What matters in inquiry-based science instruction?

Man Hung

University of Utah

December 7, 2010

Author Note

Correspondence concerning this manuscript should be addressed to Man Hung, Assistant Professor, School of Medicine, 590 Wakara Way, Salt Lake City, Utah 84108. Email: Man.Hung@hsc.utah.edu.

Abstract

With the overarching goal of improving scientific literacy for all, the emphasis on science achievement had been expanding all over the world. In this study, I examined the total, direct and indirect effects of inquiry-based science instruction on U. S. 8th graders' science achievement through attitudes toward science. I also examined the differential mediated effects of inquiry on science achievement among female and male students. Using data from the Trend in International Mathematics and Science Study (TIMSS) 2003, a series of structural equation modeling analyses were performed. Results showed that the total effect of inquiry-based activities on students' science achievement was slightly positive, and that students' attitudes toward science fully mediated the path between inquiry and achievement. In addition, the results of this study illustrated the differential mediated effects that inquiry and science achievement could have on each other, depending on students' gender. This study added to our understanding of factors that affect science achievement and suggested ways to improve scientific literacy.

Introduction

Scientific literacy was identified by many countries around the world as the main goal in science education during the last century and it still remains as the main goal in the science education community today. Many professional organizations such as the American Association for the Advancement of Science, the National Research Council and the National Science Teachers Association are increasingly advocating the preparation of scientifically literate citizens (Abd-El-Khalick & BouJaoude, 1997; American Association for the Advancement of Science, 1990; Lederman, 1992; McComas & Olson, 1998; National Research Council, 1996; National Science Teachers Association, 2000).

Many educational reform documents had been published outlining the standards for science teaching and learning. The *Project 2061: Science for All Americans* (AAAS, 1990), the *Project 2061: Benchmarks for Science Literacy* (AAAS, 1993), and the *National Science Educational Standards* (NRC, 1996, 2000) are a few examples among these numerous reform documents. The common goal in these documents is to produce scientifically literate citizens capable of understanding science concepts and processes that occur in everyday lives (AAAS, 1990, 1993; Aldrige, 1992a, 1992b; NRC, 1996; Yager, 1996). To increase the overall level of science achievement, reform documents (AAAS, 1990, 1993; Aldrige, 1992b; NRC, 1996; OERI, 1994; Yager, 1996) called for inquiry-based instruction, and efforts to increase students' positive attitudes toward science.

Literature Review

The major reform documents placed strong emphases on how the role of inquiry-based science instruction might affect students' attitudes toward science and students' science achievement. They made explicit the need for engaging students in inquiry-based activities. The *National Science Education Standards* (NRC, 1996) envisioned that students engaging in inquiry should "describe objects and events, ask questions, construct explanations, test those explanations against current scientific knowledge, and communicate their ideas to others. They identify their assumptions, use critical and logical thinking, and consider alternative explanations" (p. 2). To engage students in inquiry, teachers place less emphasis on covering material in science textbooks and more emphasis on having students working together to explore questions/problems and to debate evidence/conclusions, which is similar to what scientists do in the lab or in the field (AAAS, 1990, 1993; NRC, 1996; OERI, 1994; Welch, 1981; Yager, 1996).

Within the framework of motivation (Bandura, 2001), attitudes are conceptualized as an individual's characteristics interacting with a particular environment (e.g., a school, a subject). In other word, a student possesses a self-regulatory system that enables the development of attitudes toward school or attitudes toward science. Such a self-regulatory system directly affects students' achievement by influencing their attitudes (Pajares, 2002; Pajares, & Schunk, 2001).

Science educators had also recognized the importance of attitudes toward science in science learning (Blosser, 1984; Koballa, 1995; Schibeci, 1984; Simpson, Koballa, Oliver, & Crawley, 1994). Positive attitudes toward science are reflected in statements such as "science is fun", "I have good feelings toward science", "I enjoy science courses", "everyone should learn about science", and "science is useful in everyday life."

Intuitively, one would assume that attitudes toward science would be positively related to science achievement. Several studies indeed found such positive correlations (House, 1996; Lee & Burkam, 1996; Rennie & Punch, 1991; Simpson & Oliver, 1990).

Relationship between inquiry-based learning and attitudes toward science. In inquiry-based learning, students are engaged in science activities through observing, hypothesizing, predicting, testing, conducting experiments, asking questioning, and making inferences much the same as real scientists do (Roth, 1992). However, the relationship between inquiry-based learning and attitudes toward science is unclear. Some studies (Alouf & Bentley, 2003; Berg et al., 2003; Bredderman, 1983; Ebenezer & Zoller, 1993; Gibson & Chase, 2002; Kahle, 1992; Lord & Orkwiszewski, 2006; Shymansky et al., 1990; Shymansky et al., 1983) indicated positive association between inquiry-based learning and attitudes toward science. Other studies either found no relationship (Roth, 1992; Smith & Anderson, 1984) or were inconclusive (Brunkhorst, 1992; DeBoer, 1991; Hofstein, & Lunetta, 1982).

Based on Shymansky et al.'s (1983) original meta-analysis of 105 studies and Shymansky et al.'s (1990) reassessment of their original meta-analysis, inquiry-based instruction was found to affect students' attitudes toward science positively and significantly. An overall mean effect size of 0.19 was reported. Hence, it is reasonable to hypothesize that inquiry-based learning improves attitudes toward science.

Relationship between inquiry-based learning and science achievement. The notion of achievement can mean different things to different people and can be assessed in numerous ways. The National Assessment of Educational Progress described science achievement as "the three basic elements of science literacy: science knowledge, scientific habits of mind, and the ability to solve problems and conduct inquiries" (NCES, 1992, p. 117). Eccles (1984) pointed out that achievement has been operationally defined as grades in schools, scores on standardized tests of achievement and task performance. In the present study, science achievement is defined as performance on a standardized test.

The relationship between inquiry-based learning and science achievement is also unclear. A number of studies found a positive relationship between inquiry-based learning and science achievement (Alofu & Bently, 2003; Amaral et al., 2002; Bredderman, 1983; Brunkhorst, 1992; Jorgenson & Vanosdall, 2002; Luckie, Maleszewki, Loznak, & Krha, 2004; NRC, 1985; Shymansky et al., 1983, 1990; Von Secker, 2002). There were also other studies indicting no significant relationship between inquiry-based learning and science achievement (Bates, 1978; Booth, 2001; DeBoer, 1991; Roth, 1992).

In Shymansky et al.'s (1990) re-synthesis of their 1983 meta-analysis, inquiry-based approaches were reported to have a significant overall mean effect size of 0.30 on students' science achievement. In Brunkhorst's (1992) study, he compared middle and high school students' test scores on the *Iowa Tests of Basic Skills, Science Supplement.* The mean score of the students (N = 280) who performed inquiry-based activities in their science classroom exceeded 87% of the scores in national sample (N = 1982). Hence, it is also reasonable to hypothesize that inquiry-based learning increases science achievement.

Relationship between attitudes toward science and science achievement. The initial research in attitudes and achievement was influenced by Bloom's (1976) theory of school learning. Bloom (1976) claimed that 25% of the variance in student achievement could be attributed to students' attitude toward the subject, their school environment and their self-beliefs. Since then, many studies had been conducted to examine the relationship between attitudes toward science and science achievement, yet the exact nature of this relationship is still unclear (Rennie & Punch, 1991; Simpson et al., 1994).

Haladyna and Shaughnessy (1982) conducted a meta-analysis of 49 studies, and found a small positive relationship between attitudes toward science and science achievement. Among the 19 synthesizable studies, attitudes toward science explained 0.01% to 12.2% of the variance in science achievement. In a study examining the relationship between high school students' attitudes toward science and achievement, Rennie and Punch (1991) found that science achievement was positively related to attitudes toward science. Mattern and Schau (2002) also suggested that positive attitudes toward science improve science achievement. Thus, it is reasonable to hypothesize that a positive relationship exists between attitudes toward science and science achievement.

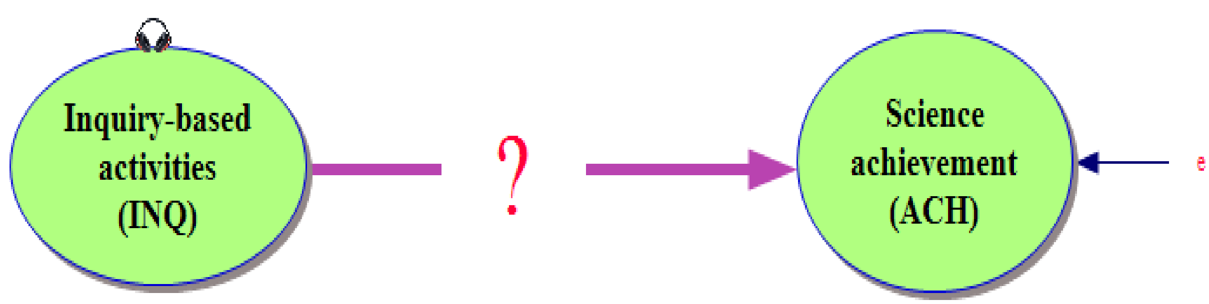
Relationship between gender and attitudes toward science. The relationship between gender and attitudes toward science seemed to be unclear. In a study conducted by Ye et al. (1998) to compare gender differentiated attitudes toward science in American and Chinese secondary school students, they reported that there were no significant differences, although boys had slightly more positive attitudes toward science. Oakes and the RAND Corporation (1990) also found more positive attitudes toward science in boys than girls. They claimed that the inquiry-based activities would benefit girls more than boys. However, there were conflicting evidences from Jarvis and Pell (2005) that inquiry-based activities led to higher attitudes toward science for girls but not for boys. In Schibeci's (1984) review of literature, he found that it was extremely difficult to draw any conclusion on the relationship among gender and attitudes toward science and he stated that gender alone might not be significant but rather the interaction between gender and other factors such subject matter and instructional practices. Based on all these studies, it is fair to say that gender may interact with inquiry-based activities to alter students' attitudes toward science.

Summary. In conclusion, research seemed to show that inquiry-based instruction predicted both attitudes toward science and science achievement, whereas attitudes toward science also predicted science achievement. In addition, gender seemed to moderate the effect of inquiry on attitudes toward science. However, the exact relationships among inquiry-based instruction, gender, attitudes toward science and science achievement were still quite unclear. As such, there was a need to study these relationships. The present study specifically addressed the following three research questions.

Research Questions

1. What is the total effect of inquiry-based science activities have on students' science achievement?

Figure 1: Research question #1 – Total effect of inquiry on achievement



2. Do students' attitudes toward science mediate the relationship between inquiry-based activities and students' science achievement? If yes, what are the direct and indirect effects? Specifically, what direct relationship does inquiry-based science activities have on students' science achievement, after adjusting for students' attitudes toward science? What is the indirect effect of inquiry-based activities on science achievement after controlling for attitudes toward science?

Figure 2: Research question #2 – Direct effect of inquiry on achievement

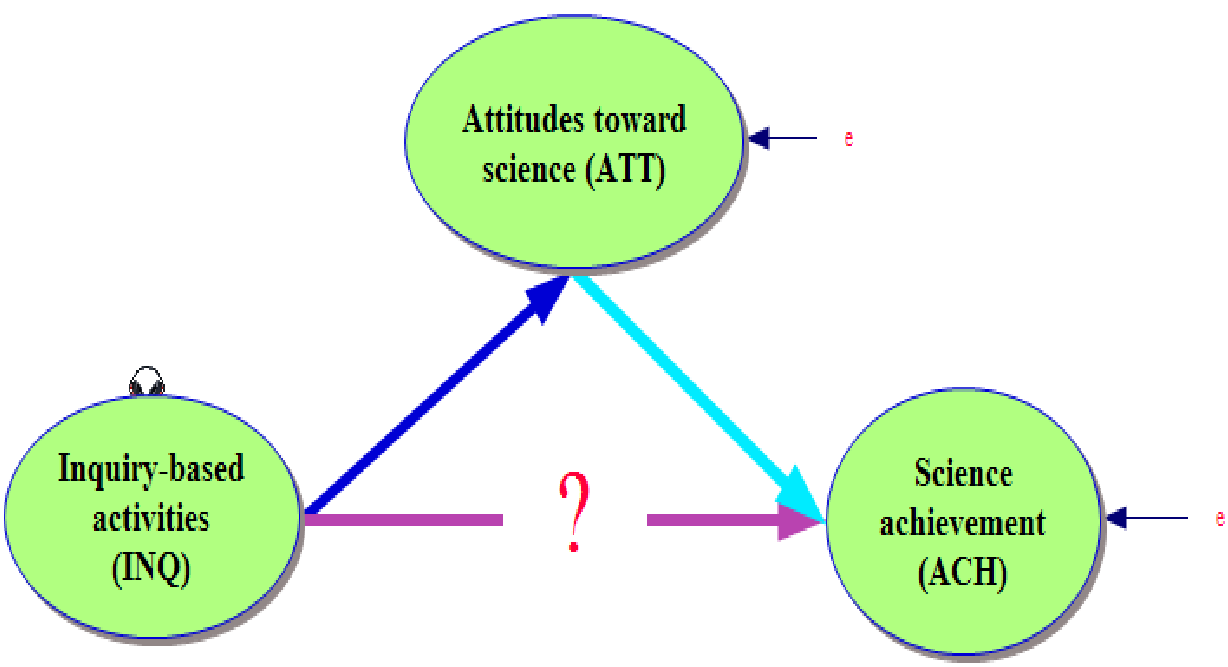
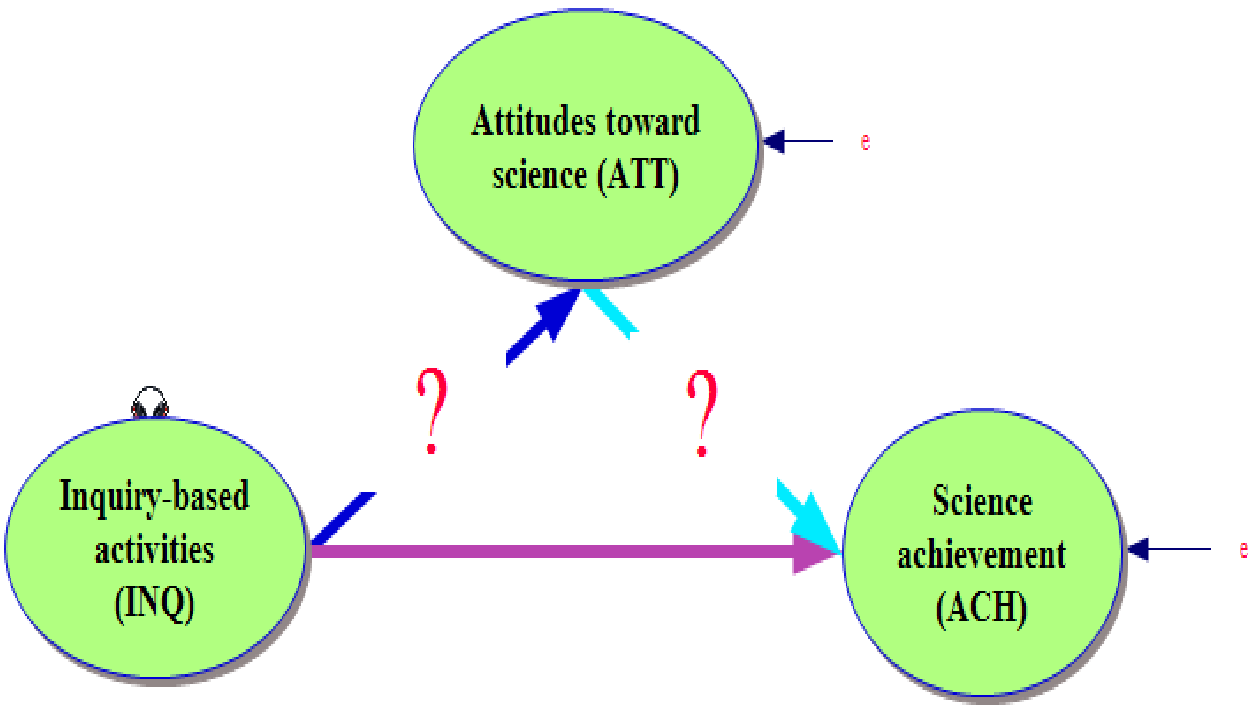
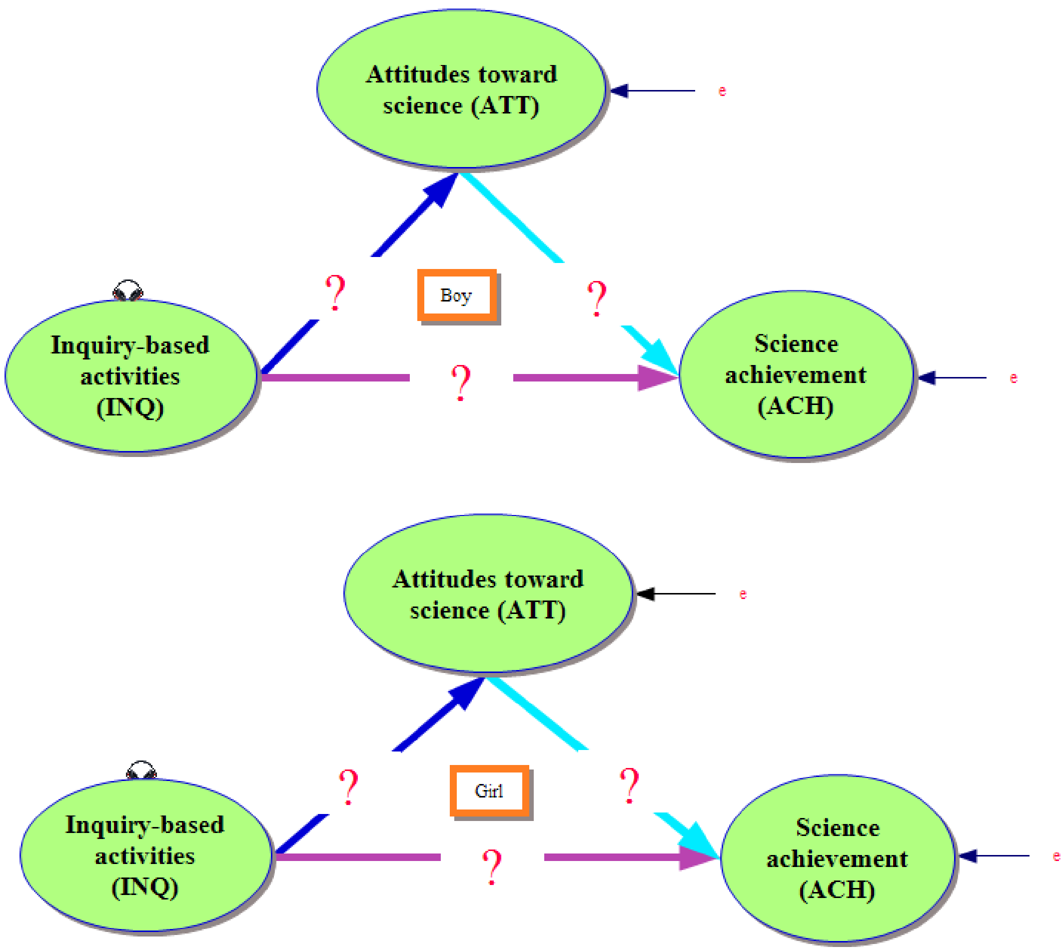


Figure 3: Research question #2 – Indirect effect of inquiry on achievement



3. Does gender moderate the overall relationship between inquiry-based activities and students' science achievement? If yes, how does the relationship between inquiry and achievement differ for boys and girls? If no, does gender moderate the mediated relationship between inquiry and achievement?

Figure 4: Research question #3 – Effects of moderated mediation



Methodology

The data for this project were obtained from TIMSS 2003, which was conducted by the International Association for the Evaluation of Educational Achievement (IEA). The TIMSS study offers an unprecedented opportunity to explore issues with a nationally representative sample, not only on student achievement in mathematics and science, but also on various contexts related to curriculum, schools, teachers, classrooms and students.

The following sections explain the TIMSS dataset as well as my intended data analysis method/procedure.

The TIMSS dataset

IEA is an international research association aimed to provide high-quality data for educational evaluation. Members of the IEA are top educational research institutions from participating countries in Africa, Asia, Australasia, Europe, Middle East, North Africa, and the Americas. Among the various international datasets collected by the IEA, the most recent one in science education was TIMSS 2003. TIMSS represents the largest and most ambitious international comparative educational study ever undertaken by IEA. It provides a rich array of information on achievement and the context in which learning occurs. As described by Schmidt et al. (1997), TIMSS is "intended to provide educators and policy makers with unparalleled multidimensional perspectives on mathematics and science curricula; their implementation; the nature of student performance in mathematics and science; and the social, economic, and educational context in which they occur" (p. ix).

Participants. A total of 8,912 8th graders (4,283 boys and 4,629 girls) participated in TIMSS 2003 in the United States. Their average age was 14.2 years at the time of assessment. These students completed questionnaires on home background, prior experiences and attitudes related to science learning.

Instruments. Data for this study came from two types of instruments: the achievement test booklets and the student questionnaires. Analyses conducted here were focused on the 8th graders' composite science scores in the achievement test and their relationships with the selected variables in the student questionnaire.

Achievement Test. The TIMSS assessment was designed so that no students answered all questions. It consisted of approximately 400 mathematics and science items assembled in 12 booklets, with 28 blocks of items rotated among those 12 booklets. Each student was assigned only one booklet, such that a representative sample of students answered each item. At grade eight, 90 minutes were allowed for this test.

Approximately 75% of the items were in multiple-choice and 25% were constructed-responses. Correct responses in multiple-choice items were award one point each, while constructed-response items could have partial credits with fully correct answers being awarded two points.

Plausible Values. The plausible values (PVs) were generated in two stages. In the first stage, students' achievement scores were estimated from their item responses based on a 3-parameters logistic Item Response Theory (IRT) model. IRT enabled achievement scores of the students to be summarized on a common scale even when different students were taking different tests and no one was tested on every item. These achievement scores were actually latent parameter estimates that reflected how students would have performed on the test had all the items administered to all of them. In the second stage, utilizing multiple imputations those achievement scores from IRT were combined with a large array of background characteristics about the students to generate a distribution of possible scores for each student. Although slightly different, we could think of this ability distribution as the dependent variable in a regression where the independent variables were students' item responses and their background characteristics. From each student's ability distribution, five random draws were made and each of these random draws constituted a PV. In this study, the TIMSS 8^{th} graders' five science PVs were used to represent their science achievement.

Student Questionnaire. The TIMSS 2003 questionnaires collected a wide array of data on educational context of student achievement. In the student questionnaire, students answered questions pertaining to home background factors, their activities and attitudes, their academic self-concepts, and so on. The present study made use of raw data available from the student questionnaire in the TIMSS 2003 dataset.

Sampling Weights. Since it was impossible to test everyone on everything, sampling was necessary. As such, sampling weights were needed in analyses to adjust for non-responses so that the results from the sample could be generalized to the population. There were three types of student weights (i.e., total weight, house weight, and senate weight) in TIMSS. For national analyses, all three types of student weights would produce the same results. In the present study, house weight would be used since it was expected to sum to the sample size.

Validity of the test. TIMSS 2003 is perhaps the largest and most comprehensive international study. It involved 49 countries over the world. The instruments used in TIMSS 2003 were carefully developed by international experts in education.

Through concerted efforts, all participating countries submitted items that were reviewed by subject-matter experts. Additional items were also written to ensure adequate coverage of science topics. Items were pilot tested, their results were reviewed, and then new items were written and pilot tested (TIMSS International Science Report, 2004). For TIMSS 2003, half of the items were newly developed and half were from

TIMSS 1995 and TIMSS 1999 for measuring trend. All new items were reviewed by subject-matter experts and pilot tested in almost all participating countries.

The TIMSS science curriculum framework specified two dimensions: the science content domain and the science cognitive domain. Six science content domains were tested: (1) life science, (2) chemistry, (3) physics, (4) earth science, (5) environmental science, and (6) scientific inquiry and nature of science. And three science cognitive domains were integrated: (1) factual knowledge, (2) conceptual understanding, and (3) reasoning and analysis.

Reliability. The TIMSS 2003 assessment was designed in 12 booklets that contained mathematics and science questions. Each student was tested in one booklet only. Although questions might vary from booklet to booklet, they were designed with similar format, content and level of difficulty.

Since each student was administered a portion of the items, TIMSS made use of the plausible values. Plausible values represented estimates of how students would perform on a test that composed of the entire item pools. They constituted random draws from each student's ability distribution. The variance between these plausible values primarily reflected the imputation error.

Because of the tremendous complexity of the TIMSS instruments and data sets, raw scores for individual items in science achievement would not be used in the analyses of this study. Instead, I will use the five plausible values generated in the TIMSS dataset. Using SPSS 14, I had computed Cronbach's alpha reliability coefficients for science achievement, inquiry-based activities, and attitudes toward science. The reliability statistics are presented in Table 1. These alpha coefficients were all sufficiently high, especially for the Science Achievement scale.

Table 1: Reliability coefficient for each subscale

Latent Construct	Chronbach's alpha
Science achievement	0.979
Inquiry-based activities	0.860
Attitudes toward science	0.846

Data Analysis Method

For the purpose of this study, selected variables from TIMSS 2003 would be used for analyzing the hypothesized structural models (see Table 2). The latent causal variable, inquiry-based activities, was created using four observed scores: (1) make hypothesis or prediction, (2) plan experiment or investigation, (3) conduct experiment or investigation, and (4) work in group or investigate science. The scale of these four observed inquiry variables had been recoded from the original TIMSS 2003 dataset such that 0 = "never", 1 = "some lessons", 2 = "about half the lessons", and 3 = "every or almost every lesson." The latent mediator variable attitudes toward science was created similarly using the following five observed variables: (1) enjoy learning science, (2) like to take more science, (3) would help in daily life, (4) need science to learn, and (5) like science jobs. Again, the scales of these five observed attitudes variables had also been recoded such that 0 = "disagree a lot", 1 = "disagree a little", 2 = "agree a little", and 3 = "agree a lot."

The latent outcome variable science achievement was composed of the five science plausible values. Lastly, the moderator gender was an observed variable such that 0 = boy and 1 = girl. Table 2 lists the characteristics of all the variables in the hypothesized models. Table 3 presents descriptive statistics of all the indicator variables.

Table 2: Characteristics of all the variables in the models.

Variable name	Variable type	Variable description
Science achievement (ACH)	Latent	Overall indicator of science achievement
1st PV in science (PV1)	Latent	1st random draw from examinee's science ability distribution
2nd PV in science (PV2)	Latent	2nd random draw from examinee's science ability distribution
3rd PV in science (PV3)	Latent	3rd random draw from examinee's science ability distribution
4th PV in science (PV4)	Latent	4th random draw from examinee's science ability distribution
5th PV in science (PV5)	Latent	5th random draw from examinee's science ability distribution
Learning through inquiry (INQ)	Latent	Overall indicator of inquiry learning
Make hypothesis or prediction (x1r)	Observed	How often ask/make hypothesis or prediction
Plan experiment or investigation (x2r)	Observed	How often ask/plan experiment or investigation
Conduct experiment or investigation (x3r)	Observed	How often ask/conduct experiment or investigation
Work in group or investigate science (x4r)	Observed	How often ask/work in group or investigate science

Attitudes toward science (ATT)	Latent	Overall indicator of attitude toward science
Like to take more science (y1r)	Observed	I would like to take more science
Enjoy learning science (y2r)	Observed	I enjoy learning science
Science helps in daily life (y3r)	Observed	Science would help me in daily life
Need science to learn (y4r)	Observed	I need science to learn other subjects
Like science jobs (y5r)	Observed	I would like job investigating science
Gender (GEN)	Observed	What is your gender?

Note: PV represents "Plausible Value."

Table 3: Descriptive statistics of indicator variables (Overall N=8,912)

Variables	Overall Means (SD)
Y1R	1.733 (1.034)
Y2R	1.957(0.980)
Y3R	1.981(0.886)
Y4R	1.735(0.925)
Y5R	1.498 (1.083)
X1R	1.659 (0.927)
X2R	1.522 (0.958)
X3R	1.682 (0.951)

X4R	1.900 (0.965)
PV1	526.972 (80.587)
PV2	526.733 (81.284)
PV3	527.749 (79.965)
PV4	527.884 (81.107)
PV5	527.153 (80.451)

Structural Equation Modeling. The structural equation modeling approach was used to analyze data in this study. It is an extremely influential technique which grew from the general linear model. It serves the purpose of many analytic techniques such as multiple regression, path analysis, factor analysis, time series analysis, survival analysis, multilevel analysis, latent class analysis, IRT modeling, and so on, but in a much more powerful way. It can be applied to cross-sectional data, longitudinal data, experimental data, non-experimental data, and quasi-experimental data. It has more flexible assumptions comparing to multiple regressions. It even allows for interpretation in the presence multi-collinearity. It takes into account measurement error, correlated error, correlated IVs, nonlinearity, interaction, etc. Most of all, it helps me to think clearly about causality.

In structural equation modeling, models can be drawn as path diagrams. The manifest or observed variables are signified by rectangular or square boxes, where the latent or unobserved variables are signified by circles or ovals. Straight arrow represents the assumption that the variable at the base of the arrow ‘causes’ the variable at the head of the arrow. Unenclosed variables represent a disturbance term (i.e., variation that can not be explained by the model).

Since latent variables are created variables, they need to be defined by a scale (Kline, 1998). This can be done by fixing the path from the latent factor to one of its observed variable to 1, or by fixing the factor variance to 1 (Bollen, 1988). For example, the path from attitudes toward science to like to take more science can be fixed to 1, or the variance of attitudes toward science can be fixed to 1.

Moderated Mediation. Mediation has been of great interest to experimental as well as non-experimental psychologists. It allows investigators to elucidate the intervening mechanism by which an independent variable affects the dependent variable. Likewise, moderation has also been very popular and it allows researchers to examine whether certain variables affect the magnitude and direction of the treatment effect. Therefore, as of today there exist lots of studies that focused on mediation or moderation. Occasionally, there are also studies that combined the processes of mediation and moderation. By combining these two processes into either mediated moderation or moderated mediation, researchers can often gain a great deal of useful information beyond simple mediation and simple moderation.

"Moderated mediation happens if the mediating process that is responsible for producing the effect of the treatment on the outcome depends on the value of a moderator." (Muller, Judd, & Yzerbyt, 2005, p. 8) That is to say, the mediating process differs as a function of individual differences or contextual differences. It implies the existence of full or partial mediation to begin with. More importantly, moderated mediation requires that "there is an overall treatment effect and the magnitude of this effect does not depend on the moderator. However, the potency of the mediating process depends on the moderator." (Muller, Judd, & Yzerbyt, 2005, p. 12) Therefore, in the present study of moderated mediation, I would first need to test whether there is an overall effect of inquiry instruction on science achievement. If it is yes, then I would test whether attitudes toward science is a mediator between inquiry and science achievement. If it is yes again, my next step would be testing whether gender is a moderator of the overall effect of inquiry on achievement. A non-significant overall moderation of the inquiry effect is the key that would allow us to say that our moderated mediation model is indeed a moderated mediation model, but not mediated moderation model. The final step is, of course, to test whether there is a significant moderated mediation effect.

In contrast, mediated moderation can only happen when there is moderation of the overall treatment effect (Muller, Judd, & Yzerbyt, 2005). This enables us to examine whether "the overall moderation of a treatment effect is reduced once the mediating process is controlled." (Muller, Judd, & Yzerbyt, 2005, p. 33) Intuitively, both the mediated moderation and the moderated mediation models appear the same (and indeed they are the same analytically), yet there is a clear distinction between them.

In the present study, I analyzed the hypothesized structural models by using a specialized structural equation modeling software MPlus. Personally, I think MPlus is quite user friendly and it contains many special features such as it accommodates missing data, sampling weights, Monte Carlo simulation, maximum likelihood estimation for all outcome types and multiple group analysis. It utilizes powerful methods to deal with missing data. Its modeling framework allows for analyzing a combination of continuous and categorical latent variables.

To test the hypothesized model, I first started by verifying the validity of the latent subscales by checking the correlations of each of their observed variables with other related variables and checking the loadings of the indicators on their respective latent constructs. Since structural equation models assumed population data and would not allow for missing data, I employed full maximum information likelihood to impute all of the missing data.

The following is an outline summarizing the basic procedure that I followed in conducting the analysis:

1. Hypothesize structural models specifying the relationships among the latent variables.
2. Create measurement models identifying how the latent variables are to be measured in terms of observed variables.
3. Put together the full structural equation models by combining the structural and measurement models.
4. Data extraction
   - ➢ Use the International Database Analyzer to extract only the U.S., 8th graders' data from TIMSS 2003.
5. Data cleaning
   - ➢ Visually inspect the dataset looking for rare patterns, etc., and recode variables if necessary.
   - ➢ Check for correlations, normality, missing data.

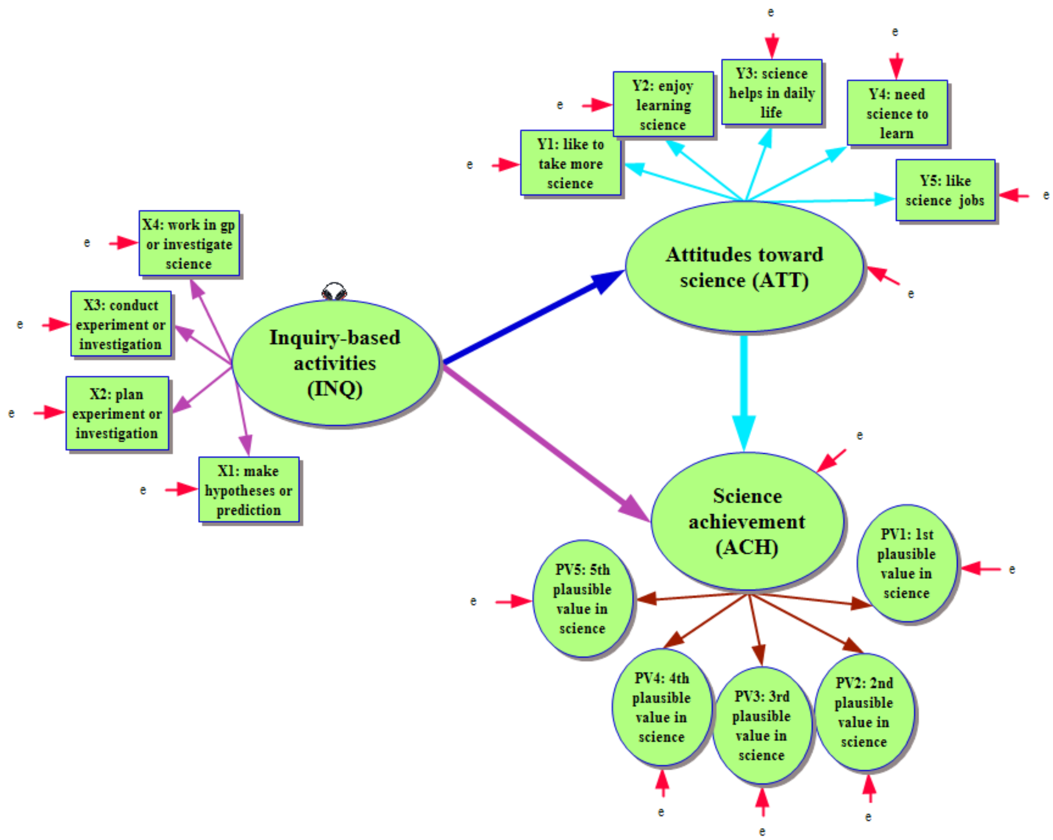
- If multivariate normal, then use maximum likelihood estimator to test the models later on.
- If not multivariate normal, then use a maximum likelihood estimator that's robust to non-normality and non-independence of observations.
- Use full information maximum likelihood to replace missing data.

6. Conduct a confirmatory factor analysis on the measurement model of inquiry-based activities.
    - Examine the goodness-of-fit statistics.
    - Check the correlations and loadings of each of the observed variables with the latent variable.
    - Modify the model if the fit is inadequate.
7. Conduct a confirmatory factor analysis on the measurement model of attitudes toward science.
    - Examine the goodness-of-fit statistics.
    - Check the correlations and loadings of each of the observed variables with the latent variable.
    - Modify the model if the fit is inadequate.
8. Conduct a confirmatory factor analysis on the measurement model of science achievement (treat each PV as an observed variable for the purpose of coding).
    - Examine the goodness-of-fit statistics.
    - Check the correlations and loadings of each of the observed variables with the latent variable.
    - Modify the model if the fit is inadequate.

9. Test moderated mediation using full structural equation modeling by combining the measurement and the structural models. The steps of conducting the tests are detailed from (a) to (d).
   a) Test for the total effect from inquiry to achievement. Examine for identifiability and goodness-of-fit statistics.
   b) If the total effect from (a) is significant, test for the mediation effect from inquiry to achievement with attitudes toward science as the mediator. Examine for identifiability and goodness-of-fit statistics.
   c) Test for the moderation effect from inquiry to achievement with attitudes toward school as the moderator. Examine for identifiability and goodness-of-fit statistics.
   d) If the mediation effect from (a) is significant and the moderation effect from (c) is non-significant, test for moderated mediation from inquiry to achievement with attitudes toward science as the mediator and the attitudes toward school as the moderator. Examine identifiability and goodness-of-fit statistics.

The following is a diagram of the hypothesized full structural equation model (Figure 5).

Figure 5: Overview of the full inquiry model of science achievement with attitudes toward science as the mediator



Results

The hypothesized model consisted of Inquiry-based activities leading to Attitudes toward science and Science achievement. Attitudes toward science was expected to lead to Science achievement. In addition, differential effects of inquiry-based activities and attitudes toward science were also expected, depending on students' gender. This model was tested using the structural equation modeling approach. Table 4 contains the pearson product moment correlations among the items. Each item correlated reasonably high with the others within their subscales.

Table 4: Observed bivariate correlations among variables (Overall N = 8,912)

	PV1	PV2	PV3	PV4	PV5
PV1	1.000				
PV2	0.903	1.000			
PV3	0.903	0.902	1.000		
PV4	0.901	0.903	0.902	1.000	
PV5	0.903	0.900	0.901	0.900	1.000
Y1R	0.189	0.177	0.188	0.182	0.179
Y2R	0.148	0.145	0.151	0.147	0.143
Y3R	0.106	0.104	0.108	0.102	0.101
Y4R	0.076	0.063	0.071	0.066	0.061
Y5R	0.216	0.205	0.212	0.209	0.205
X1R	-0.028	-0.026	-0.022	-0.018	-0.027
X2R	-0.056	-0.050	-0.045	-0.046	-0.049
X3R	0.084	0.085	0.085	0.090	0.085
X4R	0.080	0.087	0.090	0.091	0.086
SEX	-0.104	-0.097	-0.104	-0.101	-0.104

	Y1R	Y2R	Y3R	Y4R	Y5R
Y1R	1.000				
Y2R	0.709	1.000			
Y3R	0.511	0.507	1.000		
Y4R	0.438	0.427	0.626	1.000	
Y5R	0.557	0.520	0.528	0.476	1.000
X1R	0.153	0.176	0.196	0.196	0.170
X2R	0.153	0.164	0.198	0.208	0.156
X3R	0.156	0.171	0.192	0.183	0.166
X4R	0.129	0.151	0.154	0.143	0.125
SEX	-0.060	-0.081	-0.016	-0.034	-0.056

	X1R	X2R	X3R	X4R	SEX
X1R	1.000				
X2R	0.648	1.000			
X3R	0.611	0.671	1.000		
X4R	0.516	0.533	0.657	1.000	
SEX	-0.032	-0.064	-0.024	0.021	1.000

Loadings of indicator variables on their latent factors and latent factor univaiate statistics as well as bivariate correlations are presented in Table 5. Variables were ordered and grouped by their underlying latent constructs to facilitate interpretation. Lambdas represented the loadings of each indicator variable on the underlying construct. Phi's represented the correlations between the latent constructs. As we can see, factor correlations varied from very low to moderate and were all positive here. In addition, all loadings were positive and substantially large (>0.60), providing support that all the measurement models fit quite well.

Table 5: Analysis of TIMSS 2003 Data Structures (N = 8,912)

	Λ (lambdas)		
	INQ	ATT	ACH
x1r	0.750	0	0
x2r	0.799	0	0
x3r	0.850	0	0
x4r	0.719	0	0
y1r	0	0.795	0
y2r	0	0.778	0
y3r	0	0.715	0
y4r	0	0.639	0
y5r	0	0.709	0
pv1	0	0	0.951
pv2	0	0	0.950
pv3	0	0	0.950
pv4	0	0	0.949
pv5	0	0	0.949

Φ (phis)

	INQ	ATT	ACH
INQ	1.000		
ATT	0.284	1.000	
ACH	0.038	0.211	1.000

Univariate statistics

	INQ	ATT	ACH
Mean	0.000	0.000	0.000
Std dev	0.696	0.822	76.621

Since all three of the measurement models were over-identified (df ≥ 1), the measure of fit was obtained. Table 6 contains fit indices for the measurement models. Both the measurement models INQ and ATT showed significant chi-square values, suggesting poor fit. However, the chi-square goodness of fit is really not trustworthy here under large sample size (N). As N increases, power also increases and hence we are more likely to see the model is wrong even if it is not. On the contrary, the Tucker-Lewis Index (TLI) and the Standardized Root Mean Square Residual (SRMR) are more representative. TLI provides unbiased estimate of model fit regardless of sample sizes. SRMR is a stand alone index that has simple interpretative meaning and does not penalize the model for the lack of parsimony (Hu & Bentler, 1997). Therefore we should rely on TLI and SRMR to evaluate model fit here. As we can see, all three of the measurement models fit reasonably well. TLI for INQ and ACH exceeded 0.9, indicating extremely good fit. SRMR for INQ and ACH were all less than 0.05, also indicating excellent fit. ATT had a

TLI of 0.778 and a SRMR of <0.08, it too was considered adequate. Taken together, these results suggest that the measurement models were properly specified, and it was reasonable to go forward to examine the full structural equation models.

Table 6: Fit Indices for the measurement models (N = 8,912)

Model	$\chi 2$	df	p	SRMR	TLI
Measurement Models					
INQ	408.716	2	0.000	0.023	0.913
ATT	1865.425	5	0.000	0.052	0.778
ACH	5.061	5	0.5641	0.001	1.000

Table 7 presents fit indices of the full structural equation models. All seven of the full SEM models seemed to fit relatively well, with SRMR < 0.08 and TLI > 0.90. Again, we would disregard the chi-square statistic here due to its sensitivity to large sample size. Among the five moderated mediation models, the unrestricted model that allowed differential mediated effect on all paths seemed to fit the best (SRMR = 0.045, TLI = 0.967).

Table 7: Fit Indices for the full structural models (N = 8,912)

Model	$\chi 2$	df	p	SRMR	TLI
1. INQ → ACH	754.963	26	0.000	0.040	0.987
2. mediated INQ → ACH	2987.295	74	0.000	0.042	0.964

3. moderated mediation INQ → ACH (restricted)	3283.309	173	0.000	0.049	0.967
4. moderated mediation INQ → ACH (freed INQ → ATT)	3260.181	172	0.000	0.047	0.967
5. moderated mediation INQ → ACH (freed ATT → ACH)	3276.424	172	0.000	0.048	0.969
6. moderated mediation INQ → ACH (freed INQ → ACH)	3277.994	172	0.000	0.050	0.967
7. moderated mediation INQ → ACH (unrestricted)	3240.154	170	0.000	0.045	0.967

There was a significant overall total effect (see Table 8) of inquiry on science achievement ($\beta = 0.040$, $p < 0.05$). The direct effect between inquiry and attitudes toward science ($\beta = 0.284$, $p < 0.05$) in the first component of the mediated model as well as the direct effect between attitudes toward science and achievement ($\beta = 0.218$, $p < 0.05$) in the last component of the mediated were also significant (see Tables 9 & 10). That means, the influence of attitudes toward science could potentially be a mediated process.

Table 11 contained the mediation estimates through attitudes toward science. The Sobel Test for the indirect effect ($\beta=0.062$; $b=6.838$, $p < 0.05$) was shown to be statistically significant with a critical ratio of 12.970. Taken together, these results were all in favor of the argument for mediation. After controlling for the mediator students' attitudes toward science, the direct effect of inquiry on science dropped to non-significant (β = -0.025, $p > 0.05$). This indicated that the effect of inquiry on achievement was consistently and fully mediated by students' attitudes toward science.

Table 8: Total effects from INQ to ACH. All relationships are standardized regression coefficients. Numbers in parentheses represent unstandardized regression coefficients. (N=8,912)

		CAUSE
		INQ
EFFECT	ACH	0.040 (4.437)*

Note: *p<0.05

Table 9: Direct effects between exogenous and endogenous variables (Gamma Matrix). Rows depict effects and columns depict causal variables. All relationships are standardized regression coefficients. Numbers in parentheses represent unstandardized regression coefficients. (N=8,912)

		CAUSE
		INQ
EFECT	ATT	0.284(0.336)*
	ACH	-0.025(-2.701)

Note: *p<0.05

Table 10: Direct effects between endogenous variable and itself (Beta Matrix). Rows depict effects and columns depict causal variables. All relationships are standardized regression coefficients. Numbers in parentheses represent unstandardized regression coefficients. (N=8,912)

		CAUSE	
		ATT	ACH
EFFECT	ATT	0	0
	ACH	0.218(20.314)*	0

Note: *p<0.05

Table 11: Indirect effects from INQ to ACH. All relationships are standardized regression coefficients. Numbers in parentheses represent unstandardized regression coefficients. (N=8,912)

Indirect	
ACH	
ATT	
INQ	0.062 (6.838)*
Sobel Test	
Critical Ratio	12.970*

Note: *p<0.05

The following diagrams were graphical representation of Tables 8-11 (Figure 6) and Tables 5, & 8-11 (Figure 7). They provide alternative views for appreciating the model results.

Figure 6: Path Diagram of inquiry-based activities on science achievement. Standardized regression coefficients are shown with unstandardized regression coefficients in parenthesis (*p<0.05)

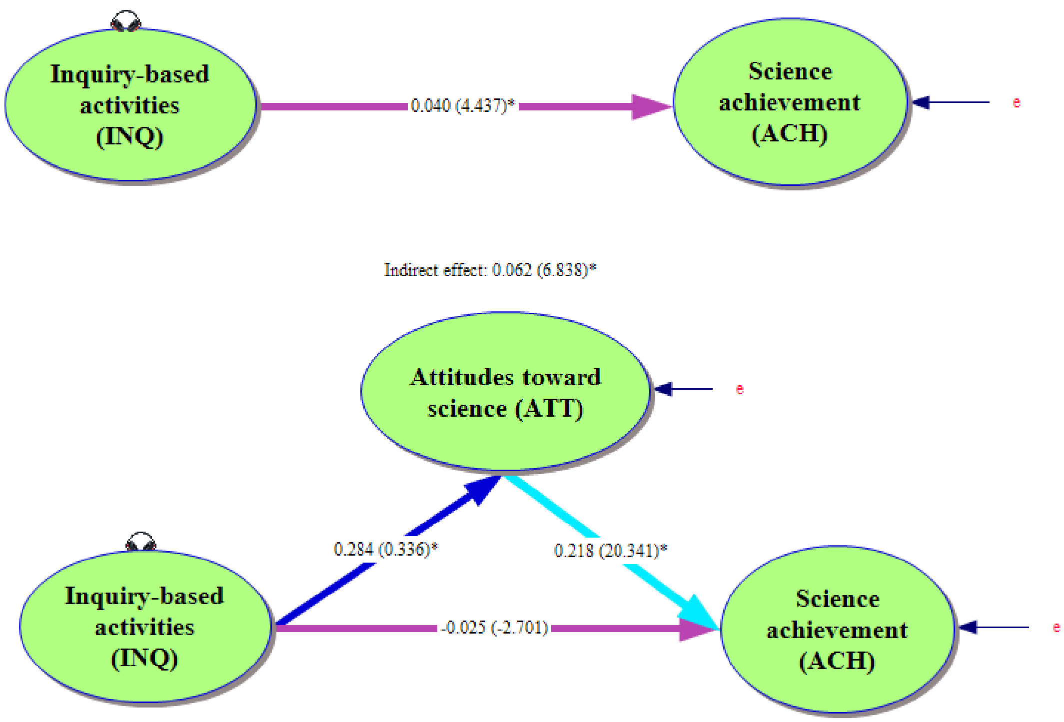
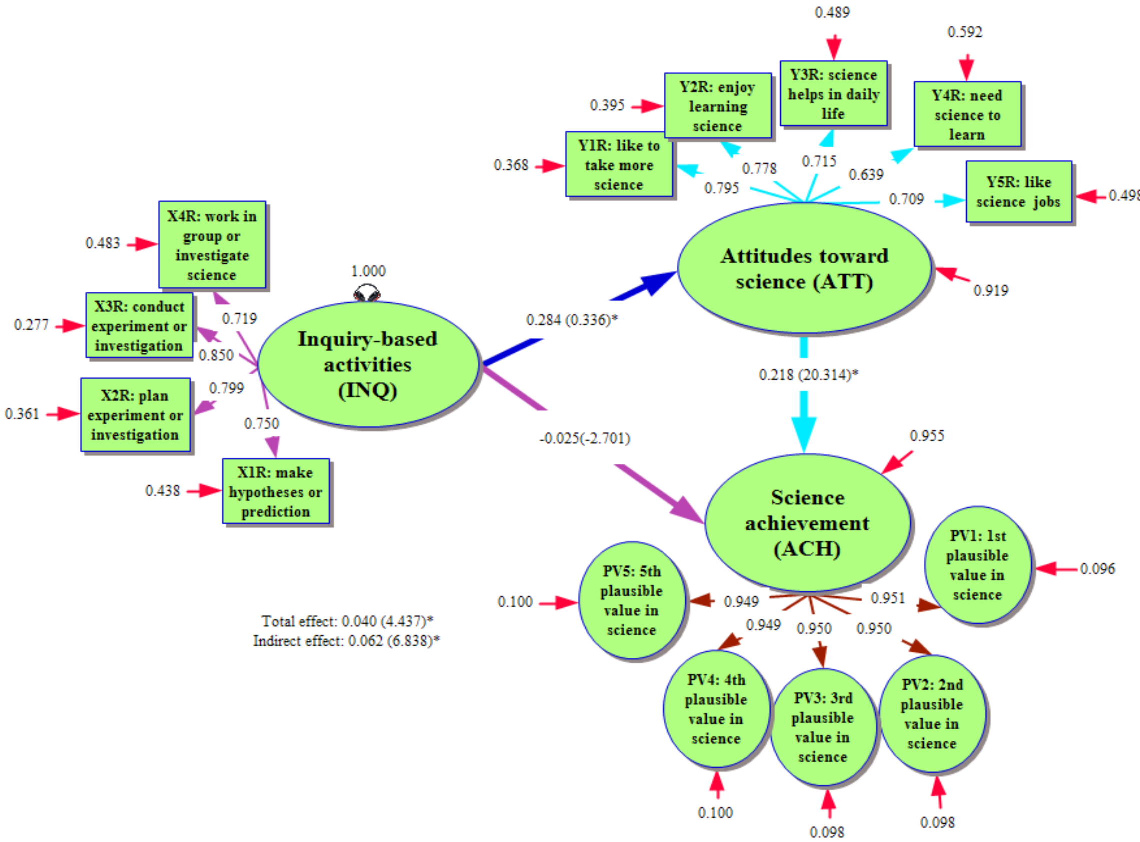


Figure 7: Inquiry model of science achievement with attitudes toward science as the mediator (*p<0.05)



In order to rule out the possibility of mediated moderation, we would have to confirm that the magnitudes of the overall effect of inquiry on science achievement did not depend on the moderator – gender. This was done by constraining all paths to be equal in both groups (baseline model), and then released the paths to see if boys and girls differed. The results in Table 12 indicate that the restricted baseline model did not differ significantly from the unrestricted model ($\Delta\chi2 = 2.464$, $\Delta df = 1$, $p > 0.05$), implying that the overall effects of inquiry on achievement were unmoderated by gender. Hence, mediated moderation was unlikely to happen here.

Table 12: Stacked model analyses testing differential gender overall moderation effect of INQ on ACH (Overall N=8,912).

Model	χ2 (df)	Δ χ2 (Δdf)
1. Baseline model (all paths constrained to be equal)	901.848 (67)	---
2. INQ → ACH (unrestricted)	899.384 (66)	2.464 (1) n.s.

Once we had ruled out the possibility of mediated moderation, the next step was to check whether moderated mediation was plausible. First, a series of stacked models was used to test the moderated mediation hypothesis. Then, chi-square difference tests were conducted to see if the restricted baseline model differed from any of the less restricted models.

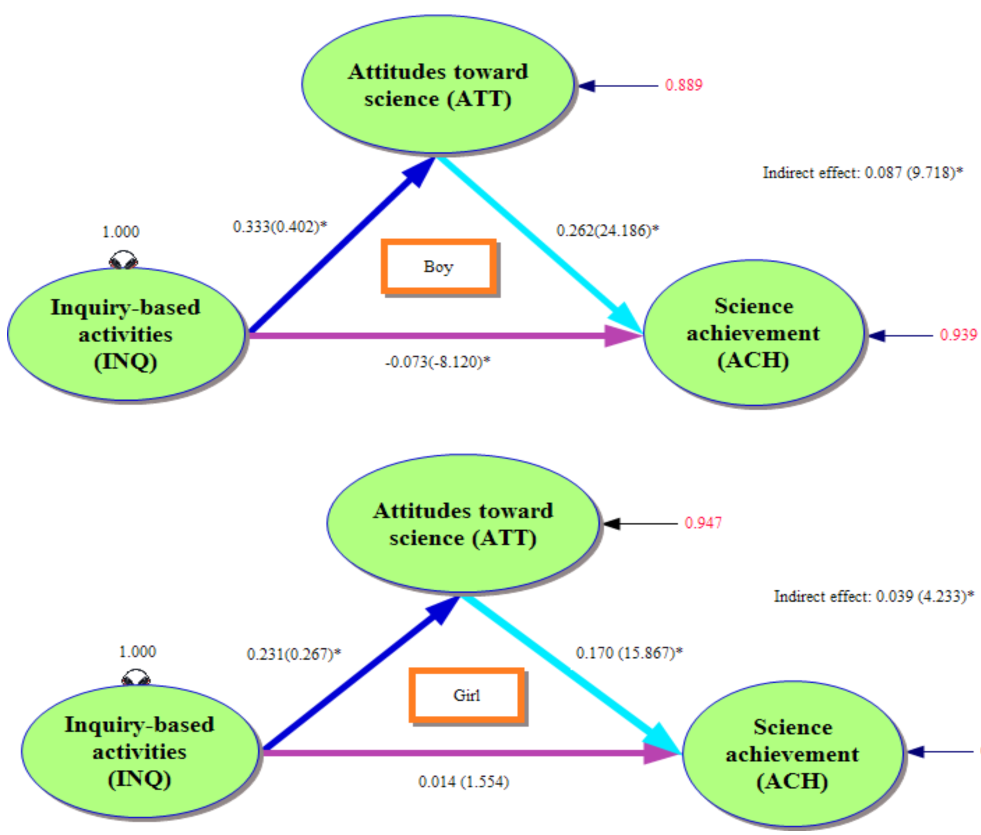
Table 13: Stacked model analyses testing differential gender mediated effects of INQ on ACH (Overall N=8,912).

Model	χ2 (df)	Δ χ2 (Δdf)
1. Baseline model (all paths constrained to be equal)	3283.309 (173)	---
2. INQ → ATT (freed)	3260.181 (172)	23.128*(1)
3. ATT → ACH (freed)	3276.424 (172)	6.885*(1)
4. INQ → ACH (freed)	3277.994 (172)	5.315*(1)
5. Unrestricted model (all paths freed)	3240.154 (170)	43.155*(3)

Note: *p<0.05

Given the significant chi-square difference test results of the effect from inquiry to attitudes toward science depend on the moderator gender ($\Delta\chi 2 = 23.128$, $\Delta df = 1$, $p < 0.05$) as well as the effect from attitudes toward science to achievement depend on the moderator gender ($\Delta\chi 2 = 6.885$, $\Delta df = 1$, $p < 0.05$) and no overall moderation of inquiry on achievement ($\Delta\chi 2 = 2.464$, $\Delta df = 1$, $p > 0.05$), we could conclude that the mediated effect of inquiry on achievement, was moderated by gender. In fact, attitudes toward science mediated the relationship of inquiry-based activities and science achievement in boys (Indirect effect: $\beta = 0.087$; $b = 9.718$, $p < 0.05$) much stronger than in girls (Indirect effect: $\beta = 0.039$; $b = 4.233$, $p < 0.05$).

Figure 8: Differential mediated effect of inquiry on science achievement (*p<0.05)



Discussion

The purposes of this study were three fold: (1) to examine the overall effects of inquiry on science achievement, (2) to study the mediating role of science attitudes on the relation between inquiry and science achievement, and (3) to see if this relationship was moderated by gender. Results indicated that there was an overall inquiry effect on science achievement, fully mediated by students' attitudes toward science. Moreover, the mediated effects were much stronger for boys than for girls. After controlling for attitudes toward science, inquiry activities increased science achievement more for boys than for girls.

These results were consistent with findings from various researchers that inquiry-based activities lead to increased science achievement (Amaral, et al., 2002; Change & Mao, 1998; Jakupcak, et al., 1996; Jorgenson & Vanosdall, 2002; Luckie, et al., 2004; Schneider, et al., 2002; Von Secker, 2002), and that students developed more positive attitudes toward science while engaging in inquiry-based activities (Alouf & Bentley, 2003, Berg, et al., 2003; Booth, 2001; DiPasquale, et al., 2003; Von Secker, 2002). What was surprising were the full mediated relationships found among inquiry, attitudes toward science and science achievement.

This study was the first to investigate a theoretical moderated mediation model of inquiry on science achievement. As such, this study added to previous studies from numerous researchers by examining mediation and moderated mediation effects of inquiry, gender, attitudes toward science, and achievement. Settlage (2003) indicated that a major reason behind the failure of teachers in implementing inquiry instruction was the lack of evidence regarding how to teach using inquiry. Now, this study had provided direct empirical evidence for enhancing inquiry instruction (i.e., boosting students' attitudes toward science).

Unclear at this point are the underlying mechanisms and strategies that can alter students' attitudes toward science. Further research should direct efforts toward examining that. As with any studies, cross validations and replications are extremely important. Future studies can be designed by using different and/or multiple measures of the constructs as well as examining a different grade level sample to see if the same results still hold. Deeper understanding of the effects of inquiry on science achievement can potentially direct curricular and policy changes to support science learning. Additionally, inquiry learning seemed to explain only a small amount of variance in science achievement. It is recommended that other factors should also be placed in the models to see which factor or which combination of factors could explain the most variance in science achievement. Lastly, more studies are needed to see if the results of this study hold similarly across different countries in the world.

The findings of this study also need to be considered in light of some limitations. First of all, the data came from an existing dataset. I had no control over coding mistakes, missing values and inaccurate derived variables. Secondly, development of the models was constrained to variables contained in the dataset. Future studies would benefit from using a different dataset to assess the hypothesized models. Third, because the data for the present study were cross-sectional, different causal interpretations might have fit the data equally well. We could only conclude that the latent constructs being studied covaried, not necessarily one construct caused the other. If and only if we had longitudinal data or a time machine that could rewind and forward to any time, we would be able to disentangle causal relationships. Lastly, since this was the first study in examining a moderated mediation model among inquiry, gender, attitudes toward science, and science achievement, it was far from developing a consistent theoretical model. Hence, this study should be considered exploratory.

Nonetheless, the research results here have practical significance for improving the quality of science teaching and learning. They provide ideas for policy makers as well as teachers to design educational programs to enhance science learning outcomes. What seems apparent is that in order to help students learning science through inquiry, teachers should work on strengthening their students' attitudes toward science. Negative attitudes toward science can be modified (Hembree, 1990) and by the same token, positive attitudes toward science can be enhanced too. Therefore, it is critical for teachers to plan lessons that excite students' interest toward science rather than lessons that embrace scientific concepts that are too difficult which drive students away. It is only until students themselves realize that science would help them in their daily life, or realize that science is really entertaining, their scientific interest would spark. Indeed, cultivating strong interest toward science is really the key for leading the pathway from inquiry to high levels of science achievement!

References

Abd-El-Khalick, F., & BouJaoude, S. (1997). An exploratory study of the knowledge base for science teaching. *Journal of Research in Science Teaching, 34*(7), 673-699.

Aldrige, B. G. (1992a). Essential changes in secondary science: Scope, sequence, and coordination. In M. K. Pearsall (Ed.), *Scope, sequence, and coordination of secondary school science, Volume 2.* (Washington, DC: National Science Teachers Association), 1-7.

Aldrige, B. G. (1992b). Project on scope, sequence, and coordination: A new synthesis for improving science education. In M. K. Pearsall (Ed.), *Scope, sequence, and coordination of secondary school science, Volume 2*. (Washington, DC: National Science Teachers Association), 1-7.

Alouf, J. L., & Bentley, M. L. (2003). *Assessing the impact of inquiry-based science teaching in professional development activities, PK-12.* Paper presented at the annual meeting of the Association of Teacher Educators, Jacksonville, FL. (ERIC Document Reproduction Service No. ED475577)

Amaral, O.M., Garrison, L., & Klentschy, M. (2002). Helping English learners increase achievement through inquiry-based science instruction. *Bilingual Research Journal*, *26*(2), 213-239.

American Association for the Advancement of Science. (1990). *Project 2061: Science for all American*. New York: Oxford University Press.

American Association for the Advancement of Science. (1993). *Project 2061: Benchmarks for science literacy*. New York: Oxford University Press.

Bandura, A. (2001). Social cognitive theory: An agentive perspective. *Annual Review of Psychology, 52*, 1-26.

Bates, G. C. (1978). The role of the laboratory in secondary school science programs. In M. B. Rowe (Ed.), *What Research says to the science teacher, Volume 1.* (Washington, DC: National Science Teachers Association), 55-82.

Berg, C. A. R., Bergendahl, V. C. B., & Lundberg, B. K. S. (2003). Benefiting from an open-ended experiment? A comparison of attitudes to, and outcomes of, an expository versus an open-inquiry version of the same experiment. *International Journal of Science Education, 25*(3), 351-372.

Bloom, B. S. (1976). *Human characteristics and school learning.* (New York: McGraw-Hill).

Blosser, P. E. (1984). *Attitude research in science education*, (Information bulletin No. 1). Columbus, OH: Eric Clearinghouse for Science, Mathematics, and Environmental Education. (ERIC Document Reproduction Service No. ED 259941).

Bollen, K. A. (1988). *Structural equations with latent variables.* (John Wiley & Sons, Inc).

Booth, G. (2001). Is inquiry the answer? *The Science Teacher, 68*(7), 57-59.

Bredderman, T. (1983). Effects of activity-based elementary science on student outcomes: A quantitative synthesis. *Review of Educational Research, 53*, 499-518.

Brunkhorst, B. J. (1992). A study of student outcomes and teacher characteristics in exemplary middle and junior high science programs. *Journal of Research in Science Teaching, 29*, 571-583.

Change, C., & Mao, S. (1998). *The effects of an inquiry-based instructional method on earth science students' achievement.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Diego, CA. (ERIC Document Reproduction No. ED418858)

DeBoer, G. E. (1991). *A history of ideas in science education: Implications for practice.* New York: Teachers College Press.

DiPasquale, D. M., Mason, C. L., & Kolkhorst, F. W. (2003). Exercise in inquiry: Critical thinking in an inquiry-based exercise physiology laboratory course. *Journal of College Science Teaching, 32*(6), 388-393.

Ebenezer, J. V., & Zoller, U. (1993). Grade 10 students' perceptions of and attitudes toward science teaching and school science. *Journal of Research in Science Teaching*, *30*, 175-186.

Eccles, J. (1984). Sex differences in achievement patterns. In R. A. Dienstbier (Series Ed.) & T. B. Sonderegger (Vol. Ed.), *Nebraska symposium on motivation:. Psychology and gender, Volume 32*. (Lincoln, NE: University of Nebraska Press), 97-132.

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.

Haladyna, T., & Shaughnessy, J. (1982). Attitude toward science: A quantitative synthesis. *Science Education, 66*, 547-563.

Hembree, R. (1990). The nature , effects and relief and mathematics anxiety. *Journal for Research in Mathematics Education, 21*, 33-46.

Hofstein, A., & Lunetta, V. N. (1982). The role of laboratory in science teaching: Neglected aspects of research. *Review of Educational Research, 52,* 210-217.

House, J. D. (1996). Student expectancies and academic self-concept as predictors of student achievement. *Journal of Psychology, 130,* 679-681.

Jakupcak, J., Rushton, R., Jakupcak, M., & Lundt, J. (1996). Inclusive education. The *Science Teacher, 63*(5), 40-43.

Jarvis, T., & Pell, P. (2005). Factors influencing elementary school children's attitudes toward science before, during, and after a visit to the UK National Science Centre. *Journal of Research in Science Teaching,* 42 (1), 53-83.

Jorgenson, O., & Vanosdall, R. (2002). The death of science? What we risk in our rush toward standardized testing and the three R's. *Phi Delta Kappan, 83*(8), 601-605.

Kahle, J. B. (1992). Why girls don't know. In M. K. Pearsall (Ed.), *Scope, sequence, and coordination of secondary school science, Volume 2.* (Washington, DC: National Science Teachers Association), 111-124.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* (New York: The Guilford Press).

Koballa, T. R. (1995). Children's attitudes toward learning science. In S. Glynn and R. Duit (Eds.), *Learning science in the schools: Research reforming practice.* (Mahwah, NJ: Erlbaum), 59-84.

Lederman, N. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*(4), 331-359.

Lee, V. E., & Burkam, D. T. (1996). Gender differences in middle grade science achievement: Subject domain, ability and course emphasis. *Science Education, 80,* 613-650.

Lord, T., & Orkwiszewski, T. (2006). Moving from didactic to inquiry-based instruction in a science laboratory. *The American Biology Teacher, 68,* 342-345.

Luckie, D. B., Maleszewki, J. J., Loznak, S. D., & Krha, M. (2004). Infusion of collaborative inquiry throughout a biology curriculum increases student learning: A four-year study of "Teams and Streams." *Advances in Physiology Education, 28*(4), 199-209.

Mattern, N., & Schau, C. (2002). Gender differences in science attitude-achievement relationships over time among white middle-school students. *Journal of Research in Science Teaching, 39*(4), 324-340.

McComas, W. F., & Olson, J. K. (1998). *The nature of science in International Science Education Standards Documents.* In McComas (Ed.), The nature of science in science education: Rationales and strategies. Kluwer Academic Publishers: The Netherlands, 41-52.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology, 89*, 852-863.

National Center for Education Statistics. (1992). *The 1990 science report card: NAEP;s assessment of fourth, eight, and twelfth graders* (NCES Report No. 92-064). Washington, DC: U. S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics.

National Research Council. (1985). *Mathematics, science, and technology education: A research agenda.* Washington, DC: National Academy Press.

National Research Council. (1996). *The national science education standards.* Washington, DC: National Academy Press.

National Research Council. (2000). *Inquiry and the national science education standards.* Washington, D. C.: National Academy Press.

National Science Teachers Association. (2000). *Earthquakes: A teacher's package for K-6.* Arlington, VA: Federal Emergency Management Agency, Washington, DC.

Oakes, J., & the RAND Corporation. (1990). *Opportunities, achievement, and choice: Women and minority students in science and mathematics*. In C. B. Cazden (Ed.), Review of Research in Education (Vol. 16, pp. 153-222). Washington, DC: American Educational Research Association.

Office of Educational Research and Improvement. (1994). *Regional Educational Laboratory Program*. Washington, DC: Programs for the Improvement of Practice.

Pajares, F. (2002). Self-efficacy belies in academic settings. *Review of Educational Research, 66*, 543-578.

Pajares, F., & Schunk, D. H. (2001). *Self-beliefs and school success: Self-efficacy, self-concept, and school achievement.* In R. Riding & S. Rayner (Eds.), Self-perception (pp. 239-266). London: Ablex Publishing.

Rennie, L. J., & Punch, K. F. (1991). The relationship between affect and achievement in science. *Journal of Research in Science Teaching, 28,* 193-209.

Roth, K. J. (1992). Science education: It's not enough to 'Do' or 'Relate'. In M. K.Pearsall (Ed.), *Scope, sequence, and coordination of secondary school science, Volume 2.* (Washington, DC: National Science Teachers Association), 151-164.

Schibeci, R. A. (1984). Attitudes to science: An update. *Studies in Science Education, 11,* 26-59.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). Many aims, many visions. *A cross-national investigation of curricular intentions in school science, Volume 2.* (Dordrecht: Kluwer Academic Publishers).

Schneider, R. M., Krajcik, J., Marx, R. W., & Soloway, E. (2002). Performance of students in project-based science classrooms on a national measure of science achievement. *Journal of Research in Science Teaching, 39*(5), 410-422.

Settlage, J. (2003). *Inquiry's allure and illusion: Why it remains just beyond our reach.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia.

Shymansky, J. A., Hedges, L. V., & Woodworth, G. (1990). A reassessment of the effects of inquiry-based science curricula of the 60's on student performance. *Journal of Research on Science Teaching, 27*, 127-144.

Shymansky, J. A., Kyle, W. C., & Alport, J. M. (1983). The effects of new science curricula on student performance. *Journal of Research in Science Teaching, 20,* 387-404.

Simpson, R. D., Koballa, T. R., Jr., Oliver, J. S., & Crawley, F. E. (1994). Research on the affective dimensions of science learning. In D. Gabel (Ed.), *Handbook of research on science teaching and learning.* (New York: Macmillan), 211-234.

Simpson, R. D., & Oliver, J. S. (1990). A summary of major influences on attitudes toward and achievement in science among adolescent students. *Science Education, 74*, 1-18.

Smith, E. E., & Anderson, C. W. (1984). Plants as producers: A case study of elementary science teaching. *Journal of Research in Science Teaching, 21*, 685-695.

Martin, M. O. (2004). *TIMSS 2003 International Science Report.* Chestnut Hill, MA: TIMSS & PIRLS International Study, Center, Boston College, Lynch School of Education.

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

Welch, W. W. (1981). Inquiry in school science. In N. C. Harms & R. E. Yager (Eds.), *What research says to the science teacher, Volume 3*. (Washington, DC: National Science Teachers Association), 53-72.

Yager, R. E. (1996). History of science/technology/society as reform in the United States. In R. E. Yager (Ed.), *Science/technology/society as reform in science education.* (Albany, NY: State University of New York Press), 3-15.

Ye, R., Wells, R. R., Talkmitt, S., & Ren, H. (1998). *Student attitudes toward science learning: A cross-national study of American and Chinese secondary school students.* Paper presented at the Annual Meeting of the National Science Teachers Association, Las Vegas, Nevada.