Summer 8-2010

Assessing the Effectiveness of Studio Physics at Georgia State University

Brianna M. Upton
Georgia State University

Recommended Citation

Upton, Brianna M., "Assessing the Effectiveness of Studio Physics at Georgia State University." Thesis, Georgia State University, 2010.
https://scholarworks.gsu.edu/phy_astr_theses/10

ASSESSING THE EFFECTIVENESS OF STUDIO PHYSICS AT GEORGIA STATE UNIVERSITY

by

BRIANNA M. UPTON

Under the Direction of Brian Thoms

ABSTRACT

Previous studies have shown that many students have misconceptions about basic concepts in physics which persist after instruction. It has been concluded that one of the challenges lies in the teaching methodology. To address this, Georgia State University (GSU) has begun teaching studio algebra-based physics. Although many institutions have implemented studio physics, most have done so in calculus-based sequences. Additionally, the unique environment of GSU's population as a diverse, urban research institution is considered. The effectiveness of the studio approach for this demographic in an algebra-based introductory physics course was assessed. This five-semester pilot study presents demographic survey results and compares the results of student pre- and post-tests using the Force Concept Inventory (FCI). FCI results show that 1) the studio approach yields higher learning gains than the conventional course, 2) there are significant performance differences among ethnic groups, and 3) a gender gaps exists regardless of instructional method.

INDEX WORDS: Physics education research, PER, Studio physics, Interactive engagement, Force Concept Inventory, FCI, Gender performance, Gender gap, Demographic survey, Ethnic differences, Urban institution, Algebra-based physics, Georgia State University, Item analysis, Two-way ANOVA, Normalized gain, Instructional method, Pre-test post-test design

ASSESSING THE EFFECTIVENESS OF STUDIO PHYSICS AT GEORGIA STATE UNIVERSITY

by

BRIANNA M. UPTON

Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2010

ASSESSING THE EFFECTIVENESS OF STUDIO PHYSICS AT GEORGIA STATE UNIVERSITY

by

BRIANNA M. UPTON

Committee Chair:	Brian Thoms
Committee:	Ramon Lopez
	Cherilynn Morrow
	John Evans

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
August 2010

DEDICATION

This thesis is dedicated to my heavenly Father, to my husband Jeremy of 11 years, and to my children, Brejah, Aria, and Lex. During the past three years, I have lived between two very different worlds, at home as a mother and wife, and on campus as a physics graduate student. Without the unfailing love and support of my hero and husband, Jeremy Upton, and my three beautiful, amazing children, this work would not be possible.

ACKNOWLEDGEMENTS

I want to thank my parents who supported me when I told them after six years of being a stay-at-home mom that I felt led to go back to school. My mom came and lived with us during the last few weeks of the writing process and defense. I cannot thank her enough.

I also want to thank my advisor, Dr. Brian Thoms, for accepting me as his first PER grad student. His direction, patience, and inspiration have been invaluable.

Lastly, I want to thank my committee, Dr. Cherilynn Morrow, Dr. Ramon Lopez, and Dr. John Evans for their encouragement and support. They taught me the basics of PER. I am grateful for their knowledge, understanding, and guidance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

1.1 History of Physics Education Research

Physics Education Research, as an area of academic study, is a relatively new field. The pioneering work in the field of Physics Education Research can be traced back to the founders of the American Association of Physics Teachers: Paul Klopsteg of the University of Minnesota, Homer Dodge of the University of Oklahoma, and F. T. Richtmeyer of Cornell University (Phillips, 1977). They were interested in the challenges of teaching physics. The dawn of the "space race" and the Atomic Age at the end of the Second World War spurred on a national interest in improving the quality of science education and curricular improvements that would benefit both students and teachers of physics. It soon was recognized in the 1960's and 1970's that there was a need for curriculum reform on a national scale that would require the input of those involved in the behavioral sciences and the field of education to help better understand the needs of the student and to better train a cadre of quality teachers. As early as 1956, Jerrold Zacharias, professor of physics at Massachusetts Institute of Technology (Lopez, R. & Schultz, T., 1991) also had deep concerns about science education. Jerrold along with physicists Francis Friedman of MIT, Philip Morrison of Cornell University, and Bob Karplus of University of California at Berkeley, led K-12 science education reform. As time progressed into the 1990s, educators and professors in the field of physics started to focus research and collect data around the processes and methodologies of teaching physics. This body of research began to be referred to as "Physics Education Research" or simply, PER.

"Physics education research differs from traditional education research in that the emphasis is not on educational theory or methodology in the general sense, but rather on student understanding of science content" (McDermott, 2001). Physics Education Research as a field seeks to improve existing techniques and to bring about new and innovative methods for the teaching of physics. Those involved in PER aim to make physics instruction synonymous with student learning of physics and how to apply it. They seek to maximize the effectiveness in methodology and educational value for the sake of the student. In other words, PER seeks to teach and communicate physics in a pedagogically sound manner that has a demonstrable positive effect on student learning. At its center, the field focuses on getting the students to understand the applications of physics; it is not focused merely on problem-solving ability. One underlying PER goal is that the science students begin to share the interest and enthusiasm that physicists have for their work, whether those students intend to become future physicists or aspire to other goals.

What has been discovered around the world is that the traditional, didactic approach to teaching does not yield the highest student learning gains (Reddish, E., Saul, J., & Steinberg, R., 1998). Physics teachers from the primary to the post-secondary level of the educational process have observed the pedagogical limitations of the lecture approach alone. Pedagogy which utilizes an integrated, collaborative, activity-based learning environment has repeatedly been shown to be the way students learn best, thus, the onset of different teaching methods for physics curriculum, such as Peer Instruction (PI), Technology Enabled Active Learning (TEAL), Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP), and Studio Physics.

Physics Education Research is motivated to increase student engagement and conceptual grasp of physics while not negating its appreciation for conducting and reporting its research in a rigorous, scientific manner. It is the intense desire of those who study and practice the ideas of

PER to throw off the stigma that this type of research should be considered as "soft science" (McDermott, 2001). Indeed, it is this demand for scientific rigor and credibility that places PER within the mandate of physics departments, rather than education departments.

David Hestenes, Oersted Medal recipient and professor emeritus of physics at Arizona State University stated, "I have since seen PER emerge as a credible discipline in its own right, with a growing body of reliable empirical evidence, clarification of research issues, and most important of all, an emerging core of able and committed researchers within physics departments across the country…. It is a serious program to apply to our teaching the same scientific standards that we apply to physics research" (Hestenes,1998).

Lillian McDermott, also an Oersted Medal recipient and professor of physics at the University of Washington is considered by many scholars to be the recognized leader in physics education reform efforts. University and colleges across the country widely use the textbooks developed by McDermott's Physics Education Group: *Physics by Inquiry* (Wiley, 1996) and *Tutorials in Introductory Physics* (Prentice Hall, 2002). She has also been instrumental in developing accurate testing procedures. Robert Beichner, who consulted for GSU in the design of its studio classroom, is the developer of the Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project (Beichner, 2006). It is a studio-style environment that promotes scientific investigation of physics problems through group interaction. Nearly 100 institutions around the world have adopted SCALE-UP as a model to implement in their learning environments.

Another alternative to conventional instruction methods has been introduced by Professor Eric Mazur of Harvard University: Peer Instruction (PI) model for collaborative learning methods. The Peer Instruction approach is student-centered and inquiry-based learning that is supplemental to class lectures. Mazur's research has led to the development of an instructional supplement known as ConcepTests (Mazur, 1997). His approach is to use Peer Instruction along with the ConcepTests during the lecture. The ConcepTest is administered throughout his lectures to allow his students to interact in small groups about the concepts covered in the lecture. Subsequently, they are given a form of "pop quiz" to see how the lecture, along with the student interaction, has impacted the students' grasp of the concept presented. As a methodology, PI's adoption increases annually across the county (Lasry, 2008). Similar to studio physics, PI's results continue to show significant increases in students' conceptual understanding as compared to the conventional, didactic approach (Lasry, 2008).

1.2 Philosophical Values of PER

As an academic discipline, PER derives its epistemological framework (how we know what we know when we study physics education) from constructionism (currently, used interchangeably with "constructivism") in which, it is believed that students learn best when they derive meaning, and hence, learning, as a community (Crotty, 2003). E. F. Redish (1999) calls it the constructivism principle: "Individuals build their knowledge by processing the information they receive, building patterns of association to existing knowledge." It is from this sense, then, that studio physics seeks to challenge students to think scientifically through inquiry, critical thinking in a collaborative setting, and communicating with each other throughout the process of scientific investigation (Wilhelm, 2007). Utilizing this framework, students are able to grasp the concepts much faster and with much better results.

With this in mind, the traditional approaches to transferring knowledge from instructor to student flies in the face of the values of today's student. Studio physics and other collaborative, interactive, peer-based instructional models that are based on constructivism are more adequately aligned with the generational values of today's university student. If the goal of education is based upon the desire for the student to gain mastery of the subject, the methods of how that is done must be re-examined from time to time to ensure that the instructor is adequately transferring our instruction and adequately received by the student. This is what PER in general, and studio physics in particular, are all about: physics must be student-centered, not instructor-centered.

The cultural diversity and differing epistemological frameworks may not be accounted for in a traditional, lecture-based classroom. For example, a student raised in a culture of group and family interaction informing learning, or an experiential approach to learning, would find the traditional physics classroom a cold and uninviting place to learn. The ensuing struggle to grasp concepts in this type of environment would significantly hinder that student from performing well. Studio physics seeks to provide that alternative to the conventional approach that would provide that student with an environment and a community to aid learning at the conceptual level.

1.3 Studio Physics

"Studio physics" is the name given to various types and models of collaborative, activity-based, interactive, student-centered learning environments. The pedagogical method is based upon the pioneering work of Jack Wilson of Rensselaer Polytechnic Institute (Wilson, 1994). Because of the success of studio methods, many institutions have revamped the way introductory physics courses are taught. A full understanding of what "studio physics" is comes with the understanding of what the "conventional" physics course is. In the conventional course, the instructor gives a typical lecture in a lecture room. It is characterized primarily by passive-learning in which there is minimal student interaction and engagement. However in the studio classroom, the students become active-learners and the instructor lectures minimally. Studio physics is also marked by PER-based materials, computer simulations to enhance conceptual understanding, and graduate assistants on hand for individual student help.

Classroom Design

The design of the GSU studio classroom is largely based on Beichner's (2007) SCALE-UP model. Located in Georgia State's Classroom South building, the studio/workshop classroom is a multimedia, interactive learning environment. The most important technological presence in the studio classroom is the roundtables (Gaffney, 2008). There are six of these tables, each fitted with 9 chairs, equipped with internet-accessible laptops. Every wall is covered with whiteboards to allow optimal space for working on concepts and problems and for displaying that work to peers and the instructor. The classroom is also fitted with cameras and microphones throughout. Figure 1.1 shows portions of the studio classroom at GSU.

Figure 1.1 The studio classroom at Georgia State University. The room is characterized by roundtables and whiteboards. Each of the six roundtables seats 3 groups of 3 students for a total of 54 students. Fitted with cameras, microphones, and laptop computers, the room is well-equipped for active-learning.

Evidence for the efficacy of active-learning pedagogies, like the studio format, is vast. (Buck, J. & Wage, K., 2005; Johnson, D., Johnson, R., & Smith, K., 1998; Prince, M., 2005; Springer, L., Stanne, M., & Donovan, S., 1999). Some of the most noted programs of universities that have adopted studio-style classrooms are the Workshop Physics Project at Dickinson College, the TEAL project at Massachusetts Institute of Technology, the physics program at University of California at Davis, and SCALE-UP at North Carolina State University. Other institutions with research groups offering PER in its PhD physics programs are Harvard University (Eric Mazur), Ohio State University (Ken Wilson), University of Maryland (Joe Redish), University of Washington (Lillian McDermott), University of Colorado at Boulder, and North Carolina State University (Robert Beichner). Georgia State University (GSU) has now joined in on the quest to improve introductory physics courses. In the Fall 2008 semester, studio physics was implemented, modified, and elaborated on based on Beichner's SCALE-UP model.

1.4 Georgia State University and PER

Georgia State University, the second-largest research institution in the University System of Georgia, has embarked upon a significant opportunity to research the effects of the studio learning environment on a unique demographic. There has been very little study devoted to an undergraduate student demographic where Blacks and Asians constitute a majority. There is even less research on the success of a studio physics environment implemented at a non-residential, urban institution. Such an opportunity exists here at Georgia State. To capture and illustrate this uniqueness, demographic data from the past 16 years is provided in Appendix C. The general trends are that White student enrollment is decreasing and Black and Asian student enrollment is increasing. For the Fall 2009 semester, Appendix D reveals that there were 22,385 undergraduate students enrolled. Of this amount, 41% were White, 36% were Black, and 12%

were Asian (other ethnic groups are not presented in this study because of their low percentages in enrollment and in registration in Physics 1111). The unique demographic of GSU's undergraduate population can open up a world of possibilities in terms of studio physics' effects on the performance and conceptual understanding of these two ethnic groups.

The fact that GSU is a predominantly non-residential institution introduces another variable in evaluating the efficacy of studio physics at Georgia State. The opportunity to see what can be learned about student learning in an open, transient environment like the one that exists here at GSU could have implications on the university itself, the Physics department, and the field of PER research as a whole.

2 DATA AND ANALYSIS

2.1 Experimental Methods and Design

Comparisons were made using gains in student understanding based on mechanics concepts. Results showed that both conventional and studio student populations began the algebra-based introductory physics course (PHYS1111) with similar content knowledge. The questions to be answered in this study are:

1. Does instructional method (conventional versus studio) affect learning gains? If so, by how much?
2. What is the effect of studio physics instructional methods on various ethnic groups?
3. Are there differences in the effects of these instructional methods based on gender?

Study Design

The sample consisted of students from 5 consecutive semesters (Fall 2008 – Spring 2010) who took both pre- and post-tests of the Force Concept Inventory (Hestenes, D., Wells, M., & Swackhamer, G., 1992). The data include 29 classes taught by 8 different instructors. The students (N = 785) were enrolled in Physics 1111, the first sequence of GSU's algebra-based introductory physics course. The study employed a pre-test - post-test design. The groups were named as conventional students (N_c = 431) and studio students (N_s = 354). There were no statistically significant differences between the two groups. The pre-test was given within the first week of the semester, and the post-test was given during the last three weeks.

Two types of studies are being conducted. The first study is a longitudinal study in which comparisons between conventional and studio test scores are made for all five semesters. This assesses the success of studio physics' implementation at GSU, overall. The second study is a cross-sectional study in which a snapshot of the sample, using the Fall 2009 semester, is investigated to assess ethnic and gender test score differences. In this study, the sample subset is comprised of conventional and studio students, where the sample sizes are $N_c = 121$ and $N_s = 136$, respectively.

2.2 Assessment Tools and Statistical Tests

As previously stated, the main purpose of Physics Education Research is to increase the effectiveness of physics teaching practices. But how is this effectiveness accurately measured? The test chosen to assess GSU's sample is the Force Concept Inventory. Its wide use, accessibility, credibility, and validity make it the appropriate instrument (Savinainen, A., & Scott, P., 2002). The revised FCI is being used in the Physics 1111 course in the form of a pre-test and post-test to determine the improvement, or lack of improvement, in students' conceptual understanding.

The Force Concept Inventory

The Force Concept Inventory, better known as FCI, was created by Halloun and Hestenes of the Department of Physics at Arizona State University (Hestenes, Wells, & Swackhamer, 1992). It is a multiple-choice test designed to assess students' understanding of basic Newtonian physics. Emphasizing qualitative reasoning, the FCI consists of 30 items related to force and motion concepts in which there is no numerical computation involved. The six areas it primarily focuses on are kinematics, Newton's three laws, the principle of superposition, and types of forces (Savinainen & Scott, 2002). Each question on the FCI offers only one correct Newtonian solution (Hestenes et al, 1992). The incorrect possible answers that are presented as "common-sense distracters" are based upon student's misconceptions about that particular topic.

The precursor to the FCI was the Mechanics Diagnostic Test (MDT). The MDT's purpose was to assess students' understanding and knowledge of basic mechanics concepts. The Kuder-Richardson reliability coefficient was 0.86 for the MDT pre-test and for post-test use it was 0.89, which indicated that the MDT was extremely reliable. Though the MDT was shown to be a consistently reliable test, the developers desired an inventory that tested students' misconceptions of Newtonian physics (Savinainen & Scott, 2002). Therefore, the FCI was developed and initially published in 1992. The FCI provided researchers with a reliable tool that gave a more complete view of the many misconceptions in Newtonian physics, in a systematic format (Hestenes et al, 1992). In the beginning phases of its use, the FCI revealed dramatic results on students completing an introductory college level physics course: "Nearly 80% of the students could state Newton's Third Law at the beginning of the course. FCI data showed that less than 15% of them fully understood it at the end" (Hestenes, 1998). Needless to say, having been in use over the last 15 years, the FCI "is now credited with stimulating reform of physics education" (Evans, L. & Hestenes, D., 2001), and the instrument is a powerful tool for improving both the learning and teaching mechanics.

The latest version of the FCI was developed in 1995 (Savinainen & Scott, 2002). This revised test is claimed to have "fewer ambiguities and a smaller likelihood of false positives" (Hake, 1998). The FCI is available in nine languages: Chinese, English, Finnish, French, German, Malaysian, Spanish, Swedish, and Turkish.

R. R. Hake is famous for his work of surveying over six thousand students (Hake 1998) and for defining the normalized gain. Also called the Hake factor, or Hake gain, the normalized gain is the ratio of the actual gain to the maximum possible gain (Sahin, 2009):

$$g = \frac{actual\_gain}{\max imum\_possible\_gain} = \frac{(posttest\_score) - (pretest\_score)}{(\max imum\_score) - (pretest\_score)}$$

This gap-closing measure is based on the work of Fran Gery (Bao, 2006). Hake used normalized gain to describe the change in a student's performance, after instruction. When g = 0 (same pre- and post-test score), the indication is that the student has not learned any Newtonians concepts. On the other hand, g = 1 implies that the student has learned everything they need to know (in Newtonian concepts).

T-test of the Differences

T-tests are used to compare the means of two populations (students in conventional physics and students in studio physics). Because neither population mean is known, a sample has been taken from both populations to perform a two-sample t-test. The type of t-test performed is the independent t-test: the students in both traditional and studio populations have no connection of consequence to each other.

ANOVA Statistical Tests

Analysis of Variance (ANOVA) is a statistical method that compares group means on a dependent variable. Used by most PER researchers, it is a powerful statistical test that is sensitive to the differences among the groups being compared. There are different types of ANOVAs depending on the number of independent variables (or factor) that the study incorporates. Appropriately called "factorial analysis of variance", a multiple factor ANOVA is used when two or more independent variables and their interaction are being analyzed. For the cross-sectional study, three separate two-way ANOVAs are performed using the software SPSS 17.0 (Statistical Package for the Social Sciences 17.0). The output tables display the quantities used to calculate the F statistic. Although these terms are comparatively unimportant, their values determine the most important numbers: the three F-ratios. F-ratios for ANOVAs are always upper-tailed, so critical values for F are provided only for $\alpha = .05$ and $\alpha = .01$.

Two-way ANOVAs are performed on the data. Performing a two-way ANOVA not only answers if there is a single effect on the independent variables and the gain, but it demonstrates if there exists an interaction between the independent variables themselves. An interaction effect is the effect of the two independent variables (gender and ethnicity) working together on the dependent variable (gain score) (Huck, 2008). This is distinguished from the main effect of gender and ethnicity individually.

Post hoc Tests

A post hoc test has to be performed when an ANOVA reveals a statistically significant difference between the means of the populations being considered and the null hypothesis is rejected. These tests employ statistical procedures that analyze every possible pair of means. It then determines if the differences are significant (Huck, 2008). There are a variety of post hoc tests (all named after their developers) to choose from. Typically, they are chosen based on the design of the ANOVA.

The three post hoc tests used in this study are the Tukey Honest Significant Difference (HSD) Test, Simple Effects Analysis, and the Student-Newman-Keuls test. In the Tukey post hoc test, the differences between the means of all of the groups will be determined first. Then for each, the difference score is compared to a critical value (the HSD value) to determine if the difference is statistically significant. The Simple Effects Analysis is a post hoc test in which single degree of freedom comparisons are done on the subeffects. This test is done when there is a significant interaction effect. The Student-Newman-Keuls (SNK) post hoc test is recommended when there are only three means being compared (Cardinal, R. & Aitken, M., 2005). In this test, the sample means are ordered from smallest to largest and tested at specific levels of significance.

3 RESULTS AND DISCUSSION

3.1 Conventional Versus Studio

The longitudinal part of the study has a design in which FCI and demographic data are collected on multiple occasions from the same population, Physics 1111 students. It is an ongoing effort in assessing the effectiveness of studio physics. Thus far, this five-semester pilot study has shown that studio physics is an effective instructional method at GSU. Figure 3.1 and Table 3.1 display the normalized FCI gains.

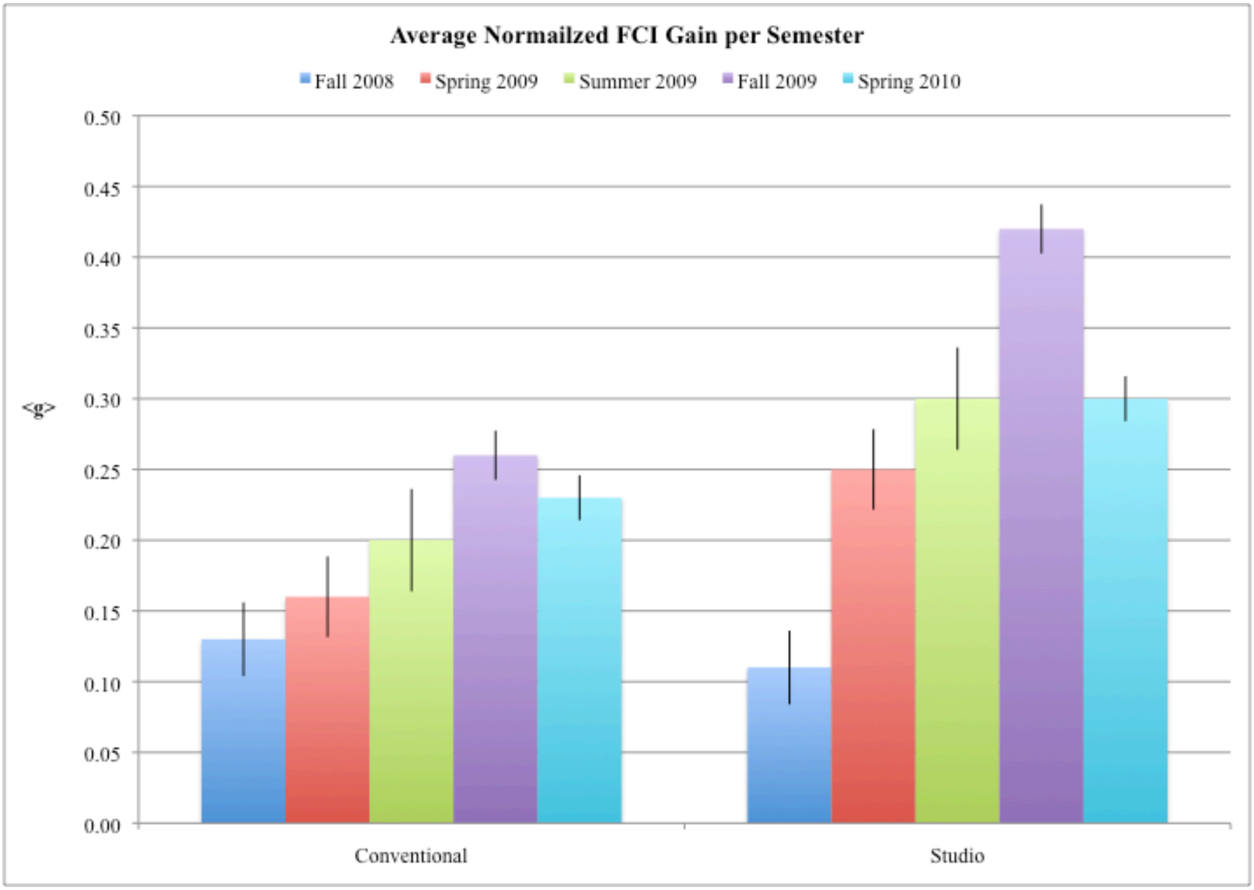


Figure 3.1 Average normalized FCI gain per semester. This figure compares the average gains of conventional and studio classes. The first semester of studio implementation reflects similar gains in both modes; each successive semester reflects increased studio gains as compared to conventional.

Table 3.1 Average normalized FCI gain per semester with standard error of the mean. <g> = average normalized gain. The table displays results for conventional and studio classes.

Semester	Conventional <g>	Studio <g>
Fall 2008	0.13 ± .03	0.11 ± .03
Spring 2008	0.16 ± .03	0.25 ± .03
Summer 2009	0.20 ± .04	0.30 ± .03
Fall 2009	0.26 ± .02	0.42 ± .02
Spring 2010	0.23 ± .02	0.30 ± .02

Figure 3.1 shows upward trends in learning gains for both conventional and studio classes. However, the increase in learning gain of the conventional classes is not statistically significant. Over time, the gains may level out. Additionally, the phasing out of part-time instructors and instructor awareness of the study may have contributed to the improvement in both conventional and studio classes.

The average normalized gain <g> for conventional instruction nationally is 0.25, and for interactive engagement instruction is between 0.36 and 0.68 (Mazur, 1997). Table 3.1 shows that GSU's gains are comparatively low. A contributing factor to this is students having little prior physics knowledge coming into the course: average pre-score for GSU is 7.6 out of 30.

The t-test found that the difference in normalized gain between conventional students (M = .21, SD = .21) and studio students (M = .33, SD = .24) was statistically significant at the .05-level, $t(700.898) = -4.431$, $p < .001$. GSU's results mimic Hake's results from his 6000-student survey of the FCI (Hake, 1998). His study, which compared traditional and interactive engagement classes, showed that conventional classes had normalized gains in the range 0.19-0.27 and interactive engagement classes had normalized gains in the range 0.34-0.62.

The effects size for this independent-samples t-test is:

$$\hat{d} = \frac{\left|\overline{Y_1} - \overline{Y_2}\right|}{\sqrt{\frac{S_{Y_1}^2 + S_{Y_2}^2}{2}}}$$

where $\overline{Y_1}$ and $\overline{Y_2}$ are the sample means and S_{y_1} and S_{Y_2} are the standard deviations. The numerator is the absolute value of the mean difference, and the denominator is the square root of the average of the variances. Effect size is a descriptive statistic that measures the extent of the difference under investigation (Huck, 2008). For this study, the result is $\hat{d}$ = .54 which is a medium effect. Cohen's (1998) proposed values for *d* are:

0.2 ≡ small effect
0.5 ≡ medium effect
0.8 ≡ large effect

2 x 5 ANOVA Results (Instructional Method Over 5 Semesters)

A two-way 2 x 5 ANOVA was performed to see the effect of semesters as an independent variable. The ANOVA output is shown in Table 3.2. All three effects were statistically significant: the main effect of instructional method, F(1, 775) = 16.288, p < .001, MSE = 0.046, the main effect of semester, F(4, 775) = 15.933, p < .001, MSE = 0.046, and the interaction effect of both factors, F(4, 775) = 2.992, p < .05, MSE = 0.046. When the interaction effect is statistically significant in an ANOVA, the main effects have to be dismissed. This signifies that the effect on one factor is not the same at the levels of another; thus the statistical significance varies across semesters. Because of this significant interaction, we must test simple effects (Maxwell, S. & Delaney, H.). A post hoc test using single degree of freedom comparisons identify which subeffects (semesters) are contributing to the significant interaction effect. (Myers, J. & Weel, A., 2003).

Table 3.2 2 x 5 ANOVA statistics for normalized gain. IM = instructional method, df = degrees of freedom, F=F-ratio, Sig.= p-value. IM and semester are the two main effects and IM x semester is the interaction effect. All three effects have significant F-ratios (p < .05).

	Sum of Squares	df	Mean Square	F	Sig.
IM	0.742	1	0.742	16.288	0.000
Semester	2.905	4	0.726	15.933	0.000
IM x Semester	0.533	4	0.133	2.922	0.020
Error	35.323	775	0.046		
Total	41.470	784			

Figure 3.2 shows the gap between conventional and studio normalized means. The first semester of implementation of studio was the only semester where the average normalized gain, <g> = .11, was lower than the that of the conventional class, <g> = .13. Table 3.3 shows the single degree of freedom comparisons identifying the statistical significance of the gaps. Equal variances are not assumed in this simple effects analysis.

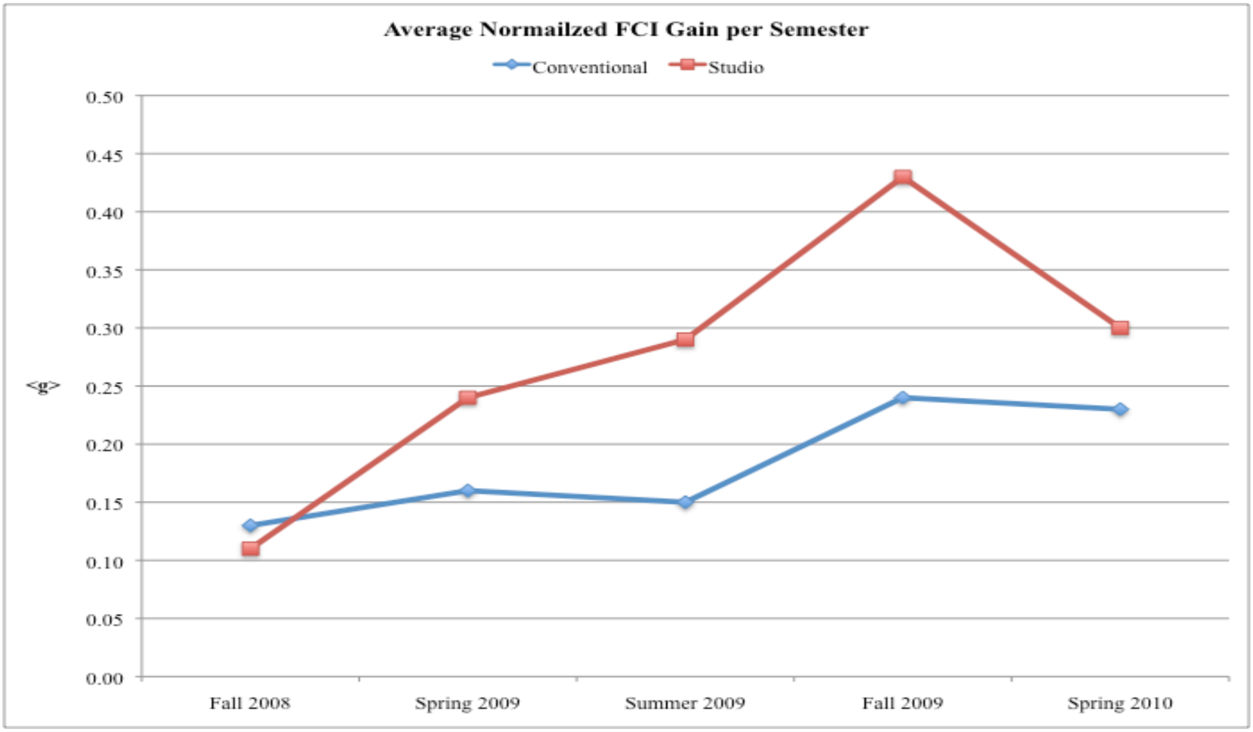


Figure 3.2 Average normalized FCI gain per semester (as a line graph). This figure depicts the gap in conventional and studio learning gains.

Table 3.3 Simple effects analysis. The table shows three semesters with statistically significant differences ($p < .05$) in normalized gains. IM = instructional method, C = conventional students, S = studio students, SD = standard deviation, df = degrees of freedom, Sig. = p-value.

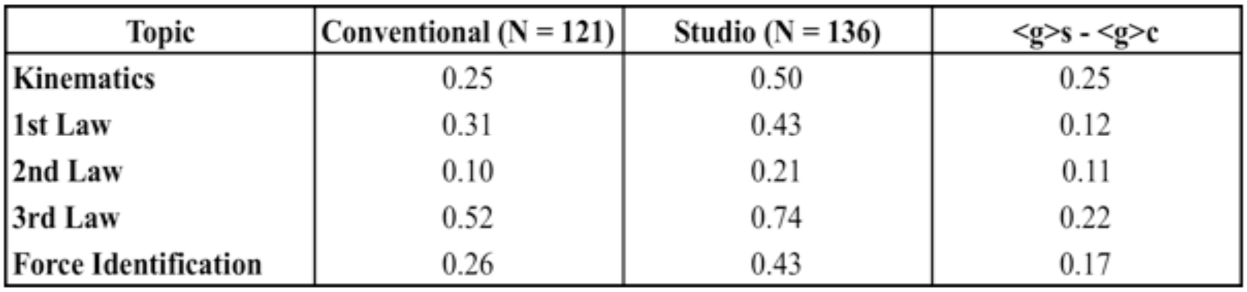
Semester	IM	N	Mean	SD	SEM	t	df	Sig.
Fall 2008	C	77	0.1337	0.2289	0.0261	0.578	37.771	$p > .05$
	S	16	0.1097	0.1292	0.0323			
Spring 2009	C	48	0.1595	0.1975	0.0285	-2.133	91.768	$p < .05$
	S	48	0.2531	0.2311	0.0334			
Summer 2009	C	35	0.1983	0.2143	0.0362	-1.958	64.997	$p > .05$
	S	32	0.2967	0.1970	0.0348			
Fall 2009	C	127	0.2595	0.1959	0.0174	-5.696	259.723	$p < .001$
	S	143	0.4199	0.2650	0.0222			
Spring 2010	C	144	0.2308	0.1908	0.0159	-2.717	246.645	$p < .01$
	S	115	0.2949	0.1868	0.0174			

The analysis reveals that the semesters that have statistically significant differences in normalized gain (between studio and conventional classes) are Spring 2009, Fall 2009, and Spring 2010. The sample sizes for conventional and studio classes were equivalent for Spring 2009. For the last two semesters of the study, the studio physics classroom was fully functional.

Topic Differences

Table 3.4 shows the average normalized gain by topic for both conventional and studio students. The normalized gains were taken by averaging the scores for the items of each specific topic. It reveals that there were differences on the FCI related to the topics covered. While it is still evident that studio gains are higher overall, there is a substantial learning gain ($\langle g \rangle = 0.74$) for Newton's third law concepts. Conversely, it is also shown that both instructional methods are weak in teaching Newton's second law concepts. The normalized learning gains for conventional and studio are 0.10 and 0.21, respectively. Topic item analysis has revealed where instructors need to intervene.

Table 3.4 Fall 2009 average normalized gain by topic. The five basic areas covered on the FCI yield different learning gains. The difference in learning gains between conventional and studio students is $<g>_s$ - $<g>_c$.



Topic	Conventional (N = 121)	Studio (N = 136)	<g>s - <g>c
Kinematics	0.25	0.50	0.25
1st Law	0.31	0.43	0.12
2nd Law	0.10	0.21	0.11
3rd Law	0.52	0.74	0.22
Force Identification	0.26	0.43	0.17

Just as studio students showed higher average normalized gains on overall FCI scores, the same holds true when looking at student gains by specific topic. There are two topics that stand out above the others. There is a 25% difference in normalized gain on kinematics and a 22% difference in normalized gain on Newton's third law items. It appears that studio physics is more successful in its ability to help students grasp and understand specifically on these two Newtonian topics than methods employed in the conventional physics classroom. Further discussion about topic differences continues in the following section on gender gap. A look at the historical and current views on gender performances prompts this study to probe into a specific item analysis to detect differences.

3.2 Ethnic and Gender Differences

Fall semester of 2009 was chosen as the semester to do a more in-depth investigation of conventional and studio learning gains. Not only were the differences in gains statistically significant, but this particular semester was chosen because it had the largest number of classes to do the statistics and because it had an equivalent number of classes (four conventional classes and four studio classes) participating in the study. Additionally, the students who participated signed informed consent forms, and they each completed a demographic survey (see Appendices

A and B to view the forms). Moreover, issues associated with implementation of studio during the first two semesters were resolved, and the format was fully functional by the Fall 2009 semester.

ANOVA tests were performed to compare instructional method, ethnicity, and gender with gain scores. Both normalized gain and raw gain were calculated for each student. Normalized gain is most widely used in PER FCI studies; however because of the nature of the normalized gain calculation, students with higher pre-test scores get higher normalized gain even if their raw gain is equivalent to another student's. For example, Student 1 gets a raw score of 8 on the pre-test and a raw score of 24 on the post-test, yielding a raw gain of 53%. Student 2 gets a raw score of 5 on pre-test and 21 on the post-test, again yielding a raw gain of 53%. Conversely, when the normalized gains are calculated, their values are 73% for Student 1 and 64% for Student 2. It measures the fraction of the available improvement that is obtained. There has been ongoing discussion and debate over the use of normalized gain to assess learning gain (Coletta, V. & Phillips, J, 2005; Marx, J. & Cummings, K., 2007; Willoughby, S. & Metz, A., 2009); however, it is still considered the standard measurement in PER to calculate normalized learning gain. Normalized gain is stable against random guessing and can be used without a mathematical correction (Bao. 2007). Thus, although we present raw gains in the cross-sectional study, we use normalized gains to quantify and assess most of the results.

Table 3.5 shows the average FCI scores for each gender and ethnic group, based on conventional and studio instruction. Several trends are identified. Female students come into the Physics 1111 course with lower pre-test scores ($\langle \text{Pre} \rangle_{\text{conventional}} = 0.23$, $\langle \text{Pre} \rangle_{\text{studio}} = 0.20$) than the male students ($\langle \text{Pre} \rangle_{\text{conventional}} = 0.33$, $\langle \text{Pre} \rangle_{\text{studio}} = 0.30$). Black students come into the course with the lowest prior physics knowledge ($\langle \text{Pre} \rangle_{\text{conventional}} = 0.20$, $\langle \text{Pre} \rangle_{\text{studio}} = 0.19$)

compared to the White students (<Pre>$_{\text{conventional}}$ = 0.32, <Pre>$_{\text{studio}}$ = 0.26) and Asian students (<Pre>$_{\text{conventional}}$ = 0.25, <Pre>$_{\text{studio}}$ = 0.27). Another observation is that Asian students have the lowest raw and normalized learning gains. Their gains are substantially lower in the conventional classes (<G> = 0.13, <g> = .17). The statistical significance of these gender and ethnic differences was determined.

Table 3.5 Average FCI scores by gender and ethnicity. Demographic data from the Fall 2009 sample is tabulated for average raw pre-test scores <Pre>, average raw post-test scores <Post>, average raw gain <G>, average normalized gain <g>, and sample size.

Fall 2009 Average FCI Scores by Ethnicity and Gender
Total Sample Size: N=257

	<Pre>	**<Post>**	**<G>**	**<g>**	**N**
All - Conventional	0.27	0.46	0.18	0.26	121
Male - Conventional	0.33	0.52	0.19	0.30	50
Female - Conventional	0.23	0.41	0.18	0.23	72
White - Conventional	0.32	0.52	0.20	0.32	56
Black - Conventional	0.20	0.40	0.20	0.25	32
Asian - Conventional	0.25	0.38	0.13	0.17	27
All - Studio	0.23	0.56	0.33	0.42	136
Male - Studio	0.30	0.63	0.33	0.47	44
Female - Studio	0.20	0.52	0.32	0.40	92
White - Studio	0.26	0.61	0.34	0.47	36
Black - Studio	0.19	0.53	0.33	0.41	49
Asian -Studio	0.27	0.56	0.29	0.39	30

2 x 2 ANOVA Results (Gender Differences)

A two-way 2 x 2 ANOVA showed that the main effect of instructional method on normalized gain scores was statistically significant, F(1, 253) = 30.768, $p < .001$, MSE = 0.056. Studio students received higher normalized gain scores than conventional students. The ANOVA also revealed a statistically significant main effect of gender such that male students received higher normalized gain than female students, F(1, 253) = 4.995, p < .05, MSE = 0.056. There was no interaction.

Table 3.6 2 x 2 ANOVA statistics for normalized gain. IM = instructional method, df = degrees of freedom, F=F-ratio, Sig.= p-value. IM and gender are the two main effects (both are significant, p < .05) and IM x gender is the interaction effect (not significant, p > .05).

	Sum of Squares	df	Mean Square	F	Sig.
IM	1.709	1	1.709	30.768	0.000
Gender	0.278	1	0.278	4.995	0.026
IM x Gender	0.000	1	0.000	0.002	0.965
Error	14.057	253	0.056		
Total	16.047	256			

2 x 3 ANOVA Results (Ethnic Differences)

A two-way 2 x 3 ANOVA found a main effect of ethnicity on normalized gains, F(2, 223) = 4.091, p < .05, MSE = 0.057, a main effect of instructional method, F(1, 223) = 29.824, p < .001, MSE = 0.057, and no interaction. Analyzing raw gains, a 2 x 3 ANOVA did not identify any statistically significant main effect of ethnicity or interaction effect between instructional method and ethnicity. It did however find a main effect of instructional method on raw gains, as it did with normalized gains.

Table 3.7 2 x 3 ANOVA statistics for normalized gain. IM = instructional method, df = degrees of freedom, F=F-ratio, Sig.= p-value. IM and ethnicity are the two main effects (both are significant, p < .05) and IM x ethnicity is the interaction effect (not significant, p > .05).

	Sum of Squares	df	Mean Square	F	Sig.
IM	1.697	1	1.697	29.824	0.000
Ethnicity	0.465	2	0.233	4.091	0.018
IM x Ethnicity	0.049	2	0.025	0.431	0.650
Error	12.685	223	0.057		
Total	14.682	228			

As previously stated, two-way ANOVAs were performed for raw gains also. The same design was incorporated: a 2 x 2 ANOVA for instructional method by gender, and a 2 x 3 ANOVA for instructional method by ethnicity. The main effect of instructional method was statistically significant, $p < .001$; however, gender and ethnicity main effects were not significant. Moreover, the interaction effect of instructional method and gender (similarly, instructional method and ethnicity) was not statistically significant for raw gain scores, $p > .05$.

Given the significant results of the two-way ANOVAs, it becomes important to analyze more details about the FCI scores. The following analyses will look at specific information concerning topics on the FCI, gender gaps on pre-test and post-test scores, and ethnic distinctions as evidenced by group averages.

Ethnic Differences

The 2 x 3 ANOVA (Table 3.7), showed that the differences in normalized gain between ethnic groups (main effect) were statistically significant ($p < .05$). Differences in raw gain were not statistically significant ($p > .05$). Table 3.5 displays ethnic group averages for FCI pre-test scores, post-test scores, raw gain, and normalized gain. The following observations have been made:

- Asian students have the lowest raw gains on the FCI (0.13 for conventional classes and 0.29 for studio classes in comparison to White students and Black students who had gains of 0.20 in conventional classes and 0.33-0.34 in studio classes).
- Black students have the lowest FCI pre-test scores, while White students had the highest (5.8 out of 30 for Black students, 7.7 for Asian students, and 8.9 for White students).

- White students have the highest normalized gains on the FCI. For example, in the conventional courses, White students had a normalized gain of 0.32 while the gain was 0.24 and 0.17 for Black and Asian students, respectively.

The SNK post hoc test for ethnicity did not detect exactly where the differences lie statistically, but much is to be gained by the aforementioned observations.

Black students had just as much raw gain as the White students. As GSU recruits, enrolls, and graduates many Black students, and will most likely continue to in the future, studio physics can help GSU in its appeal and the future performance of this major segment of its student body. Conversely, the raw gain scores of Asian students showed that, while they achieved better grades in the courses, the improving of their grasp of force concepts between pre- and post-tests was minimal. Moreover, although the studio format is more effective for all three ethnic groups, it is slightly less effective for Asian students. It may be that there exists a language barrier due to many Asian students being non-native English speakers. Further research needs to be done to discover what, if any, impact a language barrier may have on Asian students. This can be accomplished by adding a question on the pre-test survey to determine which students may not benefit from the group interaction due to unfamiliarity with speaking English.

Considering the fact that Asian and Black students constitute a majority in the studio classroom, the effort to improve introductory physics teaching at GSU must be continued and could have the greatest effects on students who may be underserved in conventional instructional settings. As GSU seeks to be well-rounded in its approach to offering quality education to its students, studio physics is and can continue to be a major contributor to that institutional and departmental goal.

Gender Differences

"Gender gap" is the term used to describe the difference in test scores of male and female students (Kahle, J. & Meece, J., 1994). Historically, males outperform females on science tests. Many factors have been subjected to debate as to why this gap persists: from biological differences to social differences (Jovanovic, J. & Dreves, C., 1995). Traditionally speaking, females have not grasped concepts at the same rate or levels as their male counterparts (Lorenzo, et. al., 2006). While straying away from the possible causes of this disparity, PER research has shown that the gender differences can be helped by alternative methodologies. Harvard University pioneered a study on gender gap issues and found that cooperative learning techniques decreased this gap in FCI scores (Lorenzo, M., Crouch, C., and Mazur, E., 2005). In some instances, their study demonstrated that the gender gap was totally eliminated. It has become pertinent to physics education researchers to assess the effects of interactive engagement methods on the gender gap and to determine its success at reducing or eliminating the gap. Most research on gender gap in physics is done for the calculus-based sequences. Recent studies on this issue have been done at the University of Colorado (Kost, L., Pollock, S. & Finkelstein, N., 2009; Pollock et al, 2007), Harvard University (Lorenzo et al, 2006), and the University of Minnesota (Docktor & Heller, 2008). They investigate the effects of instructional method on the gender gap. Harvard University has shown that interactive engagement methods reduce, and in some cases, eliminate the gender gap; while University of Colorado showed that it had little effect.

The current study pilots an effort to investigate these effects for the algebra-based physics course. The other unique difference in this study is that the sample is female-dominated. The majority of the research showing that males outperform females (Kahle & Meece, 1994), is done in male-dominated contexts. In these studies (Kost et al, 2009, Docktor & Heller, 2008), males account for nearly 75% of the classes sampled. However, the male make-up of GSU's physics courses is under 50%. Similar to the University of Colorado study (Kost et al, 2009), GSU's gender gap persists in the studio environment. Figure 3.3 shows the pre-test and post-test gender gaps for conventional and studio students. But unlike their study, this gap is slightly higher on post-test scores. In the conventional student group, the gap is 9.7% on the pre-test and 10.3% on the post-test. In the studio group, the gender gap is 9.3% on the pre-test and 10.7% on the post-test. These differences are not significant. Further analysis concerning the gender gap was performed on the FCI scores. Table 3.8 along with Figures 3.3, 3.4, and 3.5 show the obtained data. It cannot be said that GSU's studio physics reduces or eliminates the gender gap.

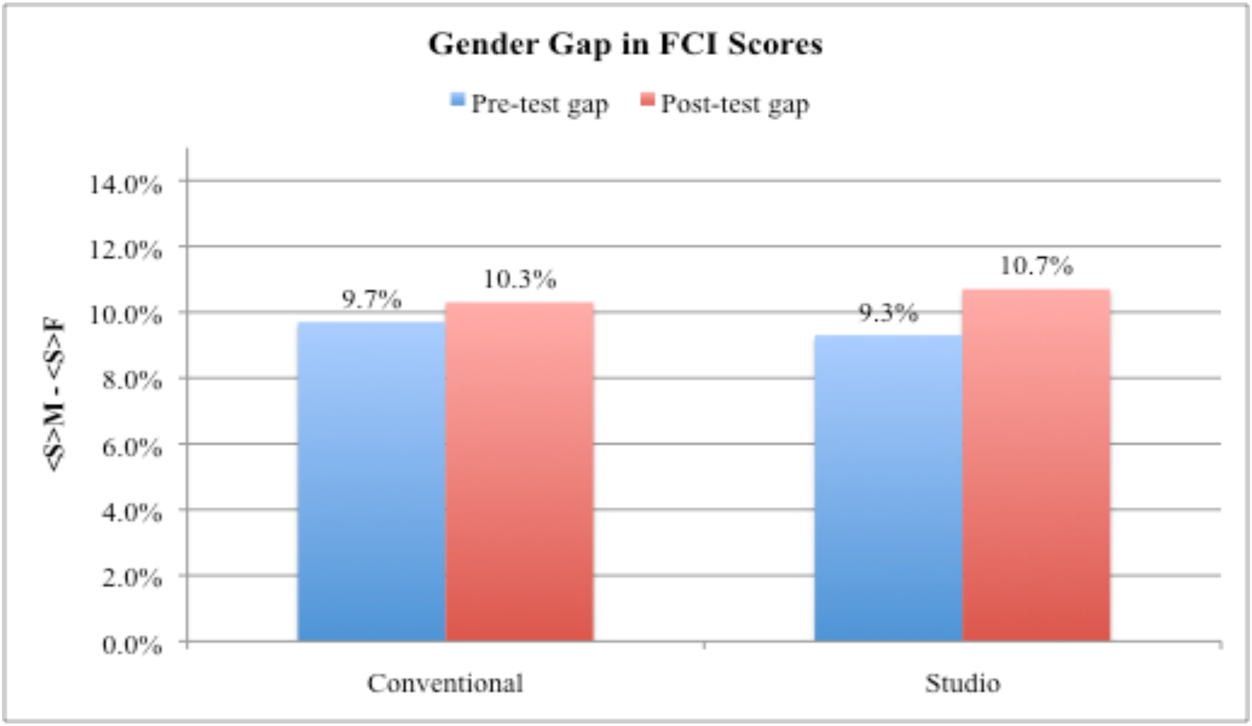


Figure 3.3 Gender gap in FCI scores. Pre-test and post-test gender gaps are shown for conventional and studio students. Student performance is averaged for the Fall 2009 semester (N=257: N_{male}=94, N_{feamle}=163).

Table 3.8 Pre-test and post-test percentage scores per item. Scores are tabulated for male and female students in conventional and studio classes for the Fall 2009 semester.

Conventional				
Item	**F Score (%)**		**M Score (%)**	
#	***Pre***	***Post***	***Pre***	***Post***
#1	44	72	72	92
#2	23	45	34	56
#3	38	38	50	56
#4	20	73	10	62
#5	4	11	8	20
#6	49	92	72	96
#7	54	68	56	74
#8	24	41	36	60
#9	37	34	38	30
#10	18	44	48	62
#11	6	27	18	38
#12	34	59	56	74
#13	1	14	12	42
#14	7	15	28	32
#15	11	48	12	38
#16	37	73	48	74
#17	8	18	0	14
#18	8	24	6	32
#19	38	30	42	50
#20	13	37	26	62
#21	14	14	30	34
#22	21	20	40	52
#23	24	23	46	52
#24	41	65	66	82
#25	11	15	4	22
#26	6	8	8	16
#27	31	44	58	56
#28	14	61	16	56
#29	25	76	36	82
#30	7	28	14	34
Conventional Female, n = 71				
Conventional Male, n = 50				

Studio				
Item	**F Score (%)**		**M Score (%)**	
#	***Pre***	***Post***	***Pre***	***Post***
#1	38	83	61	91
#2	18	45	30	59
#3	20	38	30	59
#4	14	87	11	91
#5	5	41	14	41
#6	35	75	64	91
#7	41	73	52	75
#8	30	45	34	52
#9	30	36	34	39
#10	16	71	48	86
#11	7	37	9	59
#12	43	65	57	86
#13	7	41	7	50
#14	14	43	36	66
#15	23	70	23	80
#16	32	80	52	84
#17	9	35	5	27
#18	1	48	5	68
#19	24	65	25	75
#20	17	53	20	64
#21	17	38	30	52
#22	18	35	16	41
#23	10	32	39	52
#24	48	74	61	86
#25	12	23	14	36
#26	5	13	5	16
#27	36	37	45	64
#28	8	76	14	77
#29	22	79	32	89
#30	12	35	18	45
Studio Female, n = 92				
Studio Male, n = 44				

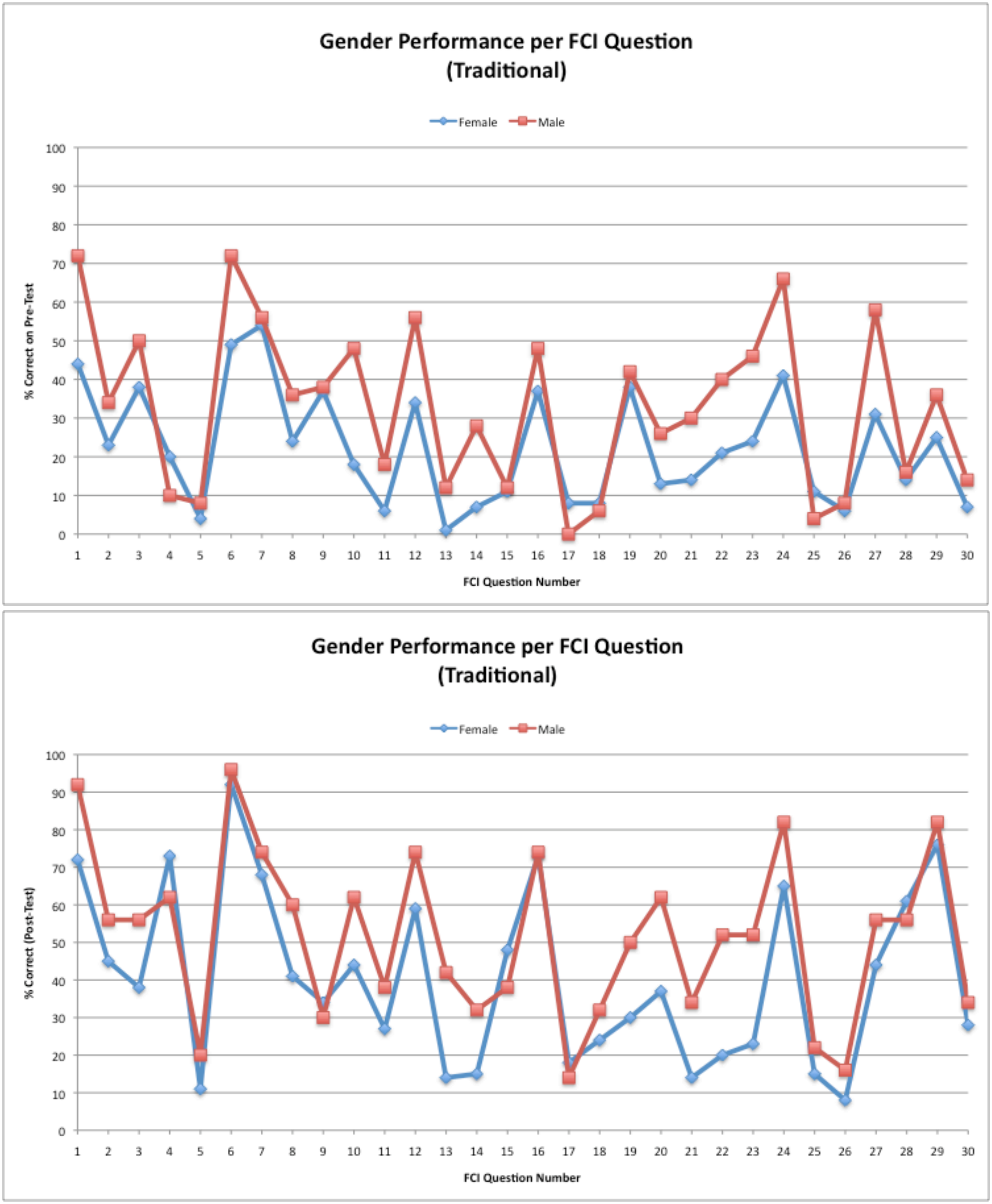


Figure 3.4 Fall 2009 gender performance on the FCI for conventional students (N=121). The top graph shows the percentage of male and female conventional students who got each question correct on the pre-test. The bottom graph shows the post-test scores. The gender gap varies per question, although males achieve higher FCI scores overall.

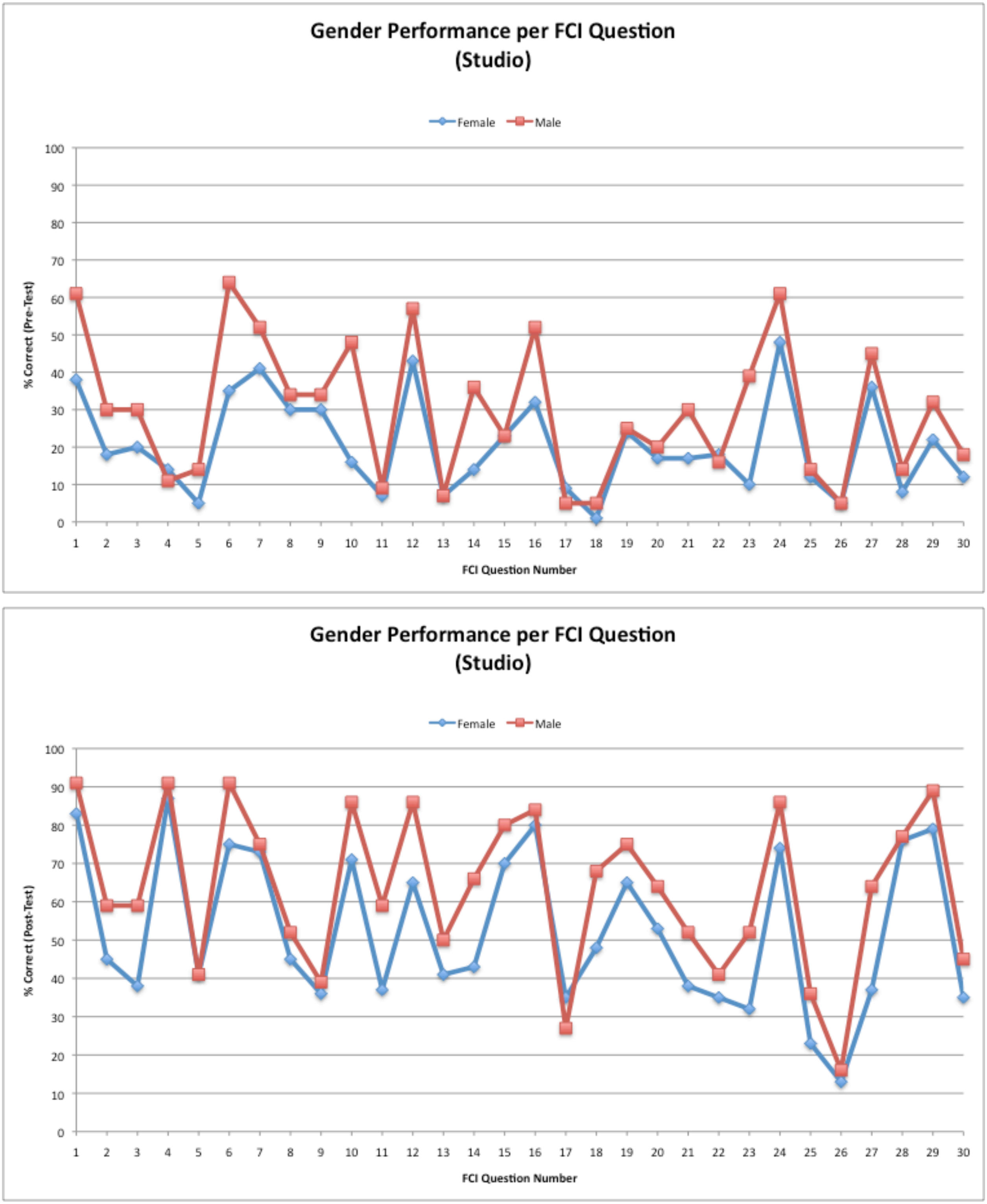


Figure 3.5 Fall 2009 gender performance on the FCI for studio students (N=136). The top graph shows the percentage of male and female studio students who got each question correct on the pre-test. The bottom graph shows the post-test scores. The gender gap varies per question, although males achieve higher FCI scores overall.

Although there is clearly an overall gender gap in pre-test FCI scores that persist in post-instruction, there is evidence that the gap is eliminated on certain items (questions) on the FCI. Table 3.9 tabulates the results of the percentage of students in conventional and studio classes that got each item correct.

Table 3.9 Fall 2009 item analysis by topic. This table shows the pre-test and post-test percentage scores for conventional and studio students. Results show percentage of students in the two samples that got each item correct.

Topic	**Item #**	**Conventional (n = 121)**		**Studio (n = 136)**	
		Pre	***Post***	***Pre***	***Post***
Kinematics	#1	55	80	46	85
	#2	27	50	22	49
	#14	16	22	21	51
	#19	40	38	24	68
	#20	18	47	18	57
1st Law	#6	59	93	44	80
	#7	55	70	45	74
	#8	29	49	32	47
	#10	31	51	26	76
	#12	43	65	48	72
	#21	21	22	21	43
	#23	33	35	19	38
2nd Law	#3	43	45	23	45
	#9	37	32	32	37
	#22	29	33	18	37
	#24	51	72	52	78
	#25	8	18	13	27
	#26	7	12	5	14
	#27	42	49	39	46
3rd Law	#4	16	69	13	88
	#15	12	44	23	73
	#16	41	74	38	82
	#28	15	59	10	76
Force Identification	#5	6	15	8	41
	#11	11	31	7	44
	#13	6	26	7	44
	#17	5	17	7	32
	#18	7	27	2	54
	#29	30	79	25	82
	#30	10	31	14	38

The gender item analysis revealed areas where the gender gap remained unaffected; but particularly, it revealed an area in which the gender gap is nearly eliminated. Based on Figure 3.5, nearly the same number of male and female students gets item numbers 4, 16, and 28 correct. Table 3.9 is used to match these item numbers with the corresponding topic. Those item numbers represent Newton's third law concepts. Evidently, studio physics is successful in decreasing the gender gap in FCI post-test scores on the topic of Newton's third law. While no other areas on the FCI seem to be affected, this study indicates that studio physics techniques can be helpful for instructors, both in the conventional classroom and the studio physics classroom, in assisting students to grasp concepts, especially as it relates to Newton's third law. Future studies are needed as the sample size increases to determine statistical significance and to determine other topics that yield potential differences.

4 CONCLUSION

4.1 Summary

This study demonstrates the effects of studio physics on student conceptual understanding of Newtonian concepts. Conventional and studio students were no different in their knowledge and understanding of basic physics concepts coming into the course. But the normalized learning gain in the FCI scores clearly shows that the conceptual understanding of students in the studio physics class is better than that of the students in the conventional class. As students are involved in the interaction and integrative methods employed in the studio physics classes, their conceptual understanding of Newtonian physics is increased between their pre-tests and post-tests. These findings indicate that a difference in approach is not simply a faddish idea that is innovation for innovation's sake. This study supports the findings of physicists all over the country, as referenced in Chapter 1: an interactive, collaborative, activity-based learning environment yields significantly higher learning gains than the traditional, didactic approach to teaching.

The current study revealed evidence that studio physics is an effective instructional method in an algebra-based introductory physics course at an urban institution. This work supports the findings of previous studies in Physics Education Research. The Force Concept Inventory has shown that student conceptual understanding of basic physics concepts is increased. This instrument will be used in a continuous manner to evaluate and improve the effectiveness of instructional strategies. Secondly, this study revealed that studio physics does play a small role in lessening the gender gap in FCI scores on Newton's third law items more than conventional physics. While not offering suggestions or drawing conclusion related to studio physics and female student performance, this study does suggest that studio physics techniques relating to integrative, collaborative, community-based learning methods can inform instructors on places within existing introductory physics courses where these methods can be used appropriately to increase student conceptual understanding. Thirdly, this pilot study was able to see some differences in the performance and conceptual grasp of physics concepts among several ethnic groups that GSU serves in its introductory physics courses. This research has shown that students exposed to studio physics, and in a larger sense, the methodologies of PER on a broad scale, increase in their understanding and application of basic concepts of physics. Increasingly, as the generational values of the students who are entering universities are rejecting authority and traditional approaches, studio physics is an effective way to meet these students "where they are" epistemologically and to see them gain mastery in those basic concepts in the contexts of community and in a constructivist framework.

4.2 Implications and Limitations

The implication of finding a difference in learning gain with studio physics is that Georgia State University can and should continue to offer the studio approach to teaching physics. As students better grasp the concepts and gain some sense of accomplishment and mastery, the implication is that studio physics can help to serve a certain demographic of student who would be underserved in the conventional classroom.

One limitation in the current study that should be considered is that only an algebra-based course, Physics 1111, has been investigated. This research adds to only a small amount of research that has been done in algebra-based physics instruction. The findings of this study will be helpful in expanding other PER investigations at the undergraduate level and lower. While that body of knowledge is still underdeveloped and under-researched, in order to expand the research efforts of GSU's Physics Education Group, studio physics needs to be implemented in Physics 2211, the calculus-based course. In a contextual and longitudinal sense, GSU and the Physics department would greatly benefit from studying how studio physics could be used in terms of student mastery at that level. Could interactive engagement techniques succeed at GSU in a calculus-based course? Would the interaction among students yield gain in a course that is more in-depth and detailed? The resulting research could even be used as a tool for identifying and recruiting students to pursue physics as major, both on the undergraduate and graduate levels, based upon a set of criteria that could be developed from student performance on some advanced form of an analysis tool.

Another limitation is that studio physics has lent itself to significant gains to student learning, retention, and performance among students at other institutions; however, it must be pointed out that these other institutions have embraced PER and its innovative and alternative methods at the institutional and departmental levels. When those schools and their physics departments embrace PER methods in the sense of funding, staffing, scheduling, and space allocations, they have been able to capitalize upon the gains that they have found. Similarly, as GSU's physics department as a whole continues to embrace PER philosophically, the research shows that students can be better serviced. As the methodologies are pursued in such a way as to have an impact on its undergraduate and graduate students within the department and across the other disciplines that take the courses, GSU could greatly improve student appreciation for and application of basic physics concepts.

4.3 Future Work

The results of this study are just the beginning of a wide range of work to be done in Physics Education Research at GSU. The opportunity is a vast one and an exciting one. This pilot study has given us an array of questions to yet be answered. As the work continues and the sample size increases by semester, we plan to gain more details about GSU's unique demographic and its performance in introductory physics.

One of the most significant results of this study is the low average normalized gains of the Asian students. An investigation into this phenomenon, of whether the difference in Asian FCI scores is due to a language barrier, needs to be performed. We speculate that the Asian students in the study are not native English speakers. We plan to incorporate an additional question on the demographic survey that identifies each student's native language.

Additionally, we hope to gain more information on what impact students' prior exposure to physics may have on the learning gains. The plan is to do correlation studies on math preparation, college major, and prior physics knowledge with learning gains. These results will add another dimension to the study.

Developing a learning gain assessment in the second semester of introductory physics (Phys 1112 – Electricity & Magnetism) is another item on the research agenda. An Electricity & Magnetism pre-assessment instrument is needed in the PER community. GSU's PER group plans to investigate the development and validation of such a tool.

REFERENCES

Bao, L. (2006). Theoretical comparison of average normalized gain calculation. *American Journal of Physics*, 74, 917-922.

Beichner, R. J., Saul, J. Mm, Abbott, D. S., Morse, J., Deardorff, D., Allain, R. J., Bonham, S. W. Bonham, Dancy, M., & Risley, J. (2007). Student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. *Reviews in PER: Research-Based Reform of University Physics,* 1, 1-42.

Buck, J. R., and Wage, K. E. (2005). Active and Cooperative Learning in Signal Processing Courses. *IEEE Signal Processing Magazine*, 22(2), 76-81.

Cardinal, R. N. & Aitken, M. R. F. (2005). ANOVA for the behavioral science researcher. Mahwah, NJ: Lawrence Erlbaum Associates.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 5th ed. Hillsdale, NJ: Lawrence Earlbaum Associates.

Crotty, M. (2003). *The Foundation of Social Research: Meaning and Perspective in the Research Process*.

Docktor, J, & Heller, K., (2008). Gender differences in both force concept inventory and introductory physics performance. *Physics Education Research Conference*, 1074, 15-18.

Evans, D. & Hestenes, D. (2001). The concept of the concept inventory assessment instrument. *Proceedings of the Frontiers in Education Conference, USA, 2,* F2A-F21.

Gaffney, J. D., Richards, E., Kustusch, M., Ding, L., & Beichner, R. (2008). Scaling up education reform. *Journal of College Science Teaching*, 37, 5, 48-53.

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

Hestenes, D. (1998). Who needs physics education research? *American Journal of Physics,* 66(6), 465-467.

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

Huck, S. W. (2008). *Reading Statistics and Research*, 5th ed. Boston: Pearson Education, Inc.

Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative Learning Returns to College: What Evidence Is There That It Works? *Change*, 30(4), 26–35.

Jovanovic, J. & Dreves, C. (1995). Math, science, and girls: Can we close the gender gap?. In Todd, C.M. (Ed.). *School-age connections*, 5(2), University of Illinois Cooperative Extension Service: Urbana, IL.

Kahle, J. & Meece, J. (1994). Research on gender issues in the classroom. *Handbook of Research on Science Teaching*. Macmillan: New York, 542-557.

Kost, L., Pollock, S., & Finkelstein, N. (2009). Characterizing the gender gap in introductory physics. *Physical Review Special Topics – Physics Education Research*, 5, 010101-1—010101-14.

Loepz, R. & Schultz, T. (1991). Two Revolutions in K-8 Science Education. Physics Today, 54(9), 44-49.

Lorenzo, M., Crouch, C., & Mazur, E. (2006). Reducing the gender gap in the classroom. *American Journal of Physics*, 76(2), 118-122.

Marx, J. & Cummings, K. (2007). Normalized change. *American Journal of Physics*, 75, 87-91.

Maxwell, S. & Delaney, H. (2004) *Designing Experiments and Analyzing Data: A Model*

Comparison Perspective, 2nd ed. Mahwah, New Jersey: Lawrence Earlbaum Associates, Inc.

Mazur, E. (1997). *Peer Instruction: A User's Manual*. Upper Saddle River, New Jersey: Prentice Hall, Inc.

McDermott, L. C. (1993). Guest comment: How we teach and how students learn – A mismatch? *American Journal of Physics,* 61(4), 295-298.

McDermott, L. C. (2001). Oersted Medal Lecture 2001: "Physics Education Research – The key to student learning". *American Journal of Physics,* 68(11), 1127-1137.

McDermott, L. C. & the Physics Education Group at the University of Washington (1996). *Physics by Inquiry*.

McDermott, L. C., Shaffer, P. S., & the Physics Education Group at the University of Washington (2002). *Tutorials in Introductory Physics*.

Myers, J. & Well, A. (2003). *Research Design and Statistical Analysis*, 2nd ed. Mahwah, New Jersey: Lawrence Earlbaum Associates, Inc.

Phillips, Melba (1977). Founder of AAPT: Paul Klopsteg. *The Physics Teacher*, 15, 4, 212-214.

Pollack, S., Finkelstein, N., & Kost, L. (2007). *Reducing the gender gap in the physics classroom: how sufficient is interactive engagement?* Physical Review Special Topics – Physics Education Research, 3, 010107-1--010104-4.

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education*, 93(3), 223-231.

Redish, E. F. (1999). Millikan Lecture: Building a Science of Teaching Physics, *American Journal of Physics*, 67, 562-573.

Redish, E. F., Saul, J. M., & Steinberg, R. N. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66(3), 212-224.

Sahin, M. (2010). Effects of problem-based learning on university students' epistemological beliefs about physics and physics learning and conceptual Understanding of Newtonian mechanics. *Journal of Science Education and Technology*, 19, 266-275.

Saul, J. M., Abbott, D. S., Parker, G. W. & Beichner, R. J. (2000). Can one lab make a difference? *American Journal of Physics, Physics Education Research Supplement*, 68(S1), S60-S61.

Savinainen, A. & Scott, P. (2002). The Force Concept Inventory: a tool for monitoring student learning. *Physics Education*, 37(1), 45-52.

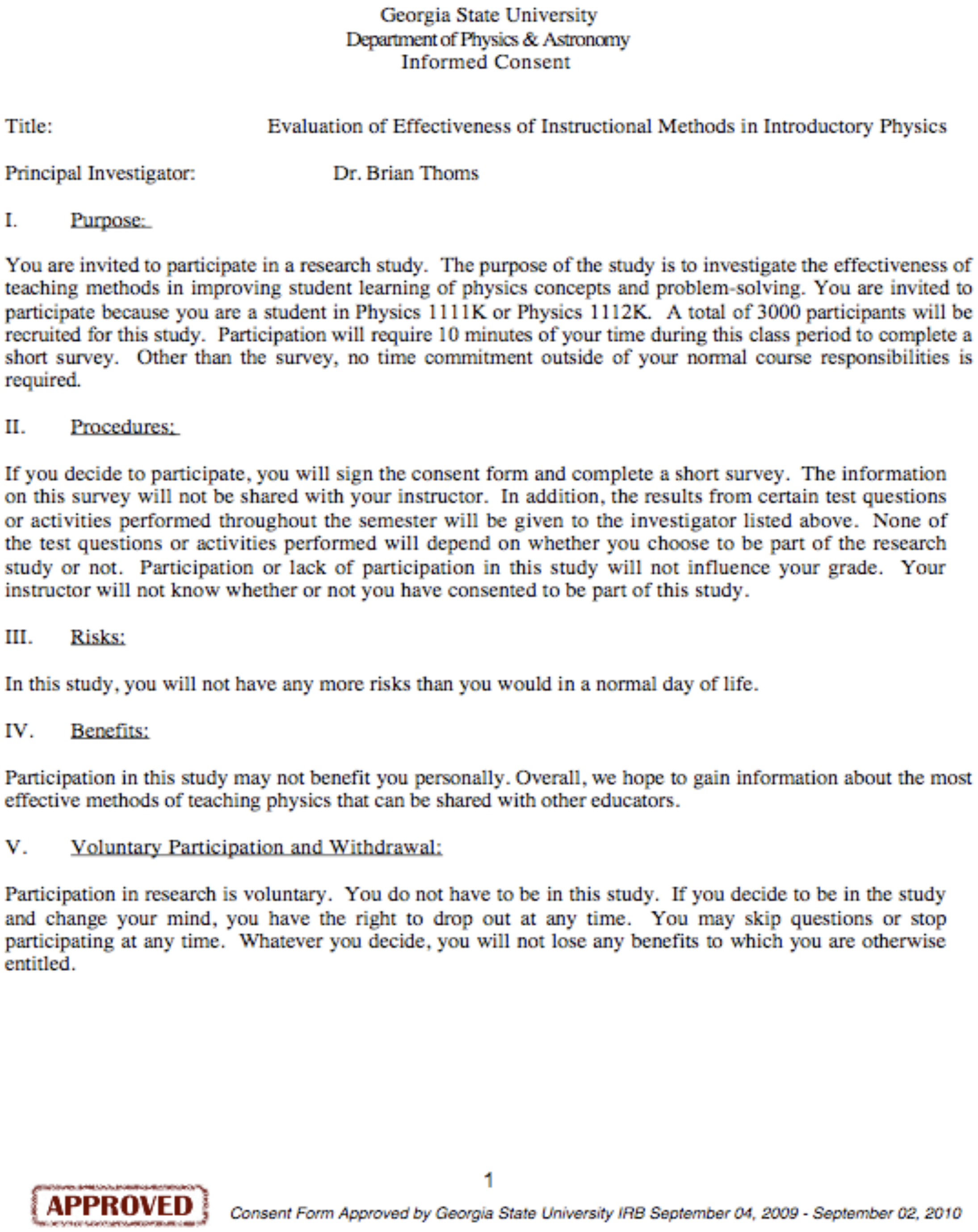
Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research*, 69(1), 21-51.

Wilhelm, J., Thacker, B, & Wilhelm, R. (2007). Creating constructivist physics for introductory university classes. *Electronic Journal of Science Education*, 11, 19-29.

Willoughby, S. & Metz, A. (2009). Exploring gender differences with different gain calculations in astronomy and biology. *American Journal of Physics*, 77(7), 651-657.

Wilson, J. (1994). The CUPLE Physics Studio. *The Physics Teacher*, 32, 518-523.

APPENDIX A



Georgia State University
Department of Physics & Astronomy
Informed Consent

Title: Evaluation of Effectiveness of Instructional Methods in Introductory Physics

Principal Investigator: Dr. Brian Thoms

I. Purpose:

You are invited to participate in a research study. The purpose of the study is to investigate the effectiveness of teaching methods in improving student learning of physics concepts and problem-solving. You are invited to participate because you are a student in Physics 1111K or Physics 1112K. A total of 3000 participants will be recruited for this study. Participation will require 10 minutes of your time during this class period to complete a short survey. Other than the survey, no time commitment outside of your normal course responsibilities is required.

II. Procedures:

If you decide to participate, you will sign the consent form and complete a short survey. The information on this survey will not be shared with your instructor. In addition, the results from certain test questions or activities performed throughout the semester will be given to the investigator listed above. None of the test questions or activities performed will depend on whether you choose to be part of the research study or not. Participation or lack of participation in this study will not influence your grade. Your instructor will not know whether or not you have consented to be part of this study.

III. Risks:

In this study, you will not have any more risks than you would in a normal day of life.

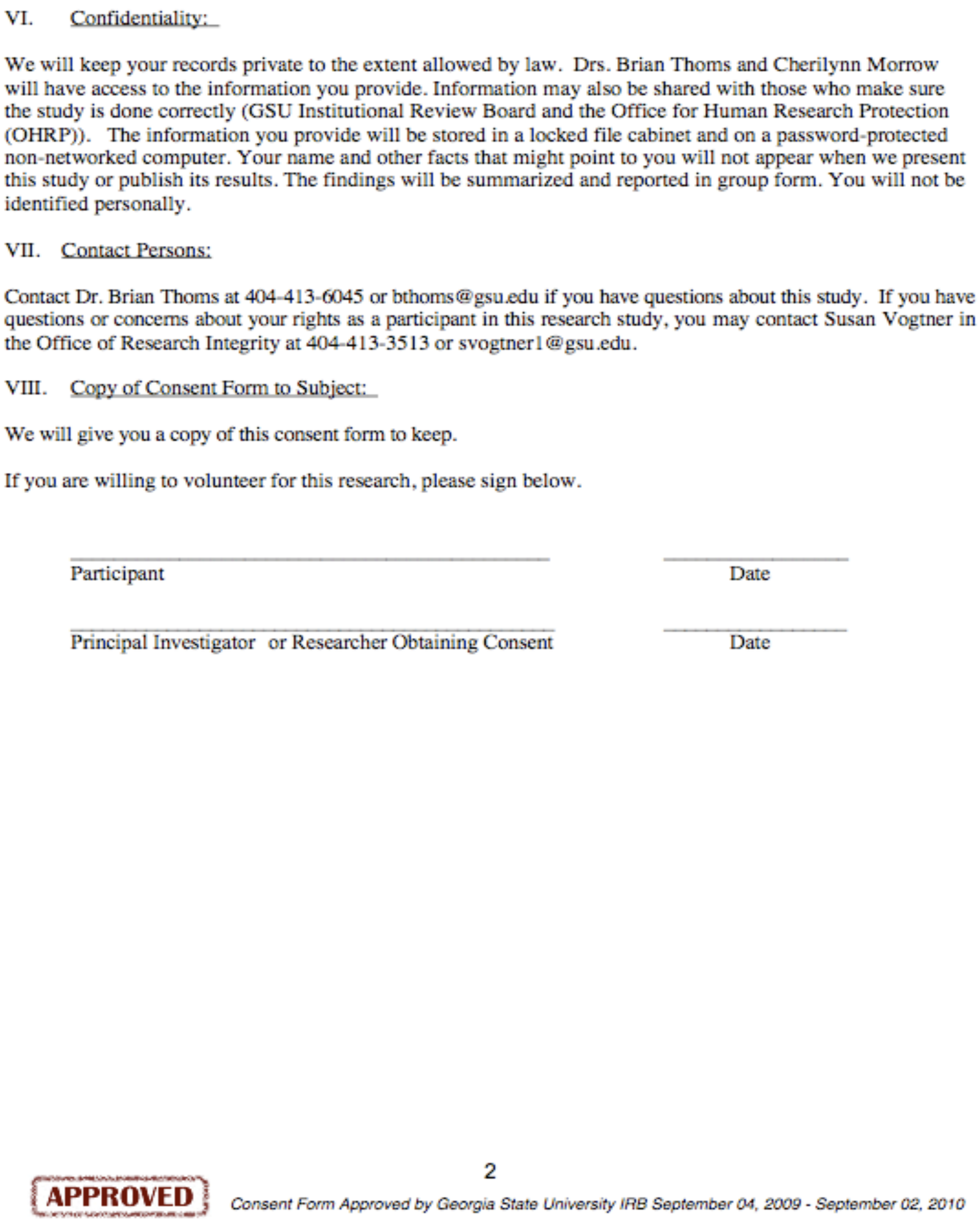
IV. Benefits:

Participation in this study may not benefit you personally. Overall, we hope to gain information about the most effective methods of teaching physics that can be shared with other educators.

V. Voluntary Participation and Withdrawal:

Participation in research is voluntary. You do not have to be in this study. If you decide to be in the study and change your mind, you have the right to drop out at any time. You may skip questions or stop participating at any time. Whatever you decide, you will not lose any benefits to which you are otherwise entitled.

APPROVED *Consent Form Approved by Georgia State University IRB September 04, 2009 - September 02, 2010*



VI. Confidentiality:

We will keep your records private to the extent allowed by law. Drs. Brian Thoms and Cherilynn Morrow will have access to the information you provide. Information may also be shared with those who make sure the study is done correctly (GSU Institutional Review Board and the Office for Human Research Protection (OHRP)). The information you provide will be stored in a locked file cabinet and on a password-protected non-networked computer. Your name and other facts that might point to you will not appear when we present this study or publish its results. The findings will be summarized and reported in group form. You will not be identified personally.

VII. Contact Persons:

Contact Dr. Brian Thoms at 404-413-6045 or bthoms@gsu.edu if you have questions about this study. If you have questions or concerns about your rights as a participant in this research study, you may contact Susan Vogtner in the Office of Research Integrity at 404-413-3513 or svogtner1@gsu.edu.

VIII. Copy of Consent Form to Subject:

We will give you a copy of this consent form to keep.

If you are willing to volunteer for this research, please sign below.

Participant _______________ Date _______________

Principal Investigator or Researcher Obtaining Consent _______________ Date _______________

2

APPROVED Consent Form Approved by Georgia State University IRB September 04, 2009 - September 02, 2010

APPENDIX B

Information Survey

Name____

What is your present school year status?
__ Freshman
__ Sophomore
__ Junior
__ Senior
__ Post-bac

What is your major?
____

Have you taken a Physics class before and if so where?
__ Never
__ High School
__ College/University

Have you taken Pre-calculus and if so where?
__ Never
__ High School
__ College/University

Have you taken Calculus and if so where?
__ Never
__ High School
__ College/University

What is your present age?
______

What is your sex?
__ Male
__ Female

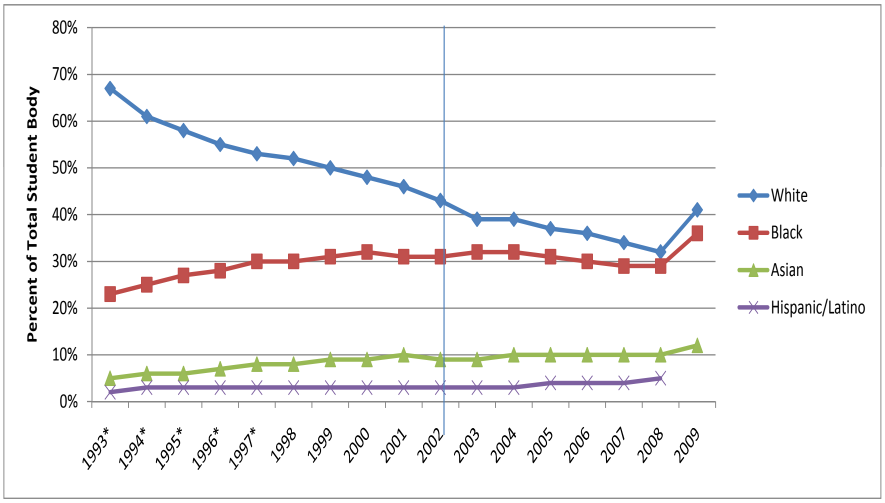
What is your race/ethnicity? Mark one or more.
__ American Indian or Alaska Native
__ Asian
__ Black or African American
__ Hispanic or Latino
__ Native Hawaiian or Other Pacific Islander
__ White

APPENDIX C

Georgia State University

Undergraduate Students by Race -- Fall Term (1993 - 2009)

	1993*	1994*	1995*	1996*	1997*	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
White[1]	67%	61%	58%	55%	53%	52%	50%	48%	46%	43%	39%	39%	37%	36%	34%	32%	41%
Black	23%	25%	27%	28%	30%	30%	31%	32%	31%	31%	32%	32%	31%	30%	29%	29%	36%
Asian	5%	6%	6%	7%	8%	8%	9%	9%	10%	9%	9%	10%	10%	10%	10%	10%	12%
Hispanic/Latino[2]	2%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	4%	4%	4%	5%	
American Indian	0.2%	0.3%	0.2%	0.3%	0.3%	0.3%	0.3%	0.2%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%	0.3%	0.2%	0.4%
2 or More Races[3]																	4%
Multiracial[4]	0%	3%	3%	4%	4%	4%	4%	5%	8%	5%	5%	4%	4%	3%	3%	3%	
Other[4]										6%	4%	3%	2%	2%	2%	2%	
Not Reported										0.0%	4%	6%	9%	12%	14%	16%	6%
Non-Res Alien[5]	2%	3%	3%	3%	2%	2%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%	3%
	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
N=	16,786	16,679	16,849	16,320	16,828	15,666	16,303	16,439	18,245	19,683	20,208	19,894	19,004	19,122	19,904	20,847	22,385



Data source: 2009 and forward IPORT - Enrollment by Race; 2008 and earlier Statware 6100 (146)

* Learning support students included in 1993-1997 figures.

[1] Prior to 2002, Other and Not Reported were included with White.

[2] Due to a change in federal reporting requirements, effective in 2009 a student can choose to report separately ethnic and racial identity.

[3] "2 or more Races" is a new federal race category effective 2009.

[4] Effective 2009 "multiracial" and "other" are no longer included in the list of federal racial reporting categories.

[5] Any non-U.S. Citizen who currently lives in the U.S. with a student visa.

APPENDIX D

Georgia State University
Fall 2009 Demographic Overview of Students

Race of GSU Students

	Undergraduate	Graduate
White	41%	59%
Black	36%	19%
Asian	12%	5%
2 or More Races	4%	2%
American Indian	0.4%	0.2%
Not Reported	6%	4%
Non-Resident Alien	3%	11%

Ethnicity of GSU Students

	Undergraduate	Graduate
Hispanic	7%	5%
Non-Hispanic	85%	90%
Not Reported	8%	5%

Average Age

overall	26
graduate	32
undergraduate	24
first-time freshmen	18

Other Demographics of GSU Students

Women	61%
Graduate	26%
GA resident	90%
International	12%
U.S. out-of-state	10%

GSU has 3,604 international [1] students from 154 countries of origin

Top 10 Countries of Origin for International[1] Students

	Undergraduate	Graduate	Total
India	196	228	424
China	91	289	380
South Korea	216	62	278
Canada	93	41	134
Vietnam	124	6	130
Jamaica	106	24	130
Nigeria	92	22	114
Colombia	79	32	111
UK	69	24	93
Mexico	71	12	83

---

[1] International students include all non-U.S. citizen resident aliens and non-resident aliens.