Do Students in General High School Physics Classes Learn as Much from Virtual Labs as from Hands-On Labs?

Jeanne Finstein
Marjorie Darrah
Roxann Humbert

Abstract: The dramatic increase in the number of students enrolled in high school physics courses, both in traditional settings and on-line, has resulted in difficulties in providing adequate lab experiences to reinforce content. This study was conducted to investigate whether or not students learn as much using virtual lab experiences as they do when conducting experiments with hands-on equipment. The study utilized a variety of statistical methods to analyze student achievement in students in three high schools. Results indicated that learning was comparable between students using the virtual labs and those using hands-on. Further study is planned to determine if students using virtual labs as supplements to hands-on show increased learning over those who use virtual or hands-on alone.

About the authors: *Dr. Jeanne Finstein is president of Polyhedron Learning Media, Inc., an educational software development company. She has decades of experience as a high school teacher and as a developer of software for projects funded by NASA and the U.S. Department of Education. Dr. Marjorie Darrah has over fifteen years of experience in higher education and K-12 outreach programs. She has expertise in the implementation, development and evaluation of educational technologies for the classroom. Dr. Roxann Humbert is Co-founder and Principal Evaluator for ProEvaluators, LLC, an educational evaluation and consulting firm. Her areas of expertise include educational software evaluation, online learning, and testing and measurement.*

Keywords: virtual lab, physics, computer-assisted, technology, integration, secondary education

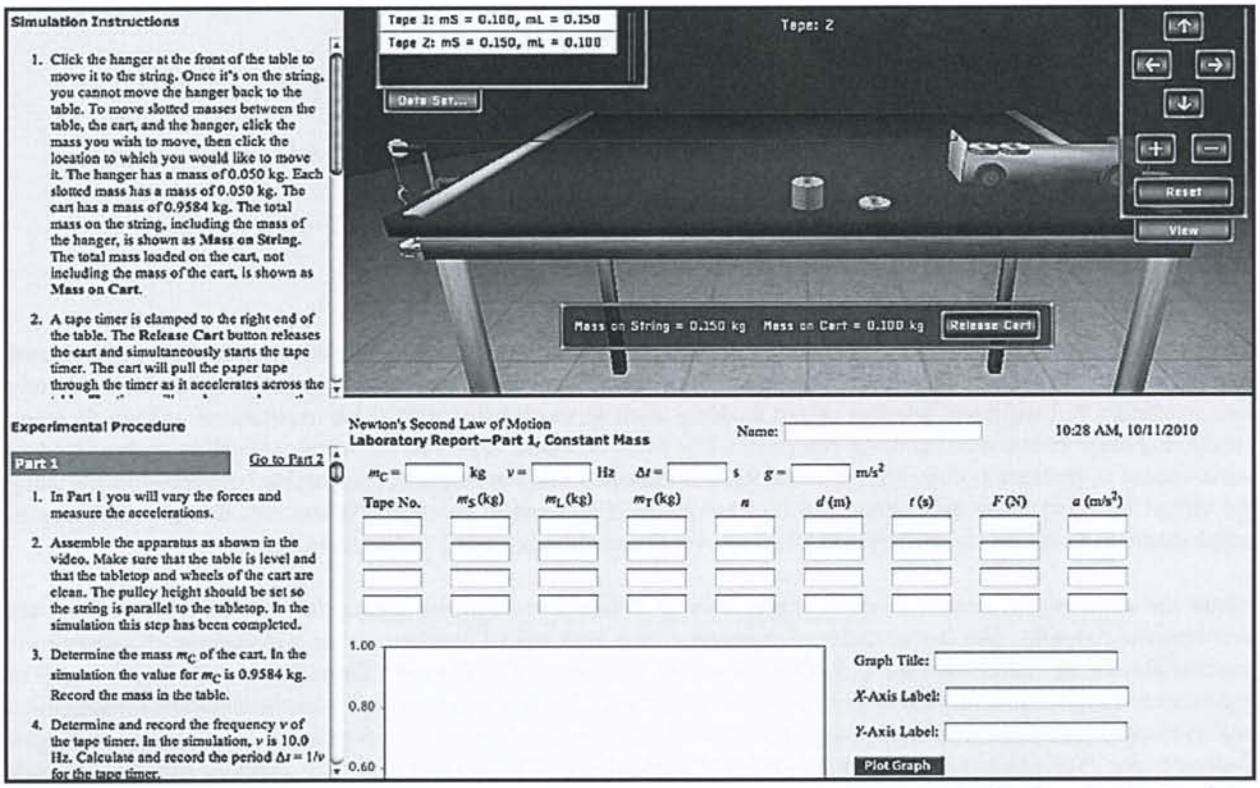
INTRODUCTION

Data show that the number of high school students taking physics nationwide has increased dramatically in the past 25 years (AIP, 2009). For these students, meaningful laboratory experiences are necessary to demonstrate and reinforce physics concepts.

High School Virtual Physics Lab was developed to address the growing numbers of physics students and the parallel struggle of school systems to provide suitable lab experiences. The virtual labs can serve as replacements for hands-on when equipment is not available or when a hands-on lab is potentially dangerous or too time-consuming for the class period, or as supplements to hands-on labs to introduce or reinforce concepts. They are also a suitable option for students in distance learning situations.

A yearlong set of labs was developed through funding from the Small Business Innovation Research (SBIR) program of the Institute of Education Sciences in the U.S. Department of Education (contract # ED-IES-11-C-0029). The 24 labs are available online or as iPad apps. Each lab includes background theory, a pre- and post-lab quiz, one or more video segments showing students conducting the experiment using hands-on equipment, a 3D simulation, post-lab thought questions, and teacher support materials. The simulations are the primary components of each virtual lab. They are not merely demonstrations; instead they allow students to manipulate equipment, gather and then analyze realistic data, prepare lab reports, and reach conclusions based on their own data and graphs.

The purpose of this study was to determine if the virtual labs are as effective as hands-on in relaying physics concepts



Newton's Second Law of Motion. This sample screen shows a cart pulled by a mass suspended below a pulley. Students vary the amount of mass and determine the change in acceleration. They enter data and the results of calculations in the provided tables.

THEORETICAL FRAMEWORK: RESEARCH QUESTION

More than two decades after the release of *A Nation at Risk: The Imperative for Education Reform*, the Business-Higher Education Forum published *A Commitment to America's Future: Responding to the Crisis in Mathematics and Science Education*. This report contends that, "Increased global competition, lackluster performance in mathematics and science education, and a lack of national focus on renewing its science and technology infrastructure have created a new economic and technological vulnerability as serious as any military or terrorist threat." (BHEF, 2005)

As the country has responded to this report and other calls to action, greater emphasis has been placed on higher-level math and science courses at the high school level. Notably, enrollment in science courses in general has been on the rise as graduation requirements have increased. (CCSSO, 2005) Additional statistics show that enrollment in physics in particular has paralleled the rise in enrollment in science courses in general. (Owings, 2008) As enrollment in high school physics courses has increased, the demand for adequate laboratory experiences has also increased. Laboratory work is seen as an essential part of the learning process, enabling students to interact with natural phenomena and analyze collected data. A position statement by the National Science Teachers Association describes the importance of the lab experience. "Throughout the process, students should have opportunities to design investigations, engage in scientific reasoning, manipulate equipment, record data, analyze results, and discuss their findings. These skills and knowledge, fostered by laboratory investigations, are an important part of inquiry—the process of asking questions and conducting experiments as a way to understand the natural world." (NSTA, 2007) Funding adequate laboratories for the ever-increasing number of physics students has become a challenge for school systems. The American Institute of Physics (AIP) reports that laboratory science is a "high-priced luxury" that far too many schools cannot afford. The group cites General Accounting Office statistics that 42% of schools surveyed reported being "not well at all" equipped for laboratory science, with even larger percentages of low-income schools not well equipped. (AIP, 2007) The Arkansas Science Teachers Association (ASTA) spent a year developing a description of what constitutes adequate science classrooms, labs, and equipment. That group estimated that the cost of equipment for a single physics lab that meets learner expectations is $40,000 - $55,000, with an additional annual replacement/repair budget of $2000 - $3000. (ASTA, 2008) This is quite a large outlay, especially for schools in low-income areas or those with small student enrollments. The AIP report points out that "Some teachers can pull together … materials and organize them into a coherent curriculum, but most have neither the time nor the capacity." (AIP, 2007)

One answer is the growing use of technology to replace and/or supplement hands-on labs. The National Education Association (NEA) and the American Federation of Teachers (AFT), as advocates for educators, firmly believe that access, adequacy, and equitable distribution of technology across schools and classrooms is critical for educators to prepare their students for success in this changing global society. These organizations further believe that teachers should be prepared to use technology to deliver alternative types of pedagogy, such as inquiry learning, models, and simulations to help students develop higher-order thinking skills. (NEA, 2008)

To meet the growing need, in some cases technology-based distance learning is replacing face-to-face instruction. As a study from the National Center for Education Statistics reports, "During the 12-month 2004–05 school year, 37 percent of public school districts had students in the district enrolled in technology-based distance education courses. This represents an estimated 5,670 of a total 15,190 public school districts in the country… Technology-based distance education courses are considered the future of distance education offerings, with online technologies looked upon by some policymakers as offering the greatest promise." (Zanberg, 2008) According to a report by the North American Council for Online Learning (NACOL), "As of September 2007, 42 states [had] significant supplemental online learning programs (in which students enrolled in physical schools take one or two courses online), or significant full-time programs (in which students take most or all of their courses online), or both." (Watson, 2007) And needless to say, students in on-line courses would not have access to a full physics lab for the laboratory component of the course.

Given this need for physics lab experiences at the high school level along with the growing use of technology, Polyhedron Learning Media, Inc. developed a set of high-quality, content rich, cost-effective, online simulations – *High School Virtual Physics Lab* (VPL) – that can serve as replacements and/or supplements in either in-school or distance learning settings. This study was initiated to assess student learning while using VPL to investigate the question: *Do students who complete the virtual labs learn as much as students who complete the traditional hands-on labs?*

METHODOLOGY

Research Design and Results – Preliminary Studies

Before the target question was addressed, formative studies of usability and feasibility were conducted to ensure that the virtual lab content was comparable to that in hands-on labs and that the labs were accessible in standard school settings. Six "Teacher Experts" and a class of "Student Experts" reviewed each virtual lab that was to be used in Round 1 of the study to assess usability. The "Teacher Experts" consisted of three high school physics teachers and three college physics teachers. The "Student Experts" consisted of five high school Advanced Placement Physics students who were in their second year of high school physics classes.

Both groups of "Experts" completed an online survey consisting of Likert scale statements with accompanying open-ended questions to gather additional comments. The statements in the teacher survey focused on content for the college level experts and both functionality and content for the high school teacher experts. These reviews revealed items such as interface weaknesses or pedagogically undesirable features and were used to resolve interface questions and uncover subtle programming problems. Subsequent revisions were made, and if they were extensive, further reviews were conducted. Improvements to the lab simulations were made based on these reviews.

Next, the "Student Experts" reviewed each lab simulation while the developers observed. The students then used an online survey to report on their experience. Comments from the online surveys were compiled and reviewed by the development team, who then made any necessary revisions.

Teacher Experts reported that the labs were easy to use and would fit well in the classroom schedule. The system requirements were accurate, and using the labs would not require an undue burden on the teacher. The teachers could see many practical uses for the labs and made some suggestions for improvements to them. Those improvements were made, when possible. Student Experts reported that they found the labs were well organized, the directions were clear, the graphics were realistic, and the lab content was very similar to what they had done in the same hands-on lab in Physics I.

Research Design – Primary Study, Round 1

Four virtual labs were integrated into the regular classroom activities during the first half of the 2011-2012 school year (September – November). Several data sources related to lab usability, feasibility, and student learning were collected during classroom testing. These measures included observations of the students as they completed the labs, video clips of the students while they completed the labs, collection of pre/post lab quiz data, and survey results from an online survey that asked students questions about their experience using the labs. Also, to inform the development team about classroom practicality, a time log was kept to determine if the virtual lab could be completed in the allotted time. Finally, a teacher interview was conducted to determine the practicality of using the labs in a high school setting. To test student learning, the Force Concept Inventory was administered to all students at the beginning of the school year. The grades on this assessment were used as a baseline indicator of the knowledge each student had as he/she entered the physics course. Before beginning each virtual lab, the students took a pre-quiz. A post-quiz was administered after each virtual lab was completed.

To ensure content validity the developers created a test blueprint of the key concepts taught in the lab. *Validity* refers to the accuracy of an assessment – whether or not it measures what it is supposed to measure. (Gronlund, 2008) "A very important first step in developing a summative assessment is to define the domain of assessment. The domain of assessment represents the scope and depth of test coverage; it is directly derived from the defined goals and objectives for the unit. One way to define the domain of assessment is to use a two-dimensional table commonly called a test grid or test blueprint." (Liu, 2010) The items for the quizzes were created by the developers based on the blueprint. A panel of experts, consisting of three high school physics teachers and three college-level physics teachers, then reviewed the questions for validity, and the quizzes were finalized.

After all four of the virtual labs had been completed by the students, including pre- and post-lab quizzes, the Force Concept Inventory was again administered to see if the students' overall physics knowledge had increased during this time.

Results – Round 1

Fifty students completed the four virtual labs during this phase. (47 of these agreed to participate in the study and have their data included in the analysis.) Not all students completed all assessments. The largest number of students possible was used in the analysis of each individual assessment.

A two-tailed paired t-test was used to compare each lab's pre-quiz to its post-quiz. The individual t-tests showed that there was a significant increase from the pre- to post-quiz grades for all four labs. A two-tailed paired t-test was also used to compare the pre/post

Figure 1: Regression Analysis of FCI versus Combined Post Lab Quizzes

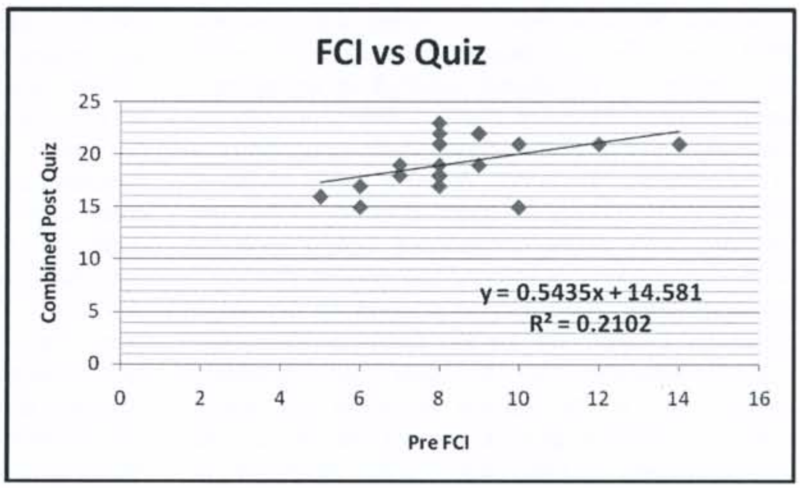


Table 1: Two Tailed t-test Results

Assessment	N	Pre Avg	Post Avg	p-value
Lab 1: Newton's Second Law of Motion	39	5.72	7.10	***p*<.01**
Lab 2: Hooke's Law	30	3.87	4.77	***p*<.01**
Lab 3: Conservation of Energy	40	2.48	3.93	***p*<.01**
Lab 4: Centripetal Force	43	1.81	2.81	***p*<.01**
Force Concept Inventory	44	8.09	10.5	***p*<.01**

Force Concept Inventory. The t-tests showed that there was a significant increase from pre to post on the Force Concept Inventory. Table 1: Two Tailed t-test Results1 below summarizes the results. The significant p-values are highlighted in bold italics. They indicate that the gain in student learning from pre- to post-lab quiz was statistically significant for all four of the labs and also on the Force Concept Inventory.

The post-quiz lab grades were combined to give an overall quiz grade. This grade reflected what knowledge the students acquired after completing the labs. The pre Force Concept Inventory (FCI) was used as the independent variable, or indicator of the knowledge that the students had before completing any of the labs. A regression analysis was completed to determine if there was any relationship between the pre FCI and the post-quiz lab grade. The results are in Figure 1: Regression Analysis of FCI versus Combined Post Lab Quizzes1.

From the trend line it can be seen that there is a positive relationship between the FCI and the combined post-quiz lab grade. Students who did better on the pre FCI (baseline indicator of knowledge) seemed to do a little better on the post quizzes. From the value of R^2 (0.2102), it can be seen that there is not a strong correlation between the FCI and the post-quiz lab score; only about 21% of the variation in the post-quiz lab grade can be explained by the previous knowledge demonstrated on the FCI, therefore 89% is either random or has another explanation. The regression equation for the Combined Post Quiz Lab Grade is POSTQUIZ = 0.544 (PREFCI) + 14.581 with a standard error of the estimate about 2.3 points.

These results indicate that students showed a significant increase in content knowledge after completing the virtual labs and that prior knowledge is not the likely reason for that increase. However, these results don't compare learning using virtual labs with learning using hands-on equipment. Those comparisons were conducted during the next round of the study.

Research Design – Round 2

In this round of testing, the focus was on student learning to answer the primary question of interest, *Do students using the virtual labs learn as much as those using hands-on labs?* The hypothesis was that there would be a positive answer to this question, demonstrating that schools using virtual labs for whatever reason (lack of equipment, student absences, distance learning) could feel confident that their students would have a similar experience to those who had access to actual equipment. During the spring semester of 2012, data were collected from 168 students who agreed to participate in the study. The students were enrolled in high school physics classes in three schools in the states of Florida, Texas, and West Virginia. All students at each school were under the instruction of the same teacher, and each teacher instructed three sections of physics. The classes were assigned at random to one of the following treatments to complete their physics labs, *High School Virtual Physics Lab* (VPL), hands-on lab (HO), or supplemental (SUPP) (using the virtual labs to supplement hands-on labs.) The students' data were partitioned into three groups, VPL, HO and SUPP. Four of the labs were completed by students from all three schools; one additional lab was completed at one school. The VPL and HO groups included students from all three schools. The SUPP group included students from two schools; one school chose to use two class sections as VPL and no SUPP. Students did not switch between sections, so for the semester they were in only one of the treatment groups. Table 2 below summarizes the number of students who completed each lab with the different methods they used to complete the lab.

Here we will note that because one school assigned two classes to the VPL treatment and no class to the SUPP treatment, the number of students in the VPL group was large and the number of students in the SUPP group was small. This was not the original design of the experiment and as will be noted in sections below, caused some problems for drawing conclusions from the data.

Round 2 - Data and data collection methods

Demographic data were collected to determine the similarity of the groups. The following characteristics were collected: grade level, sex, race, and ethnicity. Each student took the Force Concept Inventory as close to the beginning of the school year as possible, given the

Table 2: Number of students who completed labs

Lab	Total	VPL	HO	Supp
Lab 1: Lenses	168	78	55	35
Lab 2: Refraction	164	77	53	34
Lab 3: Ohm's Law	168	78	55	35
Lab 4: Resistors	158	68	55	35
Lab 5: Specific Heat of Metal	53	22	11	20

constraints of the various schools. The FCI score was used as a baseline for student knowledge. The students completed the eight-question pre-test before completing any work for each lab. The students also completed the eight-question post-test after completing their assigned version of the lab. Not all students completed all labs; the largest number of students possible was used in the analysis of each individual lab. A Total Lab Score was computed for all students who completed Labs 1, 2, 3, and 4. This score was the sum of the post-test for all four labs (Total Lab Score = Post-Test Lab 1 + Post-Test Lab 2 + Post-Test Lab 3 + Post-Test Lab 4.) The post-test score for Lab 5 was not included because only one school completed it.

Round 2 - Data analysis methods

Demographic data were collected and analyzed. Then the following methods were used to analyze the test data:

- Paired t-test was used to compare the pre and post-test grades for each lab.
- One-way Analysis of Variance was used to compare the three treatment groups' post-test grades for each lab to determine if there was any difference.
- Regression Analysis was used with the FCI as the independent variable and Total Lab Score as the dependent variable to determine if there was a relationship between the FCI and the Total Lab Score.
- Regression Analysis for each individual treatment group was completed to determine a relationship between FCI and Total Lab Score for each treatment group.
- An Analysis of Covariance was used with to compare differences between the three treatment groups' Total Lab Score, after controlling for Force Concept Inventory scores.

Round 2- Results of Data Analysis
Demographic Data

Demographic data for the three treatment groups are summarized in Table 3: Demographic Data for Treatment Groups below. As noted above the VPL group contained more students from one school than the other two groups. The SUPP group did not contain students from one of the schools. The VPL and HO groups contained students from all three schools, however these groups vary in gender make-up. The average grade level of the three groups is similar.

Two Sample t-test for Pairs of Groups for Force Concept Inventory Score

A two-tailed, two-sample t-test with samples considered to have different variance was used to compare the FCI scores between two different groups. Table 4: Comparing FCI Scores for Two Groups below shows that there was a significant difference between the FCI scores for SUPP

Table 4: Comparing FCI Scores for Two Groups

	VPL vs. HO	HO vs. SUPP	VPL vs. SUPP
Force Concept Inventory	$p = .14$	$\boldsymbol{p < .01}$	$\boldsymbol{p < .05}$

Table 3: Demographic Data for Treatment Groups

Attribute	American Indian or Alaskan Native	Asian or Pacific Islander	Black, not of Hispanic origin	Hispanic	White, not of Hispanic origin	Male	Female	Avg Grade Level	Avg FCI Score
VPL Group	1 (1%)	1 (1%)	4 (5%)	6 (8%)	66 (84%)	26 (33%)	52 (67%)	11.13	9.35
HO Group	0 (0%)	1 (2%)	2 (4%)	10 (18%)	42 (76%)	37 (67%)	18 (33%)	11.38	8.31
SUPP Group	0 (0%)	3 (9%)	3 (9%)	4 (11%)	25 (71%)	20 (57%)	15 (43%)	11.31	12.31

Table 5: Mean and Standard Deviation with Two Tailed t-test Results

Lab	Group	N	Pre-Test Mean	SD	Post-Test Mean	SD	p-value
Lab 1: Lenses	VPL	78	2.38	1.26	4.04	2.02	***p* <.01**
	HO	55	2.38	1.39	3.73	1.41	***p* <.01**
	SUPP	35	2.46	1.62	3.77	1.46	***p* <.01**
Lab 2: Refraction	VPL	77	3.34	1.69	4.99	2.20	***p* <.01**
	HO	53	2.26	1.47	4.15	1.78	***p* <.01**
	SUPP	34	2.94	1.80	5.45	1.42	***p* <.01**
Lab 3: Ohm's Law	VPL	78	3.74	1.49	5.36	1.74	***p* <.01**
	HO	55	4.07	1.83	4.96	2.08	***p* <.01**
	SUPP	35	4.26	1.93	6.09	2.01	***p* <.01**
Lab 4: Resistors	VPL	68	4.47	1.80	5.65	1.56	***p* <.01**
	HO	55	4.69	2.07	5.40	1.92	***p* <.01**
	SUPP	35	5.26	1.87	6.71	1.60	***p* <.01**
Lab 5: Specific Heat of Metal	VPL	22	5.55	2.48	7.18	1.62	***p* <.01**
	HO	11	4.45	2.66	6.82	1.33	***p* <.01**
	SUPP	20	4.30	2.92	6.62	1.56	***p* <.01**

Table 6: ANOVA comparing VPL, HO and SUPP Post-Test Scores

Lab		Sum of Squares	df	Mean Square	F	Sig.
Lab 1: Lenses	Between Groups	3.654	2	1.827	0.613	p = .54
	Within Groups	491.965	165	2.981		
	Total	495.619	167			
Lab 2: Refraction	Between Groups	39.143	2	19.572	5.254	***p* <.01**
	Within Groups	595.961	160	3.725		
	Total	635.104	162			
Lab 3: Ohm's Law	Between Groups	27.000	2	13.500	3.696	***p* <.05**
	Within Groups	602.619	165	3.652		
	Total	629.619	167			
Lab 4: Resistors	Between Groups	38.577	2	19.289	6.819	***p* <.01**
	Within Groups	438.461	155	2.829		
	Total	477.038	157			
Lab 5: Specific Heat of Metal	Between Groups	3.387	2	1.694	0.711	p = .50
	Within Groups	119.077	50	2.382		
	Total	122.464	52			

group and the other two groups. This may be explained by the low number of students in this group and the fact that this group only contained students from two schools. There was no significance noted between the VPL and HO groups.

Paired t-test Pre/Post Tests

A two-tailed, paired t-test was used to compare each pre-test to each post-test. The individual t-tests show that there was a significant difference on the pre- and post-test grades for all five labs for all three groups. Table 5 below summarizes the results. The improvement from pre to post-test is an expected result and clearly shows that none of the treatments caused confusion or "harm" to the learning of the students. It can be noted that the VPL and SUPP groups' averages were higher for the first four labs for which we had the larger sample of students. This result will be investigated in later sections. The significant p-values are highlighted in bold italics. They indicate that learning occurred with all treatments.

Analysis of Variance for Three Groups

A one-way Analysis of Variance (ANOVA) was used to compare the post-test scores of all three groups to determine whether there was any significant difference among the three groups. Table 6 summarizes the data and shows that there was a significant difference among the groups for Labs 2, 3, and 4. There was no significant difference found among the three groups for Lab 1 or Lab 5. The significant p-values are highlighted in bold italics.

From Table 6 it can be observed that for Labs 2, 3 and 4 the supplemental post-test scores were higher. There are two possible explanations for this. The first explanation is that the SUPP group's previous knowledge was greater, as observed in Table 3 and in Table 4. The supplemental groups scored significantly higher on the Pre-FCI than either of the other two groups; these students started with more prior knowledge so they did better on the assessments than the other two groups. A second explanation is that these students had the advantage of experiencing the lab twice; they did the hands-on lab and then they also did the virtual lab, and this double treatment gave them a better understanding of the lab material resulting in higher scores on the post-lab test. Further research studies are being designed to determine if any type of double treatment (doing hands-on twice or doing virtual labs twice) is enough to produce this kind of result, or if the double treatment with the combination of hands-on and virtual produces better results.

Two Sample t-test for Pairs of Groups for Post-test Scores

A two-tailed, two-sample t-test with samples considered to have different variance was used to compare post-test scores between two different groups. The individual t-tests show that there is a significant difference between post-test grades for some of the groups. Table 7 below shows the results of individual t-tests with the significant differences highlighted in bold italics. As stated before, the explanation of the significant differences between the SUPP group and the HO could be somewhat explained by the different make-up of the group and the fact that this group only had students from two schools. In the case of the significant difference on Lab 2 between the VPL and the HO, there is evidence to say that the VPL group did significantly better than the HO group.

Table 7: t-test Comparing the Post Test Means for Two Groups

Lab	VPL vs. HO	SUPP vs. HO	VPL vs. SUPP
Lab 1: Lenses	$p = .30$	$p = .89$	$p = .43$
Lab 2: Refraction	***p < .05***	***p < .01***	$p = .19$
Lab 3: Ohm's Law	$p = .25$	***p < .05***	$p = .07$
Lab 4: Resistors	$p = .34$	***p < .01***	***p < .01***
Lab 5: Specific Heat of Metal	$p = .50$	$p = .71$	$p = .26$

Regression Analysis

The post-test scores for Labs 1, 2, 3, and 4 were combined to give an overall Total Lab Score. (The grade for Lab 5 was not used because only one school completed this lab.) This combined quiz score reflected what knowledge the students acquired after completing the labs. The Force Concept Inventory taken at the beginning of the year was used as the independent variable, or indicator of the knowledge that the students had before completing any of the labs. A regression analysis was completed to determine if there was any relationship between the FCI and the Total Lab Score. The results are in Figure 2.

From the trendline it can be seen that there is a positive relationship between the FCI and the combined Total Lab Score, indicating that students who did better on the FCI (baseline indicator of knowledge) seemed to do a little better on the Total Lab Score. From the value of R^2 (0.1526), it can be seen that there is not a strong correlation between the FCI and the Total Lab Score. Only about 15% of the variation in the Total Lab Score can be explained by the previous knowledge demonstrated on the FCI, therefore 85% is either random or has another explanation. The regression equation can be seen in the top right corner of Figure 2: Regression Line FCI vs. Total Lab Score2.

Analysis of Covariance

The Analysis of Covariance was used to find out how the Total Lab Score varied based on the treatment (VPL, HO, SUPP) and FCI baseline score. Only 18% of the variability of the Total Lab Score is explained by the treatment or the baseline FCI score. The remainder of the variability is due to some effects that have not been or could not be measured during this experiment or ones that have not been considered in the model at this time. A significant amount of information is not explained by the ANCOVA model used. Further analyses would be necessary. Figure 3: Regression of Lab Total Score by Treatment Group3 below shows the regression analysis of the Total Lab Score by FCI for each of the three

Figure 2: Regression Line FCI vs. Total Lab Score

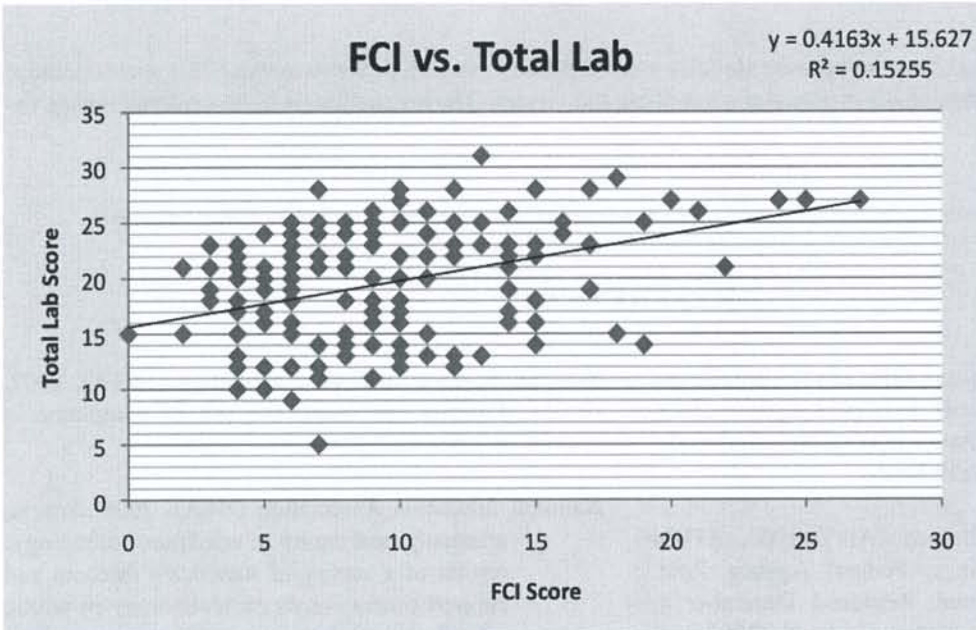
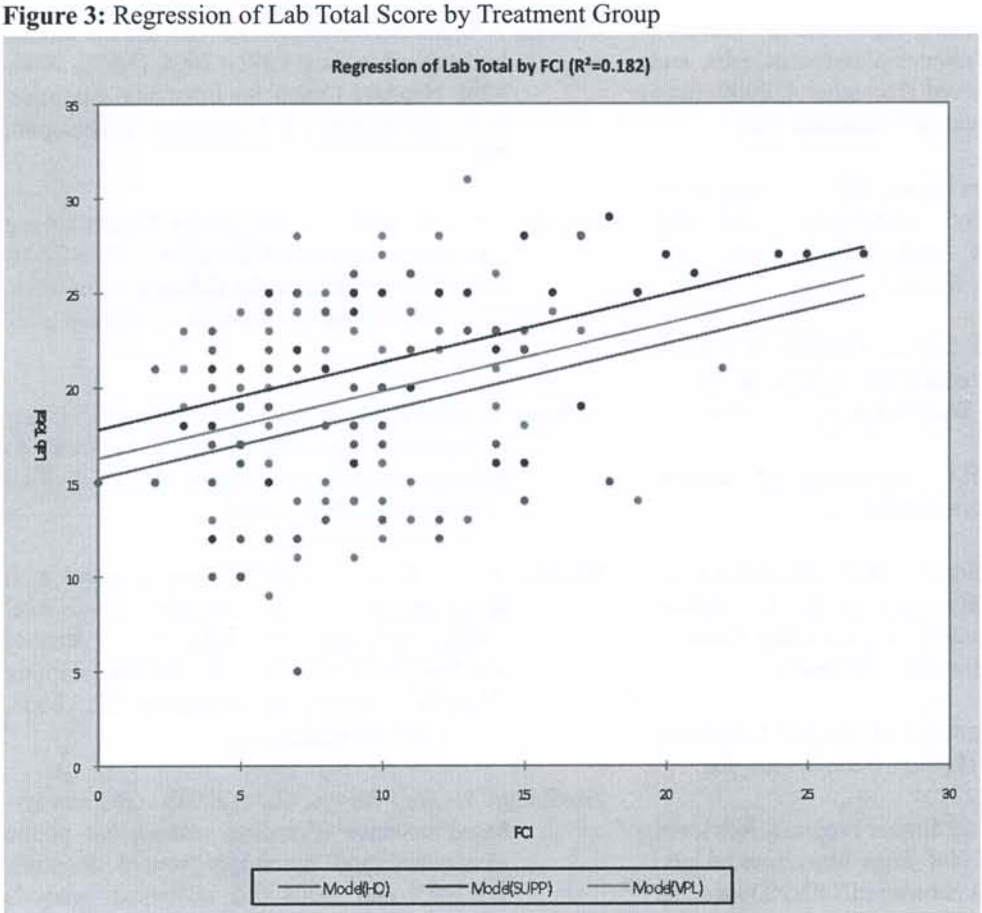


Figure 3: Regression of Lab Total Score by Treatment Group



treatment groups.

Round 2 – Conclusions and Implications

The analyses of the data did not show any conclusive evidence that one treatment was more effective or less effective than any of the other treatments, making it reasonable to conclude that the *Virtual Physics Lab* used in the high-school setting produces similar learning outcomes as the hands-on traditional labs. The developers viewed this conclusion as very positive. The implications of these findings are important as traditional schools struggle to equip enough physics labs with hands-on equipment to serve the needs of the rising number of students and as the number of online courses grows. If at least some of the hands-on labs can be replaced with equally effective virtual labs, students should expect to learn physics concepts as well as they would if funding and access problems allowed them to conduct experiments hands-on.

Of interest was the finding that for some of the labs that were tested, the supplemental group showed increased learning. Was that increased learning because students were exposed to the lab experience twice? Or was it because they had access to two forms of lab experience – hands-on and virtual? Those questions will be explored during the next round of testing.

REFERENCES

American Institute of Physics (AIP). Physics enrollment in U.S high schools. Retrieved April 7, 2013, from http://www.aip.org/statistics/trends/highlite/hs2/figure1b.htm

American Institute of Physics (AIP). 2007. STEM Education Hearings: Federal Agency Role, Laboratory Science. Retrieved December 4, 2009, from http://www.aip.org/fyi/2007/064.html

Arkansas Science Teachers Association (ASTA). 2008. Adequate science classrooms, labs, and equipment. Retrieved December 4, 2009, from http://users.aristotle.net/~asta/labs.htm

Business-Higher Education Forum. 2005. A commitment to America's future: responding to the crisis in mathematics and science education. Washington, DC: BHEF.

Council of Chief State School Officers (CCSSO). 2005. Key state education policies on PK-12 education: 2004. Washington, DC: Author.

Gronlund, N. E. (2008) *Assessment of Student Achievement*, Allyn and Bacon.

Institute of Education Sciences (IES) National Center for Education Statistics. 2008. 1.5 million homeschooled students in the United States in 2007. 2008. Washington, DC: Author.

Liu, Xiufeng (2010) Essentials of Science Classroom Assessment", 2010, Sage Publications, Inc.

National Assessment of Educational Progress. Retrieved September 23, 2008 from http://nces.ed.gov/nationsreportcard/pubs/main2005/2006466.asp

National Science Teachers Association (NSTA). 2007. Position statement: the integral role of laboratory investigations in science instruction.

National Science Teachers Association (NSTA). 2007. Position statement: the use of computers in science education.

National Education Association (NEA). 2008. Access, adequacy, and equity in education technology: results of a survey of America's teachers and support professionals on technology in public schools and classrooms. Washington, DC: Tuck, Kathy

Owings, J.A. and Schneider, M. (2008). Trends among high school seniors: 1972 – 2004. (NCES 2008-320). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Owings, J.A. and Schneider, M. (2008). Trends among high school seniors: 1972 – 2004. (NCES 2008-320). National Center for Education Statistics, U.S. Department of Education. Washington, DC.

Tunison, S. (2003). "A place among the fossils": Using metaphors from imaginative literature to manage change in our schools. *McGill Journal of Education, 38*(1), 79-91.

Watson, J. and Ryan, J. (2007). Keeping pace with K–12 online learning: A review of state-level policy and practice. [Electronic version] North American Council for Online Learning (NACOL). Retrieved September 12, 2008, from http://www.nacol.org/

Zandberg, I, and Lewis, L. (2008). Technology-based distance education courses for public elementary and secondary school students: 2002–03 and 2004–05, statistical analysis report. (NCES 2008-008). National Center for Education Statistics, U.S. Department of Education. Washington, DC.