japanese students.

Table 5

Means and Standard Deviations for Females and Males in the U.S.

Variable	Females *M* (*SD)* *n* = 55	Males *M* (*SD)* *n* = 53	ES
FCI	8.95 (4.53)	9.28 (4.79)	0.07
ASQ			
Self-Efficacy	3.43 (0.55)	3.55 (0.59)	0.20
Task Value	3.81 (0.72)	3.82 (0.74)	0.01
Mastery Goals	3.69 (0.66)	3.83 (0.64)	0.22
Ability Goals	2.71 (0.73)	2.99 (0.79)	0.37
Extrinsic Goals	2.40 (0.62)	2.54 (0.75)	0.20

Table 6

Means and Standard Deviations for Females and Males in Japan

Variable	Females *M (SD)* *n* = 203	Males *M (SD)* *n* = 413	ES
FCI*	10.34 (4.31)	12.00 (2.25)	0.51
ASQ			
Self-Efficacy*	2.59 (0.68)	2.82 (0.74)	0.32
Task Value	3.39 (0.90)	3.37 (0.93)	0.02
Mastery Goals	3.38 (0.65)	3.34 (0.76)	0.06
Ability Goals	2.68 (0.69)	2.80 (0.79)	0.16
Extrinsic Goals	2.32 (0.52)	2.35 (0.59)	0.05

* $p < .05$

Teachers' Use of Inquiry-Based Instruction

A MANOVA at an alpha level of .05 revealed a statistically significant difference between teachers in the U.S. and Japan, F (3, 16) = 13.40, $p < .001$. The univariate tests yielded statistically significant differences in the use of inquiry-based teaching practices, F (1, 18) = 19.14, MSE = 26.65, $p < .001$; and class size, F (1, 18) = 17.70, MSE = 65.63, $p = .001$. The U.S. teachers (M = 40.33) reported higher frequency of inquiry-based instruction use than did Japanese teachers (M = 30.18). The average class size in the U.S. (M = 13.44) was statistically significantly smaller than that in Japan (M = 28.72). The average year of teaching experience in physics was not statistically significantly different, $p = .24$. Table 7 presents results of descriptive analysis for teacher participants.

Table 7

Means and Standard Deviations for Teachers in the U.S. and Japan

Variable	U.S. *M (SD)* *n* = 9	Japan *M* (*SD)* *n* = 11	ES
Inquiry*	40.33 (3.71)	30.18 (6.08)	2.07
Experience	12.33 (11.83)	18.27 (10.03)	0.54
Class Size*	13.44 (4.82)	28.76 (9.98)	2.07

* $p < .05$

Inquiry-based teaching practices based on the composite scores were significantly different between two countries, yet further analysis of the ten dimensions of inquiry methods may help to better describe cultural differences. A rank order of the ten dimensions of teaching practices (using mean scores) for each country is presented in

Table 8. Note for each dimension teachers were asked to rate how often they ask students to perform each activity. For the most part, there were marked dissimilarities in ranking by country. U.S. teachers highlighted use of small group work, data analysis, and discussing ideas as evidenced by these activities having a top 5 ranking. Japanese teachers, on the other hand, stressed real life problems, examining textbooks, and formulating hypotheses. Further, examining the textbook was the least frequently used activity in the U.S. These rankings indicate a similarity across two countries in that both noted frequent use of posing questions and making observations.

Table 8

Inquiry-Based Teaching Practices: U.S. and Japan

Ranking	U.S. *n* = 9	Japan *n* = 11
1	Make observations (*M* = 4.56)	Pose question (*M* = 3.82)
2	Pose questions (*M* = 4.44)	**Real life problems (*M* = 3.82)**
3	**Small group work (*M* = 4.33)**	**Examine textbooks (*M* = 3.45)**
4	**Analyze data (*M* = 4.22)**	Make observations (*M* = 3.45)
5	**Discuss their ideas (*M* = 4.11)**	**Formulate hypotheses (*M*=3.18)**
6	Real life problems (*M* = 3.89)	Small group work (*M* = 2.91)
7	Share the results (*M* = 3.89)	Analyze data (*M* = 2.73)
8	Plan experiments (*M* = 3.78)	Discuss their ideas (*M* = 2.45)
9	Formulate hypotheses (*M* = 3.67)	Share the results (*M* = 2.18)
10	Examine textbook (*M* = 3.44)	Plan experiments (*M* = 2.18)

Interviews with teachers in the U.S. ($n = 3$) and Japan ($n = 3$) were conducted to better understand the cultural context of their inquiry teaching practices. Initially, the interview protocol had seven questions; however, there was a limited amount of time available for interview sessions (approximately 15 minutes on average). Because of this, many questions could not be asked. Nevertheless, two questions were asked for every teacher in both countries: (a) What do you ask students to do in your typical lesson and (b) what are the goals of teaching physics? The main idea each teacher conveyed are summarized in Table 9 and 10 (See Appendix F for the full interview transcripts).

U.S. teachers tended to emphasize small group work and posing questions whereas Japanese teachers had a tendency emphasizing on teacher's demonstration, students' hypotheses, and examining textbooks. However, teachers in two countries responded in a similar way regarding their goals for teaching high school physics. Both U.S. and Japanese teachers expect students to develop reasoning skills and thinking in physics courses and apply their knowledge beyond the classroom. In addition, teachers in both countries expressed their concerns about preparing students to enter college.

Table 9

Inquiry-Based Teaching Practices: U.S. Teachers

Teacher	Gender	Teaching experience	Activities activities	Teaching goals	Interview excerpts
US-1	Female	6 years	Application	Reasoning skills Hands-on	*"I try to do a lot of hands-on with those groups because I think they learn more hands-on than they are just sitting like we use to and take notes and work on problems"* *"Well, one thing I think that physics helps to develop their reasoning and logic… their thinking skills and I want them to think out of the box"*
US-2	Male	20 years	Application Real problem	Thinking skills	*"I apply things to real life. It is like you are boiling eggs or you are driving a car, you know, I always try to find something call for their attention" "I really want them to know how the universe is, how the things work, how this universe has order"*
US-3	Female	15 years	Review Problem solving	Mastery	*"…we review some of trigonometric function that they were supposed to learn" "…if I can help bring their math skills up, bring the level of physics and knowledge up to a base level that college professor can come and refine that. But, my ultimate goal is that I always tell my students, 'I am preparing you for college'"*

Table 10

Inquiry-Based Teaching Practices: Japanese Teachers

Teacher	Gender	Teaching experience	Activities activities	Teaching goals	Interview excerpts
JP-1	Male	20 years	Demo Hypotheses	Scientific thinking	*"I demonstrate experiments by focusing on prediction because I do think regular experiment is just an operation. Having students predict, hypothesize, or define the problems are true experiments" "…simply attaining knowledge is not an ultimate goal. Rather, scientific way of thinking is more important for them to acquire"*
JP-2	Male	21 years	Demo Textbook	Mastery	"*I do lecture and demonstrate experiments. A difficulty I have in this class is getting responses from the students" "I need to prepare students for college entrance exams such as Center Exam. So, my goal of teaching this class is for students to master the textbook"*
JP-3	Male	31 years	Experiments Hypotheses	Logical thinking	*"I ask students to do experiments by themselves, but I also try to demonstrate experiments in front of the students as many as possible" "I want students to predict or hypothesize what are going to happen and if there is any rule or principle" "I want students to think logically, record what they observed, and express those in words"*

Hierarchical Linear Modeling

Two-level hierarchical linear modeling (HLM) methods were used to examine the effects of gender as student level variable (Level-1) and inquiry-based teaching, teaching experience, and class size as teacher level variables (Level-2) in United States and Japan, separately. A typical HLM model involves three phases. The first phase involves fully unconditional model in which no predictor was specified at either student level or teacher level (see Tables 11 and 12 for results). The second phase involves a model examining the effects of Level-1 variable (see Tables 13 and 14 for results). The final phase involves a model examining Level-2 variables (see Tables 15 and 16 for results).

U.S. Models

Conceptual Understanding

The First Phase

The fully unconditional model (USA_00) examined differences in students' conceptual understanding on the FCI across classrooms. This model yielded an estimated reliability of 0.91 for the mean scores on the FCI at the teacher level, which indicated that there was a large variation in the FCI scores across classrooms. The final estimation test of variance component coefficient of 11.92 at teacher level was statistically significant, χ^2 (8) = 77.83, $p < .001$. This indicated that 48% of the variation in students' scores on the FCI was due to between teacher or classroom differences. The next model examined the effect of student level variable.

The Second Phase

The second model (USA_10) involving student gender as Level-1 control variable resulted in an estimated reliability of 0.91 for the average scores between classrooms. Gender was not a statistically significant variable in accounting the differences in FCI, estimated slope = 0.45, $p = 0.54$. The final estimation test of variance component coefficient of 12.05 at teacher level was statistically significant, χ^2 (8) = 77.72, $p < .001$. This indicated that 48% of the variation in students' scores on the FCI was due to differences between teachers or classrooms. For this sample, student gender appeared not to help explain the differences in student performance on the conceptual test. Therefore, the next model examined the effects of teacher level variables.

The Third Phase

The final model (USA_11) examined the effects of inquiry-based teaching practices on students' conceptual understanding by including inquiry as predictor variable and teaching experience and class size as control variables. The final model resulted in an estimated reliability of 0.92 on the mean FCI scores between classrooms. This implied that even after controlling for teaching experience and class size, the average FCI scores can distinguish among classes or teachers. Further, none of the teacher level variables had statistically significant influence on the FCI scores as well as the slope of gender, $p > .05$. The variance component coefficient of 14.27 at teacher level was statistically significant, χ^2 (5) = 59.49, $p < .001$. This indicated that 51% of the variation in students' scores on the FCI was due to between class or teacher differences. Thus, neither student gender nor teacher level variables (inquiry, teaching experience, and class size) appeared to help explain the differences in students' performance on the FCI.

Motivation: Self-Efficacy

The First Phase

The fully unconditional model (USA_Efficacy_00) resulted in an estimated reliability of 0.58 for the average scores of self-efficacy scale at the teacher level. The estimated variance component coefficient of 0.04 of difference in students' self-efficacy across teachers was statistically significant, χ^2 (8) = 19.33, p = .01. This indicated that 12% of the variation in students' self-efficacy was due to teacher variation. A larger variance component coefficient of 0.30 was estimated at the student level. Therefore, the next model examined the effect of the student level variable.

The Second Phase

The second model (USA_Efficacy_10) including student gender as Level-1 control variable resulted in an estimated reliability of 0.55 for the average scores between classrooms. Gender was not a statistically significant variable in accounting the differences in self-efficacy, estimated slope = 0.08, p = 0.48. The final estimation test of variance component coefficient of 0.03 at teacher level was statistically significant, χ^2 (8) = 18.04, p = 0.02. This indicated that 9% of the variation in students' scores on the self-efficacy scale was due to differences between teachers. For this sample, however, student gender appeared not to help explain the differences in self-efficacy. Therefore, the next model examined the effects of teacher level variables.

The Third Phase

The final model (USA_Efficacy_11) involved inquiry as predictor variable and teaching experience and class size as control variables. The final model resulted in an estimated reliability for the average score on self-efficacy scale of 0.30 at teacher level, after controlling for teaching experience and class size. According to Raudenbush, Bryk, and Congdon (2000), the precision of estimation depends on the sample size within each teacher; therefore, low reliabilities do not invalidate the HLM analysis. Nevertheless, none of the teacher level variables had a statistically significant influence on self-efficacy nor did the gender slope, $p > .05$. The estimated variance component coefficient of 0.01 at teacher level was not a statistically significant, χ^2 (5) = 7.70, $p = 0.17$. This indicated only 3% of the variation in students' scores on the self-efficacy scale was due to differences between classes or teachers after inquiry teaching practices, teaching experience, and class size were taken into account.

Motivation: Task Value

The First Phase

The fully unconditional model (USA_Value_00) resulted in an estimated reliability of 0.66 for the students' average score of task value between classrooms. The estimated variance component coefficient of 0.08 of difference in students' task value across teachers was statistically significant, χ^2 (8) = 23.64, $p = 0.003$. This indicated that 15% of the variation in students' task value was due to the difference between teachers.

A larger estimated variance component coefficient of 0.46 was at student level. Therefore, next model examined the effect of student level variable.

The Second Phase

The second model (USA_Value_10) including students' gender yielded an estimated reliability of 0.66 for the students' mean scores of task value across classrooms. Gender was not statistically significant variable in accounting the differences in task value, estimated slope = - 0.04, p = 0.79. The estimated variance component coefficient of 0.08 for the variation in students' task value at teacher level was statistically significant, χ^2 (8) = 23.67, p = 0.003. For this sample, student gender appeared not to help explain the differences in student task value.

The Third Phase

The final model (USA_Value_11) examined teacher level variables. The estimated reliability for the students' mean scores on task value across teachers was 0.05. This implied that the average scores of task value at the teacher level became indistinguishable after controlling teacher variables including inquiry-based teaching, teaching experience and class size. Teaching experience was statistically significant variable in explaining the variation in students' task value, estimated slope = - 0.02, p = 0.047. This negative relationship indicated that students in more experienced teacher had lower task value. However, inquiry-based teaching practice and class size were not statistically significant, p = .056, p = .894, respectively. The estimated variance component coefficient of 0.002 for the students' task value at teacher level was not statistically significant, χ^2 (5) = 5.33, *p* = 0.38. This indicated that 0.4% of the variation in students' task value was due to differences among teachers or classrooms.

Motivation: Mastery Goals

The First Phase

The fully unconditional model (USA_Mastery_00) resulted in an estimated reliability of 0.71 for the students' mean scores of mastery goals between classrooms. The variance component coefficient of 0.08 of difference in students' mastery goals across teachers was statistically significant, χ^2 (8) = 27.49, *p* = 0.001. This indicated 18% of the variation in mastery goals was due to teacher variation. A larger variation component coefficient was at the student level. Therefore, the next model examined the effect of student variable.

The Second Phase

The second model (USA_Mastery_10) including students' gender yielded an estimated reliability of 0.70 for the students' mean scores of mastery goals across classrooms. Student gender was not a statistically significant variable in explaining the differences in mastery goals, estimated slope = 0.09, *p* = 0.44. The estimated variance component coefficients of 0.08 of difference in students' mastery goals at teacher level was statistically significant, χ^2 (8) = 26.42, *p* = 0.001. This indicated that 18% of the variation in mastery goals was due to teacher variation. For this sample, student gender appeared not to help explain the differences in student mastery goals. Therefore, the next model included teacher variables.

The Third Phase

The final model (USA_Mastery_11) resulted in an estimated reliability of 0.43 for the average scores of students' mastery goals across classrooms after controlling for teaching experience and class size. None of the teacher level variables was statistically significant variable at alpha level of .05. The estimated variance component coefficient of the difference in mastery goals at teacher level of 0.02 was not statistically significant, χ^2 (5) = 8.47, p = 0.13. This indicated 5% of the variation in mastery goals was due to teacher variation.

Motivation: Ability Goals

The First Phase

The fully unconditional model (USA_Ability_00) resulted in a reliability of 0.65 for the mean score of students' ability goals across teachers. The estimated variance component coefficient of 0.09 for the differences in ability goals at teacher level was statistically significant, χ^2 (8) = 22.57, p = 0.004. This indicated that 15% of the variation in ability goals was at teacher level. A larger estimated variance component coefficient was at student level. Therefore, the next model examined the effect of student level variable.

The Second Phase

The second model (USA_Ability_10) involving student gender yielded an estimated reliability of 0.68 for the students' average ability goals across teachers. Gender was statistically significant variable in explaining the variation in ability goals, estimated slope = 0.30, $p = 0.04$. As female students were coded 0 and male students were coded 1, this indicated that male students reported, on average, 0.30 units higher ability goals than female students. The estimated variance component coefficient of 0.10 for the variation of ability goals at teacher level was statistically significant, χ^2 (8) = 24.52, $p = 0.002$. This indicated 16% of the variation in students' ability goals was at teacher level. Therefore, the next model included teacher variables.

The Third Phase

The final model (USA_Ability_11) yielded an estimated reliability of 0.65 for the students' average score on ability goals across teachers after controlling for teaching experience and class size. None of the teacher level variables was statistically significant at alpha level of .05. The estimated variance component coefficients of 0.08 of the difference in ability goals at teacher level was statistically significant, χ^2 (5) = 13.74, p = 0.02. This indicated that 14% of the variation in students' ability goals was due to teacher variation.

Motivation: Extrinsic Goals

The First Phase

The fully unconditional model (USA_Extrinsic_00) resulted in a reliability of 0.57 for the students' average extrinsic scores across teachers. The estimated variance component coefficient of 0.05 for the differences in extrinsic goals at teacher level was statistically significant, χ^2 (8) = 18.48, $p = 0.02$. This indicated 10% of the variation in students' extrinsic goals was at teacher level. A larger estimated variance component coefficient was observed at the student level. Therefore, the next model examined the effect of student variable.

The Second Phase

The second model (USA_Extrinsic_10) resulted in an estimated reliability of 0.59 for the students' mean extrinsic goals across teachers. Gender was not statistically significant variable in accounting the differences in extrinsic goals, estimated slope = 0.16, $p = 0.22$. The variance component coefficients of 0.06 for the students' extrinsic goals at teacher level was statistically significant, $\chi^2(8) = 19.17$, $p = 0.01$. This indicated 12% of the difference in students' extrinsic goals was due to teacher variation. Therefore, the next model examined teacher variables.

The Third Phase

The final model (USA_Extrinsic_11) yielded an estimated reliability of 0.58 for the students' average extrinsic goals across classrooms after controlling for teaching experience and class size. None of the teacher variables were statistically significant variable in accounting for the differences in students' extrinsic goals at alpha level of .05. The estimated variance component coefficient of 0.05 for the variation in extrinsic goals at teacher level was statistically significant, $\chi^2(5) = 11.38$, $p = 0.04$. Therefore, the final model did not appear to help explain the differences in students' extrinsic goals.

Japanese Models

Conceptual Understanding

The First Phase

The fully unconditional model (JPN_00) examined difference in students' conceptual understanding on the FCI across teachers. This model yielded an estimated reliability of 0.89 for the mean of the FCI scores at teacher level, which indicated that the average scores across classrooms or teachers were distinguishable. The final estimation test of variance component coefficient of 4.58 at teacher level was statistically significant, χ^2 (10) = 196.01, $p < .001$. This indicated that 19% of the variation in students' scores on the FCI was due to teacher variation. A larger estimated variation coefficient was observed at the student level. Therefore, the next model examined the effect of student level variable.

The Second Phase

The second model (JPN_10) involving students' gender as Level-1 control variable resulted in an estimated reliability of 0.92 for the students' average FCI scores between classrooms. Gender was a statistically significant variable in accounting the differences in FCI, estimated slope = 2.53, $p < .001$. As female students were coded 0 and male students were coded 1, this implied that the intercept for male students was, on average, 2.53 points higher than for female students. The final estimation test of variance component coefficient of 5.98 at teacher level was statistically significant, χ^2 (10) = 232.47, $p < .001$. This indicated that 25% of the variation in students' scores on the FCI was due to teacher variation. For this sample, student gender explained differences in the performance on the conceptual test.

The Third Phase

The final model (JPN_11) examined the effects of inquiry-based teaching practices on students' conceptual understanding by including inquiry as predictor variable and teaching experience and class size as control variables. The final model resulted in an estimated reliability of 0.93 on the mean FCI scores between classrooms or teachers. This implied that even after including teaching experience and class size, the average FCI scores among teachers or classrooms were distinguishable. Further, none of the teacher level variables was statistically significant at an alpha level of .05. The variance component coefficient of 7.32 at teacher level was statistically significant, χ^2 (7) = 219.87, $p < .001$. This indicated that 29% of the variation in students' scores on the FCI was due to teacher variation.

Motivation: Self-Efficacy

The First Phase

The fully unconditional model (JPN_Efficacy_00) resulted in an estimated reliability for the students' average self-efficacy across teachers of 0.75. The estimated variance component coefficient of 0.04 of difference in students' self-efficacy across classrooms was statistically significant, χ^2 (10) = 57.19, $p < .001$. This indicated that 7% of the variation in students' scores in self-efficacy was due to teacher variation. A larger variance component coefficient of 0.50 was observed at the student level. Therefore, the next model examined the effect of the student level variable.

The Second Phase

The second model (JPN_Efficacy_10) including student gender as Level-1 control variable resulted in an estimated reliability of 0.83 for the students' average scores between classrooms. Gender was a statistically significant variable in accounting the differences in self-efficacy, estimated slope = 0.31, $p < .001$. As female students were coded 0 and male students were coded 1, this implied that the intercept for male students was, on average, 0.31 higher than for female students. The final estimation test of variance component coefficient of 0.06 at the teacher level was statistically significant, χ^2 (10) = 73.91, $p < .001$. This indicated that 11% of the variation in students' self-efficacy was due to teacher variation. For this sample, student gender explained differences in self-efficacy. The next model examined the effects of teacher level variables.

The Third Phase

The final model (JPN_Efficacy_11) involved inquiry as predictor variable and teaching experience and class size as control variables. The final model resulted in an estimated reliability of 0.84 for the students' average self-efficacy at teacher level. Nevertheless, none of the teacher level variables had statistically significant influence on self-efficacy nor did the gender slope, $p > .05$. The estimated variance component coefficient of 0.07 at teacher level was statistically significant, χ^2 (7) = 55.05, $p < .001$. This indicated only 13% of the differences in students' self-efficacy was due to differences among teachers or classrooms.

Motivation: Task Value

The First Phase

The fully unconditional model (JPN_Value_00) resulted in an estimated reliability of 0.88 for the students' average score of task value between teachers. The estimated variance component coefficient of 0.15 of difference in students' task value across teachers was statistically significant, χ^2 (10) = 139.58, $p < .001$. This indicated that 17% of the variation in students' task value was due to the difference between teachers. A larger estimated variance component coefficient of 0.72 was observed at the student level. Therefore, the next model examined the effect of student level variable.

The Second Phase

The second model (JPN_Value_10) including students' gender yielded an estimated reliability of 0.89 for the students' mean scores of task value across teachers. Gender was not a statistically significant variable in accounting the differences in task value, estimated slope = 0.09, $p = 0.28$. The estimated variance component coefficient of 0.16 for the variation in students' task value at teacher level was statistically significant, χ^2 (10) = 145.30, $p < .001$. For this sample, student gender appeared not to help explain the differences in student task value.

The Third Phase

The final model (JPN_Value_11) examined teacher level variables. The estimated reliability for the students' mean scores on task value across teachers was 0.89 after controlling for teaching experience and class size. None of the teacher variables was statistically significant in explaining the variation in students' task value at alpha level of .05. The estimated variance component coefficient of 0.17 for the students' task value at teacher level was statistically significant, χ^2 (7) = 98.35, $p < .001$. This indicated that 19% of the variation in students' task value was due to teacher variation.

Motivation: Mastery Goals

The First Phase

The fully unconditional model (JPN_Mastery_00) resulted in an estimated reliability of 0.78 for the students' mean scores of mastery goals between teachers. The variance component coefficient of 0.04 of difference in students' mastery goals across teachers was statistically significant, χ^2 (10) = 73.80, $p < .001$. This indicated 7% of the variation in mastery goals was due to teacher variation. A larger variation component coefficient was at student level. Therefore, the next model examined the effect of student variable.

The Second Phase

The second model (JPN_Mastery_10) including students' gender yielded an estimated reliability of 0.77 for the students' mean scores of mastery goals across teachers. Student gender was not a statistically significant variable in explaining the differences in mastery goals, estimated slope = -0.03, $p = 0.64$. The estimated variance component coefficients of 0.04 of difference in students' mastery goals at teacher level was statistically significant, χ^2 (10) = 73.20, $p < .001$. This indicated that 8% of the variation in mastery goals was due to teacher variation. For this sample, student gender appeared not to help explain the differences in student mastery goals. Therefore, the next model included teacher variables.

The Third Phase

The final model (JPN_Mastery_11) resulted in an estimated reliability of 0.78 for the average scores of students' mastery goals across teachers after controlling for teaching experience and class size. None of the teacher variables was statistically significant at alpha level of .05. The estimated variance component coefficient of the difference in mastery goals at teacher level of 0.04 was statistically significant, χ^2 (7) = 39.98, $p < .001$. This indicated 8% of the variation in mastery goals was due to teacher variation.

Motivation: Ability Goals

The First Phase

The fully unconditional model (JPN_Ability_00) resulted in a reliability of 0.60 for the mean score of students' ability goals across classrooms. The estimated variance component coefficient of 0.02 for the differences in ability goals at teacher level was statistically significant, χ^2 (10) = 28.12, $p = 0.002$. This indicated that 3% of the variation in ability goals was at teacher level. A larger estimated variance component coefficient was at student level. Therefore, next model examined the effect of student level variable.

The Second Phase

The second model (JPN_Ability_10) involving students' gender yielded an estimated reliability of 0.62 for the students' average ability goals across classrooms.

Gender was not a statistically significant variable in explaining the variation in ability goals, estimated slope = 0.13, $p = 0.06$. The estimated variance component coefficient of 0.02 for the variation of ability goals at teacher level was statistically significant, χ^2 (10) = 29.40, $p = 0.001$. This indicated 3% of the variation in students' ability goals was at teacher level. The next model examined teacher variables.

The Third Phase

The final model (JPN_Ability_11) yielded an estimated reliability of 0.72 for the students' average scores in ability goals across classrooms or teachers. None of the teacher level variables was statistically significant at alpha level of .05. The estimated variance component coefficients of 0.04 of the difference in ability goals at teacher level was statistically significant, χ^2 (7) = 26.90, $p = 0.001$. This indicated that 7% of the variation in students' ability goals was due to teacher variation.

Motivation: Extrinsic Goals

The First Phase

The fully unconditional model (JPN_Extrinsic_00) resulted in a reliability of 0.49 for the students' average extrinsic goals across classrooms or teachers. The estimated variance component coefficient of 0.01 for the differences in extrinsic goals at teacher level was statistically significant, χ^2 (10) = 21.60, $p = 0.02$. This indicated 3% of the variation in students' extrinsic goals was at teacher level. A larger estimated variance component coefficient was at student level. Therefore, next model examined the effect of student variable.

The Second Phase

The second model (JPN_Extrinsic_10) resulted in an estimated reliability of 0.49 for the students' mean extrinsic goals across classrooms or teachers. Gender was not a statistically significant variable in accounting the differences in extrinsic goals, estimated slope = 0.02, p = 0.64. The variance component coefficient of 0.01 for the students' extrinsic goals at teacher level was statistically significant, χ^2 (10) = 21.17, p = 0.02. This indicated 3% of the difference in students' extrinsic goals was due to teacher variation. Therefore, the next model examined teacher variables.

The Third Phase

The final model (JPN_Extrinsic_11) yielded an estimated reliability of 0.55 for the students' average extrinsic goals between teachers after controlling for teaching experience and class size. None of the teacher variables was statistically significant in accounting for the differences in students' extrinsic goals at alpha level of .05. The estimated variance component coefficient of 0.01 for the variation in extrinsic goals at the teacher level was statistically significant, χ^2 (7) = 16.00, p = 0.03. This indicated that the final model did not appear to help explain the differences in students' extrinsic goals.

Table 11

U.S. Participants: Unconditional HLMs

	DEPENDENT VARIABLE					
	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Student level variance	13.14***	0.30*	0.46**	0.36*	0.53**	0.44*
Teacher level variance	11.92	0.04	0.08	0.08	0.09	0.05
Proportion of variance†	48%	12%	15%	18%	15%	10%
Reliability of between teacher/class differences	0.91	0.58	0.66	0.71	0.65	0.57

* *p* < .05; ***p* <.01; ****p* < .001

† These figures are proportion of variation due to between teacher variability.

Table 12

Japanese Participants: Unconditional HLMs

	DEPENDENT VARIABLE					
	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Student level variance	19.40***	0.50***	0.72***	0.49***	0.56**	0.31*
Teacher level variance	4.58	0.04	0.15	0.04	0.02	0.01
Proportion of variance†	19%	7%	17%	7%	3%	3%
Reliability of between teacher/class differences	0.89	0.75	0.88	0.78	0.60	0.49

* *p* < .05; ***p* <.01; ****p* < .001

† These figures are proportion of variation due to between teacher variability.

Table 13

U.S. Participants: Student (1st) Level HLMs

	DEPENDENT VARIABLE					
INDEPENDENT VARIABLE	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Gender†	0.45	0.08	-0.04	0.09	0.30*	0.16
Variance explained††	0%	0%	0%	0%	4%	2%

* $p < .05$

† Coefficients are estimated differences in intercept between females and males; positive values imply females < males.

†† Proportion variance explained by the model = [variance (unconditional model) – variance (compositional model)] / variance (unconditional model)

Table 14

Japanese Participants: Student (1st) Level HLMs

	DEPENDENT VARIABLE					
INDEPENDENT VARIABLE	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Gender†	2.53***	0.31***	0.09	-0.03	0.13	0.02
Variance explained††	58%	4%	0%	0%	0%	0%

***$p < .001$

† Coefficients are estimated differences in intercept between females and males; positive values imply females < males.

†† Proportion variance explained by the model = [variance (unconditional model) – variance (compositional model)] / variance (unconditional model)

Table 15

U.S. Participants: Teacher (2nd) Level HLMs

INDEPENDENT VARIABLE	DEPENDENT VARIABLE					
	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Average intercept	22.37	4.97**	6.99***	6.62***	2.98	0.50
Inquiry	-0.21	-0.04	-0.07	-0.07	0.01	0.05
Experience (control)	-0.09	0.00	-0.02*	-0.01	0.01	0.01
Class Size (control)	-0.25	0.00	-0.00	0.01	-0.06	-0.02
Average gender slope	-1.25	1.15	-0.02	0.44	-2.42	0.08
Inquiry	0.05	-0.03	-0.01	0.00	0.07	0.01
Experience (control)	-0.04	-0.00	0.01	0.01	-0.02	-0.02
Class Size (control)	-0.06	-0.00	0.01	-0.04	0.00	-0.01
Variance explained††	0%	0%	0%	0%	11%	0%

* $p < .05$; ** $p < .01$; *** $p < .001$

†† Proportion variance explained by the model = [variance (unconditional model) – variance (compositional model)] / variance (unconditional model)

Table 16

Japanese Participants: Teacher (2nd) Level HLMs

	DEPENDENT VARIABLE					
INDEPENDENT VARIABLE	FCI	Efficacy	Value	Mastery	Ability	Extrinsic
Average intercept	12.50*	3.38***	4.57***	4.11***	2.93***	2.09***
Inquiry	-0.02	-0.02	-0.02	-0.02	-0.01	0.00
Experience (control)	-0.07	0.00	0.01	0.02	0.01	-0.01
Class Size (control)	-0.00	-0.01	-0.02	-0.01	0.00	0.01
Average gender slope	-4.28	-0.40	-0.12	-0.09	-0.03	-0.02
Inquiry	0.04	0.00	-0.01	0.00	0.01	0.00
Experience (control)	0.11	0.00	0.00	-0.02	-0.01	0.01
Class Size (control)	0.08	0.02	0.01	0.01	0.00	-0.01
Variance explained††	0%	0%	0%	0%	0%	0%

* $p < .05$; ***$p < .001$

†† Proportion variance explained by the model = [variance (unconditional model) – variance (compositional model)] / variance (unconditional model)

CHAPTER IV

DISCUSSION

In this cross-cultural study, high school students' motivational orientations and their understanding of physics concepts were examined in a context of instructional design in the United States and Japan. The purpose of the cross-cultural investigation is to compare not only student achievement, motivational beliefs, and instructional practices between the two countries but also patterns in these variables in classrooms within a country. Therefore, this dissertation utilized multilevel data analysis that took in consideration individual differences and classroom differences. Furthermore, to more fully understand each culture, interviews of three teachers in two countries supplied some additional qualitative understanding to the quantitative analysis.

Conceptual Understanding

U.S. versus Japan

Students' conceptual understanding was measured by the Force Concept Inventory (FCI) toward the end of the school year. As Table 4 presents, the overall average scores in both countries were very low: USA = 9.11 (30% correct) and Japan =

11.45 (38% correct). Further, the range of the scores in each country was 20% to 57% for the U.S. and 28% to 49% for Japanese students. This confirmed previous findings that regardless of class type (i.e., regular, AP, or honor), high school students scored low; in fact lower than what previous studies reported, ranging from 48% to 57% on a posttest (Halloun & Hestenes, 1985; Hestenes, Wells, & Swackhamer, 1992). Hestenes and Halloun (1995) suggested that an FCI score of 60% be the entry threshold to Newtonian physics, and that a student scoring below the threshold will not be able to solve problems effectively. Data from this study also indicated that students' misconceptions existed after completing almost an entire year of beginning physics in high school. Therefore, high school students demonstrate persistent difficulties in understanding Newtonian concepts, a foundation for future learning in physics. In addition, data from this study indicated that in the U.S. there was a large variation in the FCI scores across classrooms. This implies the need for further investigation of student conceptual understanding in different class types (e.g., regular, AP, or honor).

Gender

Gender is another important issue in science education. In Japan, there was a significant gender difference: male students outperformed female students in conceptual understanding. In contrast, in the U.S., no significant gender difference in FCI scores was found in this data. According to the TIMSS 2003 study, there were gender discrepancies in 8^{th} graders' science achievement in both Japan and the United States (Martin, Mullis, Gonzalez, & Chrostowski, 2004). This may imply that in Japanese culture, gender difference remains consistent across grades whereas in the American culture, the gender difference becomes more equitable when students get older. Another explanation may be that because physics is an elective course in the U.S., gender differences in conceptual understanding are less likely to occur than in Japanese high schools which require physics.

Motivation

U.S. versus Japan

Overall, U.S. students had higher self-efficacy, task value, mastery goals, and extrinsic goals compared to their Japanese counterparts. The TIMSS 2003 study also reported extremely low self-efficacy of Japanese 8^{th} grade students despite their high achievement in science and mathematics (Martin, Mullis, Gonzalez, & Chrostowski, 2004). They argued that Eastern culture may encourage modest self-confidence. Markus and Kitayama (1991) pointed out the cultural difference in self perceptions between Eastern and Western cultures. Many Asian cultures place emphasis on attending others, fitting in, and being harmonious with them. Western culture, in contrast to Eastern cultures, neither assumes nor emphasizes such an overt connectedness among individuals. Therefore, it is reasonable to expect that the degree of self perception of capability is lower in Asian countries including Japan than in Western countries such as the United States although all individuals deserve to achieve and feel capable of achieving physics.

Another possible explanation for Japanese students' lower task-value and mastery goals is that high school physics is a requirement in many schools in Japan. In the U.S., however, physics is an elective course. Therefore, students in a country such as the U.S., where physics is an elective, might value physics more which would result in higher mastery goals compared to those students in a country such as Japan where students are required to take physics regardless of their academic orientation.

Gender

In the U.S., there was a significant gender difference in ability goals. U.S. male students exhibited higher ability goals than female students. This result was consistent with previous findings that boys had greater competitiveness than girls in 8th grade science (Patrick & Yoon, 2004). This may indicate that boys and girls typically have different achievement goals (Maehr, 1983). Another possible explanation for boys having higher goals to best others could derive from classroom context. It is argued that science in general is a competitive enterprise (Maehr, 1983) which attracts boys' competitiveness.

In Japan, on the other hand, there was a significant gender difference in self-efficacy. Japanese female students reported lower self-efficacy than male students. Although that is not extensive research investigating gender differences in motivational orientations in Japanese culture, this finding could be explained by the male dominated physics classrooms in Japan, which mirrors Japanese culture. Bandura (1982) argued people develop self-efficacy by observing similar others perform successfully; therefore, Japanese female students may have a disadvantage because they see a limited number of other successful female students.

Teacher Characteristics

Instructional Practices

Although teachers in both countries reported relatively high frequency of using inquiry methods in physics course, data indicated that these inquiry practices were not associated with students' understanding of Newtonian concepts. These results were not consistent with findings from previous research on inquiry teaching (Ertepinar & Geban, 1996; Von Secker, 2002). One of the possible explanations is that this finding was an artifact derived from a small sample size. This indicated the necessity of increasing the number of teachers in future study. Another possible explanation is related to validity of the instrument. The teacher survey intended to measure their use of inquiry teaching on ten dimensions using a 5-point Likert scale. It may be that the response scale (5 = all the time, 3 = sometimes, 1 = never) was vague, especially perceptions of 'often,' 'sometimes,' or 'seldom' might be different in different cultures. In addition, the teacher survey was self-report; therefore, some teachers may respond on the survey in a way that is more desirable (i.e., more frequent use of inquiry teaching). Finally, it is possible that inquiry teaching may have less of an impact with high school students compared to younger students with whom most previous research was conducted. Perhaps, there are other effective instructional practices that benefit high school students' conceptual understanding and motivation in physics.

The qualitative data from teacher interviews pointed to different instructional structures in the two countries: U.S. teachers favored small group instruction and Japanese teachers preferred whole group instruction. These approaches may simply reflect cultural differences related to individualism versus collectivism.

Teaching Experience

In Japan, teaching experience was not related to student conceptual understanding on the FCI nor any of the motivational constructs. However, in the U.S., there was a negative impact of teaching experience on students' task value. This implied that students in a classroom where the teacher had more experience showed lower ratings of value for learning science. This is not consistent with previous study findings (Ehrenberg & Brewer, 1994; Ehrenberg & Brewer, 1995). This might indicate that more experienced teachers forget beginners' perceptions of task importance compared to less experienced teachers.

Class Size

For both countries, class size did not influence students' learning in physics nor their motivational beliefs for this sample. Previous research on class size reported that the smaller the class size the more students achieved (Fergason & Ladd, 1996; Hanushek, 1992; Negishi, Elder, & Mzoughi, 2004). Moreover, Japanese classrooms were significantly larger than U.S. classrooms. This trend may explain why Japanese tend to focus on learning as a group whereas Americans emphasize small group learning environments in which students engage in hands-on activities.

Gender

Although an examination of teachers' gender was beyond the scope of this study, there was a distinctive difference between the U.S. and Japan. The high school physics teachers in the U.S. were predominately female whereas the majority of Japanese physics teachers were male. This might reflect that in the U.S., school teachers are predominantly female and that in Japan, more males go into teaching profession, and moreover, physics classes are primarily taught by male teachers. This cultural difference in teachers' gender could affect students' interest in the subject and subsequently, their future career aspirations in the field of science. It could also impact their classroom interaction and performance in science. In other words, Japanese female students do not see many female role models in physics, which could relate to their lower self-efficacy and lower conceptual understanding compared to their Japanese male counterparts.

Limitations of the Study

There are at least four major limitations in this study. First, sample size was a major concern. More specifically, the small sample size at the teacher level limited power to detect potential results using multilevel analysis. With this study, there were nine teachers in the U.S. and eleven in Japan. The sample size recommended per level for HLM analysis is 30 observations (Kreft, 1996), and therefore, an optimal study should include a minimum 30 teachers with 30 students in each classroom.

A second limitation is generalizability. The participants in this study were a convenience sample. U.S. teachers were participants of a teacher training workshop. Therefore, they might have a social desirability to report higher frequency of using inquiry teaching practices. Japanese teachers included in this study volunteered to participate in this study, and thus they may have strong interests in teaching practices relating to student misconceptions. Therefore, the samples included in this current research do not necessarily represent the populations of high school students and their teachers in each culture.

A third limitation of this study is the limited information on student background such as prior achievement in science and a pre measurement of motivational orientations. Although all students in this study took physics for the first time, previous academic achievement is likely related to their learning of physics and knowledge as measured by the FCI. Although gender issues are important in this study, additional control variables such as socioeconomic status and class type (i.e., regular, AP, or honor), or academic tracks (i.e., junior college, university, advanced university bound) could be considered in the future.

The last limitation derives from the outcome measure of student conceptual understanding. The FCI was intended to measure misconceptions which have been proved to be stable and difficult to change (Caramazza, McCloskey, & Green, 1981;

Halloun & Hestenes, 1985). Therefore, the nature of the measurement may explain students' relativelt low levels of conceptual understanding in physics.

Implications for High School Physics

By examining student learning, motivational orientations, and teaching practices in the U.S. and Japan, this dissertation provided additional information for teachers and researchers to understand the reasons why students have difficulties in physics across cultures, why boys and girls have different achievement motivations, and why teaching practices are different in the two cultures. Understanding individuals' motivational beliefs is critical to help student develop optimal goals, values, and efficacy in learning physics.

Students' misconceptions about force and motion are pervasive even after completing almost an entire year of beginning physics across cultures. It is important to raise teachers' awareness of student misconceptions and develop instructional practices to effectively deal with student understanding of Newtonian concepts that are considered to be a foundation in physics.

An additional concern in science education in Japan, not limited to physics, is gender equity. In the Japanese co-ed classrooms in this study, male students outnumbered female students by at least a ratio of two to one. Female students' self-efficacy and achievement may have been affected by the disproportional ratio of boys to girls in physics classrooms. Since student competence in advanced mathematics and physics is related to industrial productivity and economical growth of the country, both female and male students should have opportunity to enhance their interests, values, and efficacy in this field of study.

Future Research

This dissertation research highlights some important issues that deserve further investigation. First, because student misconceptions in physics are pervasive in two cultures, there is a need to examine students' use of cognitive strategies to overcome these misconceptions. Perhaps, instruction in cognitive and metacognitive skills, such as setting goals and monitoring their learning processes may foster students to overcome their conceptual difficulties. It may also be enlightening to investigate if there are certain areas in Newtonian physics concepts with which students have more difficulties than other areas in each culture in an effort to better understand how to foster conceptual change for understanding physics.

Next, additional cross-cultural investigations of best teaching practices that promote motivation to learn science as well as enhancing conceptual understanding would enrich our current knowledge of teaching practices. It is generally recommended that physics teachers place more emphasis on deep understanding and high-level thinking skills by posing questions and leading discussions as opposed to traditional lectures or hands-on activities with minimal guidance (Geelan, Wildy, Louden, & Wallace, 2004). Future research should continue to investigate classroom practices and provide practical implications for teachers and students.

REFERENCES

Bandura, A. (1982). The self and mechanisms of agency. In J. Suls (Ed.), *Psychological perspectives on the self: Vol. 1.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Bandura, A., Barbaranelli, C., Caprara, G. V., & Pastorelli, C. (1996). Multifaceted impact of self-efficacy beliefs on academic functioning. *Child Development, 67,* 1206-1222.

Canton, E., Brewer, C., & Brown, F. (2000). Building teacher-scientist partnerships: Teaching about energy through inquiry. *School Science and Mathematics, 100*(1), 7-15.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity.* Washington, DC: U.S. Government Printing Office.

Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review, 13,* 1-17.

Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960s? *Coleman* revisited. *Economics of Education Review, 14,* 1-21.

Ertepinar, H., & Geban, O. (1996). Effect of instruction supplied with the investigative-oriented laboratory approach on achievement in a science course. *Educational Research, 38*(3), 333-341.

Eryilmaz, A. (2002). Effects of conceptual assignments and conceptual change discussions on students' misconceptions and achievement regarding force and motion. *Journal of Research in Science Teaching, 39*(10), 1001-1015.

Fergason, R. F., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp.265-298). Washington, DC: Brookings Institution.

Gakushu Shido Youryou (1998) Retrieved January 1, 2006, from http://www.mext.go.jp/a_menu/shotou/youryou/main4_a2.htm

Geelan, D. R., Wildy, H., Louden, W., & Wallace, J. (2004). Teaching for understanding and/or teaching for the examination in high school physics. *International Journal of Science Education, 26*(4), 447-462.

Gibson, H. L., & Chase, C. (2002). Longitudinal impact of an inquiry-based science program on middle school students' attitudes toward science. *Science Education, 86*(5), 693-705.

Halloun, I. A., & Hestenes, D. (1985). The initial knowledge state of college physics students. *American Journal of Physics, 53* (11), 1043-1048.

Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy, 100*, 85-117.

Hestenes, D., & Halloun, I. (1995). Interpreting the force concept inventory: A response to Huffman and Heller. *The Physics Teacher, 33*, 502-506.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher, 30*, 141-158.

Hiebert, J., & Stigler, J. W. (2000). A proposal for improving classroom teaching: Lessons from the TIMSS video study. *The Elementary School Journal, 101*(1), 3-20.

Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

Kozulin, A. (2004). Vygotsky's theory in the classroom: Introduction. *European Journal of Psychology of Education, 19*(1), 3-7.

Kreft, I. G. G. (1996, June 25). *Are multilevel techniques necessary? An overview, including simulation studies.* Retrieved February 4, 2005, from http://www.calstatela.edu/faculty/ikreft/quarterly/quarterly.html

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist, 35*(2), 125-141.

Levin, I., Siegler, R. S., & Druyan, S. (1990). Misconceptions about motion: Development and training effects. *Child Development, 61*, 1544-1557.

Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., & Chrostowski, S.J. (2004). *TIMSS 2003 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224-253.

National Research Council. (1996). *National Science Education Standards.* Washington, DC: National Academy Press.

Negishi, M., Elder, A., & Mzoughi, T. (2004, November). *The effect of teacher characteristics on students' physical science and physics achievement.* Paper presented at the annual meeting of the Mid-South Educational Research Association, Gatlinburg, TN.

Patrick, H., & Yoon, C. (2004). Early adolescents' motivation during science investigation. *The Journal of Educational Research, 97*(6), 319-328.

Payne, K. J., & Biddle, B. J. (1999). Poor school funding, child poverty, and mathematics achievement. *Educational Researcher, 28*(6), 4-13.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33-40.

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167-199.

Pintrich, P. R., & Schrauben, B. (1992). Students' motivational beliefs and their cognitive engagement in classroom academic tasks. In D. Schunk & J. Meece (Eds.), Student perceptions in the classroom: Causes and consequences. Hillsdale, NJ: Lawrence Erlbaum Associates.

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Technical Report No. 91-B-004). Ann Arbor, MI: The University of Michigan.

Raudenbush, Bryk, & Congdon (2000). *HLM 6 Hierarchical linear and nonlinear modeling.* Scientific Software International, Inc.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Rowan, B., Chiang, F., & Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, *70,* 256-284.

Sinatra, G. M., & Pintrich, P. R. (Eds.) (2003). *Intentional conceptual change*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stamp, N., & O'Brien, T. (2005). GK-12 partnership: A model to advance change in science education. *BioScience*, *55*(1), 70-77.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: The Free Press.

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using video survey to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist*, *35*(2), 87-100.

Tretter, T. R., & Jones, M. G. (2003). Relationships between inquiry-based teaching and physical science standardized test scores. *School Science and Mathematics, 103*(7), 345-350.

Von Secker, C. (2002). Effects of inquiry-based teacher practices on science excellence and equity. *The Journal of Educational Research*, *95*(3), 151-160.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

APPENDIX A

FORCE CONCEPT INVENTORY

The Force Concept Inventory is secured by the authors; however, interested parties can request a copy from http://modeling.la.asu.edu/R&E/Research.html

APPENDIX B

ATTITUDES ABOUT SCIENCE QUESTIONNAIRE

Attitudes about Science Questionnaire

We would like to find out what students really think about science. It is important to answer the questions as truthfully as possible. There are no right or wrong answers; it is just what you think. Please do your best.

NAME: ______________________________________ AGE: _______________

GENDER: ___________________ TEACHER: __________________________________

Please use the following scale for responding the following items:

1 / A	2 / B	3 / C	4 / D	5 / E
Strongly disagree	Disagree	Slightly agree/ Slightly disagree	Agree	Strongly agree

1. ___ I am sure I will be able to understand what we will learn about science.
2. ___ It is important to me to learn about science.
3. ___ I would feel really good if I was the only one who could do the science experiments correctly.
4. ___ It is okay with me to make mistakes in science if I am learning new things.
5. ___ The main reason I do science experiments is because I will get into trouble if I don't.
6. ___ The main reason I do science experiments is because I can learn new things.
7. ___ I want to do better on the science experiments than the other students in the class.
8. ___ I like learning about science.
9. ___ I expect to do well when we work with science.
10. ___ I think I will be able to use what I learn about science outside of school.
11. ___ I would like to show that I'm smarter than the other students by finishing my science experiments first.
12. ___ I want to keep working on science experiments until I understand them.
13. ___ An important reason why I do science experiments is to get complimented by my teacher.
14. ___ I want to do the experiments in science because they really make me think.
15. ___ When doing science experiments, I don't want to make mistakes because mistakes make me look dumb.
16. ___ I think learning about science is useful.
17. ___ I am sure I can do an excellent job on the experiments we will do with science.
18. ___ I think learning about science is interesting.
19. ___ When I do an investigation in science, I like to know if I did better than other students.
20. ___ If I get the wrong answer when working in science, it is really important for me to figure out why.

21. ___ I do my science experiments in order to get complimented by my parents.
22. ___ Understanding science is important to me.
23. ___ I know I will be able to learn about science.
24. ___ I do my science experiments because I have to, not because I want to.
25. ___ It is important to me that my teacher knows when I get a right answer in science.
26. ___ I am sure I can do work in science even if it is really hard.
27. ___ An important reason I do the work in science is because I want to get better at doing science.
28. ___ I feel good if I am the only one who can answer the teacher's questions about science.
29. ___ The main reason I do science experiments is because the teacher says so.
30. ___ No matter how hard I try there are some things about in science that I won't be able to understand.
31. ___ I want to do well in science so the other students in my class will think I am smart in science.
32. ___ Understanding experiments I do in science is more important to me than getting the right answer.

APPENDIX C

CODING SHEET: ATTITUDES ABOUT SCIENCE QUESTIONNAIRE

Coding Sheet: Attitudes about Science Questionnaire

Category

Efficacy

1. I am sure I will be able to understand what we will learn about science.
9. I expect to do well when we work with science.
17. I am sure I can do an excellent job on the experiments we will do with science.
23. I know I will be able to learn about science.
26. I am sure I can do work in science even if it is really hard.
30. No matter how hard I try there are some things about in science that I won’t be able to understand.

Value

2. It is important to me to learn about science.
8. I like learning about science.
10. I think I will be able to use what I learn about science outside of school.
16. I think learning about science is useful.
18. I think learning about science is interesting.
22. Understanding science is important to me.

Mastery Goals

4. It is okay with me to make mistakes in science if I am learning new things.
6. The main reason I do science experiments is because I can learn new things.
12. I want to keep working on science experiments until I understand them.
14. I want to do the experiments in science because they really make me think.
20. If I get the wrong answer when working in science, it is really important for me to figure out why.
27. An important reason I do the work in science is because I want to get better at doing science.
32. Understanding experiments I do in science is more important to me than getting the right answer.

Ability

3. I would feel really good if I was the only one who could do the science experiments correctly.
7. I want to do better on the science experiments than the other students in the class.
11. I would like to show that I'm smarter than the other students by finishing my science experiments first.

15. When doing science experiments, I don't want to make mistakes because mistakes make me look dumb.
19. When I do an investigation in science, I like to know if I did better than other students.
28. I feel good if I am the only one who can answer the teacher's questions about science.

Extrinsic

5. The main reason I do science experiments is because I will get into trouble if I don't.
13. An important reason why I do science experiments is to get complimented by my teacher.
21. I do my science experiments in order to get complimented by my parents.
24. I do my science experiments because I have to, not because I want to.
25. It is important to me that my teacher knows when I get a right answer in science.
29. The main reason I do science experiments is because the teacher says so.
31. I want to do well in science so the other students in my class will think I am smart in science.

APPENDIX D

TEACHER SURVEY

Teacher Survey

This survey is intended to know how you teach physics courses. Please answer all questions based on your experience during the academic year **2005-2006**.

Name: ______________________________

School: _____________________________________

Number of years you've taught high school physics: _________

Type(s) of physics courses you taught this year: ______________________________

Number of students in each class: ________________________

In teaching physics to the students, how often do you usually ask them to do the following?

1. Make observations	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
2. Pose questions	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
3. Examine textbooks and/or other resources of information	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
4. Plan experiments	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
5. Analyze data	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
6. Formulate hypotheses	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
7. Find solutions to real problems	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
8. Share the results	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
9. Work together in small groups on experiments	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never
10. Discuss their ideas in class	5 All the time	4 Often	3 Sometimes	2 Seldom	1 Never

APPENDIX E

TEACHER INTERVIEW

Teacher Interview

This interview is intended to know how teachers teach physics courses during the academic year **2005-2006**.

Teacher: ______________________________

School: _____________________________________

1. Please tell me about your physics class.
2. What are some challenges you face in teaching physics?
3. Please tell me about the typical class activities students do.
    a. How do you get started on those activities?
    b. How long do you spend on those activities?
4. (From the inquiry-based teacher survey)
    a. Which activities are more important than others?
    b. Which activities do you feel more comfortable to do?
5. Assume your best friend became a physics teacher next classroom. What kind of advice would you give for this teacher?
6. What is the goal of teaching physics? What do you want students to learn from physics?
7. Student activities
    a. What kind of project did the student do this semester?
    b. What kind of homework did you give to the students?
    c. Could you share any artifacts (students' work, documents, lesson plans etc.)?

APPENDIX F

TEACHER INTERVIEW TRANSCRIPTS

U.S. Teacher 1 (**US-1**; Female, 6-year experience, 9 students)
Q1:
What do you ask students to do in your typical lesson?
US-1:
A lot times I do introductory activities, like you observed today, to get them thinking and so when I say something they have the point of reference. *I try to do a lot of hands-on with those groups because I think they learn more from hands-on than they are just sitting like we used to and take notes and work on problems.* We still do problems and we still do notes but not the rigor that may have been going on when I was in high school.
Q2:
What are the goals of teaching physics?
US-1:
Well, one thing I think that physics helps to develop their reasoning and logic their thinking skills and I want them to think out of the box. I want them to be creative and I try to give assignments out side of the class. For instance, right now the senior has a project. There are couples of topics that are off limit just because in the past I know that everybody wants to do those, but I made them sign up and you cannot duplicate. Once somebody has named the topic, another student cannot take it. I have one who turned in automobile racing and I have one that is working on snowboarding. And so, it's making them see physics through on every avenue of their life not just in a classroom. *And I want them to see that they can carry over what they learned here into our life.*

U.S. Teacher 2 (**US-2**; Male, 20-year experience, 11 students)
Q1:
What do you ask students to do in your typical lesson?
US-2:
I apply kind of thing to real life. It is like you are boiling eggs or you are driving a car, you know, I always try to find something call their attention. For example, because they like music, when I teach sounds I go with that. There is always question everybody wants to know how the things work, what is behind everything. Everybody needs to know that. But the thing is that it is different when you have to spend time and read a book and study. They don't want to do that. In general, regular physics students don't have the interest in the subject. That's something teacher has to do to motivate them and find different ways. It is really hard to do, but you have to.
Q2:
What are the goals of teaching physics?
US-2:
I really want them to know how the universe is, how the things work, how this universe has order. Also physics helps students develop personal skill like a will. You do not have to be strong physically, but strong will. Something you get it done. You get it from sample of life. If you want to be in good shape physically, I go to the gym and get really big muscle. Right? You do the same thing for the brain. When you teach, *you get trained to think. This is something that physics really helps. That's something I really would like to achieve with the students.*

U.S. Teacher 3 (**US-3**; Female, 15-year experience, 14 students)
Q1:
What do you ask students to do in your typical lesson?
US-3:
Students have missed out somewhere along with the trigonometry and I think that's one major obstacle to have to overcome. Once we kind of went back and *review some of trigonometric function that they were supposed to learn, they realize that they could not go any further until they actually master that.* And they start studying those I guess they call it trigonometric function. They get to know them to begin to use some feel comfortable then we can actually move on.
Q2:
What are the goals of teaching physics?
US-3:
I think one major goal that I have is making sure that I provide students with a foundations whenever they get to college they can build on that foundation. I know that I do not have time to cover all the aspects of physics because so broad, but if they know kinematics, they know electricity and magnetism, then surely *if I can help bring their math skills up, bring the level of physics and knowledge up to a base level that college professor can come and refine that. But, my ultimate goal is that I always tell my students, 'I am preparing you for college*. Keep that in mind'
Q3:
If you assume that your best friend becomes a physics teacher, teaching next door to you, what kind of suggestions would you give to teach physics effectively?
US-3:
Don't assume too much from your students. Make sure that you know where your students are and be able to pick those students up and bring them to the level that you are expect them to be at. A lot of times, new physics teachers would step into the classroom and assume too much assume that mathematical skills are there and you have to maybe go back and like I referred to before develop skills or just bring those skills back to into current memory than per se and then, knew that students to that level. I think that the major difficulty for beginning physics teacher because they are coming straight from college that college level course they took and it's like 'OK, I am jumping in and we are going to teach just like what they taught to me' and you really cannot do that. Because if you discourage a student in the beginning, he is going to quit. In another words, he'll drop the class and sign up for another class. Or the student will over a period of time, just give up because he doesn't understand anything what's going on with class, and just accept whatever grades they get. So, it's a major disadvantage for the students whenever the teacher probably could have boosted the spirits and the staying of the student by helping.

Japanese Teacher 1 (**JP-1**; Male, 20-year experience, 38 students)
Q1:
What do you ask students to do in your typical lesson?
JP-1:
I demonstrate experiments by focusing on prediction because I do think regular experiment is just an operation. Having students predict, hypothesize, or define the problems are true experiments. In my class, I use this approach, but at the same time I have to teach from the textbook, especially for those students who are going to take college entrance examinations.
Q2:
What are the goals of teaching physics?
JP-1:
I want students to experience the joy of science. Everyone before learning scientific experiments is in a same stage in terms of their knowledge about science. However, *simply attaining knowledge is not an ultimate goal. Rather, scientific ideas and scientific way of thinking is more important for them to acquire.* Also, it is important how students use what they learned in physics after graduating from high school.

Japanese Teacher 2 (**JP-2**; Male, 21-year experience, 33 students)
Q1:
What do you ask students to do in your typical lesson?
JP-2:
I do lecture and demonstrate experiments. A difficulty I have in this class is getting responses from the students. They study and they do well on exams, but they are just quiet in the class. So, sometimes I cannot tell if they really understood or not and I tend to repeat the same things. I occasionally distribute the outline of the class for students to follow.
Q2:
What are the goals of teaching physics?
JP-2:
I want students to enjoy understanding physical phenomenon in natural world. I do not believe that students have to conduct experiments by themselves. Rather, teacher should demonstrate the experiments and show the essential components to the students. But, sometimes I ask students to assist my demonstrations. I also use textbook in class because *I need to prepare students for college entrance exams such as Center Exam. So, my goal of teaching this class is for students to master the textbook.*

Japanese Teacher 3 (**JP-3**; Male, 31-year experience, 31 students)
Q1:
What do you ask students to do in your typical lesson?
JP-3:
I ask students to conduct experiments by themselves, but I also try to demonstrate experiments in front of the students as many as possible. I want students to predict or hypothesize what are going to happen and if there is any rule or principle. I would

demonstrate the experiments because I do not want students to concentrate on the operation too much. So, I show the experiments using two or three different patterns of conducting it. I want them to catch the whole idea.

Q2:

What are the goals of teaching physics?

JP-3:

The goal is discovering the essence of natural phenomenon. *I want students to think logically, record what they observed, and express those in words.* Students need to be aware of whatever the objects or phenomenon that they wonder. Then they can solve the problems by experiencing the essence of natural phenomenon.