Comparison of Simulation and Hands-on Labs in Helping High School Students Learn Physics Concepts

Matthew Rytting[1], Geoffrey A. Wright[1,*], Steven L. Shumway[1] & Jamie Jensen[2]

[1]Technology and Engineering Studies, Brigham Young University, 230 D Snell Building Provo, UT 84602, USA

[2]College of Life Sciences, Brigham Young University, 4059 LS Building Provo, UT 84602, USA

[*]Corresponding author: Technology and Engineering Studies, Brigham Young University, 230 D Snell Building Provo, UT 84602, USA. E-mail: geoffwright@byu.edu

Received: December 7, 2018 Accepted: January 9, 2019 Published: March 7, 2019

doi:10.5296/ije.v11i1.14017 URL: https://doi.org/10.5296/ije.v11i1.14017

Abstract

The purpose of the research was to determine whether PhET simulation labs or hands-on labs were more effective in helping students learn physics concepts. This measure was done by comparing quiz scores using recall, calculation, and transfer questions. Additionally, student perceptions of learning from both hands-on and simulation lab experiences were measured. Six labs were conducted with high school physics students on the topics of momentum, energy, circuits, angular momentum, pendulums, and friction. It was found that PhET simulation labs were as effective at creating student understanding, and sometimes more effective, as measured by quizzes given after the labs. Additionally, the survey data revealed that students were more engaged by hands-on lab experiences, and viewed the hands-on labs to be more effective than the simulation labs.

Keywords: Physics, Simulation lab, Hands-on lab

1. Introduction

The use of simulations has become more mainstream in both secondary and post-secondary education and industry. Although many of these simulations are true to the conditions of their physical counterparts and have the flexibility to account for many variables and factors, what can be done is ultimately limited and controlled by the creator of the simulation. The purpose of this study is to investigate whether students learn as well using simulations as they do in traditional lab experiences. This research had four sections of high school physics classes perform either simulation or hands-on versions of a lab to compare how the students learned and their attitudes toward the labs.

2. Background

Engineering jobs are a vital part of the United States' economy, and one of the most consistently under-filled professional occupations. One solution to the employment gap has been to improve and promote STEM education, specifically in K-12 settings. Various methods have and are being employed in K-12 classrooms, such as new and revised curriculum, improved STEM assessments, increased use of technology and increased funding (i.e. Engineering is Elementary, Race to the Top, Project Lead the Way, etc). One avenue that is being utilized to meet this need is the use of online courses, tutorials, and simulations (Gibson, 2007). The eLearning market increased at an average of 9.2% annually from 2010-2015 and 33% of U.S. companies use simulation applications for their employees (Pappas 2015). The use of simulations is especially applicable to STEM because many STEM courses involve learning concepts through performing labs and designing and creating products (i.e. CAD). With the rising use of simulations, the concern is whether students learn as effectively from these as from using hands-on procedures (Rutten, 2012).

The use of laboratory experiences to aid student learning has been a part of American education since the late 1800's. The prevailing viewpoint at that time was that since scientists engaged in meticulous observation of phenomena to identify and accumulate data in support of new theories, students should engage in similar activities to better prepare themselves for higher education and to understand the scientific method (National Research Council, 2005). While the exact emphasis that should be placed on laboratory experiences has shifted overtime, the National Research Council and National Science Foundation have concluded that lab activities provide important benefits to students. Specifically, they feel that "*laboratory experiences provide opportunities for students to interact directly with the material world (or with data drawn from the material world), using the tools, data collection techniques, models, and theories of science." (National Research Council, 2005)*

The College Board, the publishers of Advance Placement (AP) tests and courses, has recently – in 2012 – revamped many of its courses in an effort to better prepare students for college and careers. As part of this curriculum change they have emphasized the need for student lab work, favoring the use of hands-on activities over simulation labs citing the study done by the National Research Council in 2005 (AP Physics Teachers Manual and NSF 1996, NSF 2012, AAPT Committee on Physics in High Schools, 1992). They felt that the opportunity for students to "interact directly with the material world" was important (NRC, 2005). In fact, they now require "25 percent of the instructional time will be spent in laboratory work, with an emphasis on inquiry-based investigations" believing that this "provide[s] students with opportunities to demonstrate the foundational physics principles".

Some physics principles (i.e. universal gravity and small charge interactions) are not easily observed or reproduced in the classroom. Lab simulations have been developed to give students experiences with these phenomena. The University of Colorado developed PhET, a collection of interactive math and science simulations. While they are not the only such collection, they do offer one of the largest number of such simulations that are used by educators nationwide (PhET, 2002).

Further, proponents of simulations point out that many fields such as medicine, military, aerospace, and law enforcement use simulations to teach "real world" skills that are normally performed outside the simulation during regular work activity (Lateef, 2010). Case studies have shown that simulations can teach concepts as well or better than in-person observations (Hensberry, 2015). While several studies have been conducted analyzing the effectiveness or other characteristics of simulation labs, few have compared them directly to hands-on labs. This research project investigated these questions of which method do students prefer and which method is more effective in the context of a high school physics class.

3. Study

The purpose of this study was to investigate whether computer simulations or hands-on labs lead to better understanding of physics concepts and better perception of learning for high school students (10^{th} – 12th grades). This study investigated the purpose by asking and addressing two questions: 1. Do simulations lead to better understanding for students in high school physics classes? 2. Do simulations lead to better perceptions of learning for high school physics students?

The research methodology of this study was of a quasi-experimental design. Four classes were tracked over the course of a year. Two classes used simulations to explore physics concepts while two other classes used hands-on activities. Student understanding was measured through quizzes completed soon after the experiences. Students also took surveys on perceived learning after the labs to see how they respond to the two methods of instruction. It is hoped that data from this study will give educators a better understanding of the benefits and limitations of online and virtual lab experiences and their place in education.

Two data collection techniques were used to answer these questions. The first method involved measuring student learning perceptions by conducting brief surveys after each lab. Perkins, et al (2005) suggests that student retention is often based on student interest and participation. The second method involved measuring student achievement on concepts after participating in the intervention through quizzes given the next school day after the lab.

The sample groups consisted of four classes of Introduction to Physics at Pleasant Grove High School. All the classes were taught the same day. Two of these classes were selected to receive the computer simulations while the other two were chosen to receive the hands-on labs for the first three labs and then both switched to the other lab type for the final three labs. The computer simulations were all drawn from the PhET website developed by the University of Colorado and were largely individual learning experiences, while the hands-on labs were largely group experiences, and were created, or modified from existing labs, by the researcher.

The study was conducted at Pleasant Grove High School located in Pleasant Grove, Utah. The school has a student population of about 2000 with approximately 85% identifying as white and 10% as Hispanic. Approximately 20% qualify for free or reduced lunch. Overall, the school serves a largely suburban community (Start Class, 2016). Students were quasi-randomly assigned to the various classes. They choose when to take the class so it cannot be called truly random, but the end result of this self-selection is fairly random with regard to gender, grade level, and achievement. The students did not begin the intervention until the second unit of the year, because 1.) the scores of the students on the first unit test were also analyzed to look for distinctions between class periods, and 2.) students participated in one physical lab and one computer simulation during the first unit to give them exposure to both types and limit biases caused by unfamiliarity. After looking at the results of the first test it was seen that period A1 had a mean score of 46.4, A2 had a mean score of 51.4, A3 had a mean score of 51.8, and A4 had a mean score of 48.6. Since a different academic level was most concerning for this research, the periods were grouped so that A1 and A3 were together and A2 and A4 were together. This also led to a final distribution of slightly more boys than girls in this second group and slightly more upperclassmen in the first group. Ultimately it was felt that minimizing academic disparity was more pertinent, and when these groups were compared the mean difference between the group was only 1.5 ($p = .20$) out of 70 points and the standard deviations for the two groups were 11.2 and 10.7. This shows that the groups were not statistically different from each other in regards to their prior academic performance in the class up to the start of the intervention.

The premise of this study was that one type of lab experience may be more effective in teaching physics concepts or may lead to better student perceptions of learning. To test this, students took a survey about the lab and a quiz over the concepts covered in the lab the next class period after completing the lab. This was done so students would have to retain the information without further instruction or influence from the teacher.

Throughout the year, students participated in labs in order to explore new concepts (for example, how velocity relates to position, how force affects acceleration, or what affects friction) and developed an equation that could describe the relationship studied. The students conducted each lab either in a hands-on manner using equipment in the classroom or using a computer simulation. The computer simulations used were developed by PhET, a program created by the University of Colorado that is used by many science teachers around the country. In both cases, students completed a lab that had the same objectives and studied the same phenomena. The labs were completed, and the next class period (Pleasant Grove High

School uses an A day/B day schedule) students took the same quiz covering the concepts studied in the lab. This had the goal of measuring which lab experience taught the concepts better without other instruction, clarification, or reinforcement from the teacher. This gap in time, rather than taking the quiz directly after the lab, means that students had to transfer the knowledge from working memory to long-term memory. In spite of the name, long-term memory is technically defined as something retained even past a few minutes. (Cowan, 2008) The wait between experiencing the lab and taking the quiz forced the students retain the information from the working memory stage to long-term.

The lab process itself is similar for both the lab and computer simulation interventions, but each has unique limitations and advantages. For physical labs, students are divided into groups of four to five students. This is largely due to equipment constraints, but is also affected by the number of tasks students are required to accomplish during the lab. Division of labor is essential for the number of measurements that have to be made. Students then carry out the lab by recording their observations and data and creating a graph or performing calculations to study the relationships highlighted by the lab. This form of lab encourages team work, delegation, accuracy of measurement (or at least awareness of where error is introduced) and physical manipulation. Students are also given more autonomy in designing lab procedure. This method generally takes longer due to the extra data collected, and thus there is sometimes a smaller scope to the actions taken in the lab. For example, in the energy lab the hands-on students only had time to use the equation for gravitational potential energy to look the effect of height or mass but the simulation provided the time and means to additionally look at the effect of the gravitational field.

Simulation labs, on the other hand, are more individual. Students complete these on the computer so the only resource limitation is how many computers are available. Students were encouraged to perform the lab themselves, but were allowed to compare notes and work with neighbors. This form of lab encourages experimentation and exploration as the ease of manipulation with the simulation lets students try multiple variations. There is little concern about accuracy of measurements as the computer provides the data automatically. Students instead focus on effects of changes they make. These labs generally have more instructions in order to focus the students on examining certain effects out of the many that can be seen.

Both types have interaction with the teacher. In each type the teacher will begin the lab with a brief introduction to the concepts, guide the students through the general procedure and demonstrate the equipment or simulation. Throughout the lab, the teacher is involved with helping students work through difficult situations and answering their questions. If a certain question occurs multiple times the teacher would often make an announcement to the entire class to clarify the issue. To measure whether one type encourages more interaction with the teacher than the other, two labs (angular momentum and pendulums) were observed by an outside observer and the number of teacher-student interactions was counted.

The scope of the intervention included six lab pairs and subsequent quizzes and surveys. The six labs covered the concepts of linear momentum, energy, circuits, angular momentum, pendulums, and friction. These labs were written or modified by the researcher. The simulation labs were all based on PhET simulations. The hands-on labs were labs that the researcher had used previously that were adapted to better match the focus of the simulation labs. In all cases, the researcher sought to align the hands-on and simulation labs as closely as possible in regard to tasks and objectives. The quizzes were written by the researcher. They were modelled after questions from reputable sources such as Tippers, the Force Concept Inventory, and *Mastering Physics*. *Mastering Physics* is a database of questions designed to match the book, *Physics* by James Walker. While modeled after these sources, the questions were created by the researcher. The questions were shown to another physics teacher at the high school to ensure the quality were acceptable. This was not to claim "face validity" but rather a device to ensure quality and accuracy of contemporary physics questions. Because both groups received the same quiz, any issue with how the questions were written affected both groups and inferences about the relative effectiveness can still be drawn.

The quizzes were written with the goal of having two recall questions, two calculation/problem-solving questions, and two transfer questions for a total of six questions per quiz. Recall questions were directly related to relationships or facts expressed in the lab. Calculation questions forced students to use the equations derived from or used in the lab. Transfer is considered the ability to apply learning in a new context (Perkins and Solomon, 1988) and so transfer questions encouraged students to integrate multiple resources or apply the principles to new situations. Educators in the state of Utah have recently begun thinking of their assessment questions in terms of Depth of Knowledge (DOK). As questions move from DOK 1 to DOK 4 students are forced to take more autonomy and integrate more information to create their answer (Naylor, 2015). The quiz questions correlate roughly to recall being DOK 1, calculation being DOK 2, and transfer being DOK 3. Quiz questions were either multiple-choice or had only one possible correct response from a calculation so that no rubric was needed to grade the student responses.

Before each quiz, students were given a survey about their perceptions of the lab. The survey consisted of two questions and focused on whether students thought the lab helped them understand the concepts and if the lab was engaging. The students responded using a sliding scale from 1 to 7 rating how effective and engaging they felt the labs were. These were given before the quiz so that student performance on the quiz would not bias their responses.

The data collected through the quizzes and surveys was analyzed using independent sample *t*-tests to measure the size and significance of the differences between the means. Separate tests were conducted for the quiz score and survey questions for each lab. The quiz scores were further analyzed with respect to the question type. A combined effect was also measured by combining all the quiz scores and survey responses to see if there were statistically significant general trends in terms of student performance or learning perceptions based the type of lab completed. Effect size of differences was calculated using Cohen's d with the interpretation that an effect of .2 - .49 is considered small, .5 - .79 is considered medium, and .8 and higher is considered large (Ellis, 2009). An ANOVA was performed to see whether the averages on question type were statistically different. Further, a correlation was performed between student responses on the survey and their quiz scores to examine if there was a significant interaction between these factors.

4. Findings

The first question posed for this research was, "Do simulations lead to better understanding for students in high school physics classes?" This question was addressed through the students' performance on quizzes taken the day after they participated in a lab experience. Each lab showed that simulation labs were at least statistically equal with hands-on labs in helping students understand physics concepts. Looking at the combined quiz scores for all six labs there is a statistically significant difference between scores from students who had performed simulation labs and those who had done a hands-on lab. Simulation labs did produce better quiz scores. Yet, the size of the difference, (d = .17) is very small. The raw difference in mean scores is only .23 points. For the size of the quiz, that is only about a 5% improvement in scores. While the general trend is that simulations do lead to slightly better outcomes, the practical significance is dubious. More accurately it can be said that simulation labs perform as well as their hands-on counterparts with the difference depending on the specific lab experience. It is interesting to note, however, that while for each individual lab the difference in means was not usually statistically significant, every simulation lab had better quiz scores than its hands-on version. For the purpose of extending this to a general population, the lack of significant difference is prohibitive, but the results are important and interesting to consider for the purpose of simple case study comparisons.

When examining the data by question type, further evidence of the benefits and relative parity are manifest. Quiz questions were grouped in pairs of recall questions, calculation questions, and transfer questions. When looking at the cumulative score of the total recall and transfer questions the difference between lab types essentially disappears. Almost all the difference in total quiz scores came from the calculation type questions. The reason why simulation labs may lead to better ability to use the equations that describe the relationships studied is an interesting one. There seems to be no predicted benefit in a simulation for this improvement in calculation score – in fact, perhaps we would predict the opposite, since much of the data collection and calculations are done by the computer. Perhaps it is that the cognitive load of manually collecting the data is reduced allowing students to focus on the relationships that derive the equations. The simulation labs may be providing scaffolding – providing students the opportunity to gradually increase their zone of proximal development – in a way that the hands-on labs are not (Iris Center, 2016).

The second question addressed by the research was, "Do simulations lead to better perceptions of learning for high school physics students?". In other words, do students perceive one type of lab as more effective than the other? Do they feel more engaged and enjoy one type more than the other? Students answered these two questions before taking the quiz to prevent their performance on the quiz from introducing a potential bias on their responses. On a seven-point scale, they rated how effective and engaging they felt the lab was. Here the results were far more conclusive both on a macro level and in each case study. With the exception of one lab (angular momentum), each lab showed a significant difference in the mean rating in favor of the hands-on lab in one or both of those measures. The combined data again showed this statistically significant difference. The effect size for the difference in engagement was greater than the difference in perceived effectiveness, but both were significant. Clearly then, students enjoy hands-on experiences more, and also feel they learn better from them. It was common to hear that sentiment expressed directly by the students during labs. This is interesting as several studies have shown that students had positive impressions of simulation learning experiences (Smentana, 2012). In this study, while students still did have what could be termed a positive impression of simulations – the mean rating for effectiveness and engagement were both greater than four on a seven point Likert scale – they clearly preferred hands-on activities.

Despite students' perceptions that they learned better from hands-on lab experiences, the data does not support that conclusion. At best, it can be said that students learn equally well from either type, and at worst (for the accuracy of their perspective) they actually learn better from simulation labs. When student responses on the survey were correlated to their quiz scores there was a clear lack of correlation between how they rated the lab and how they scored. In other words, those who thought the lab was effective did not necessarily score well on the quiz and those who thought the lab was ineffective scored equally well on the quiz. The same can be said for how they rated the engagement. The only significant correlation was how they rated effectiveness and engagement. In other words, students who enjoyed the lab also usually thought the lab taught or showed the principles well. This seems to imply that students are very poor judges of how well they are learning or what learning experiences are actually effective. Instead, it seems that students equate having fun or enjoying the experience with actual learning.

5. Conclusions

There are several limitations to the conclusions that can be drawn from this study. Each lab was effectively a separate quasi-experimental study. In that sense, it is difficult to combine the data from the labs and draw firm conclusions. Rather, only general trends can be noticed. Further, the exact reason for the trend in student performance and attitude is clouded by other factors. The cognitive load of data collection or the newness and inexperience with different types of equipment could lead to difficulty in focusing on the principles being studied. Simulation labs also tended to be more individual and perhaps students were thus less distracted. This may be an inherent advantage/disadvantage between the types of labs, but it does make it difficult to know the exact cause of the difference. The students may also interact less with the teacher in one type or the other. To address this an observer counted interactions between students and the teacher during two different labs (eight class periods, four for each lab) and the number of interactions was almost exactly the same every class period (25-30). Accordingly, this does not seem to be a factor, although perhaps the quality or duration of the interaction varied. Future study could focus on these interactions.

This idea of interaction also serves as a potential explanation in student attitudes. Students may have enjoyed the hands-on labs more because of the experience of manipulating things manually and using equipment. But, because of limitations of equipment and the number of tasks involved in collecting data, hands-on labs have to often be done as group activities, and the students may have simply enjoyed the increased interaction with their peers during hands-on labs. Addressing or separating this factor out would be enlightening. This could be done by controlling lab group sizes for hands-on labs and forcing students to work in small groups for simulation labs.

One of the more unusual findings occurred with the angular momentum lab. This lab was the only one where students rated the simulation lab higher in perceived effectiveness and engagement. One possible explanation for this is that this was the first lab after students switched which type of lab they were doing (the hands-on group started doing simulation, and the simulation group started doing hands-on), where the newness of the type of lab led may have led to confusion. However, during the time frame of the first three labs all students performed two simulation labs and one hands-on lab in addition to and outside of the labs being studied, so neither type was completely unfamiliar at the time of the switch. Further, while the angular momentum lab was unusual, the standard pattern reasserted itself for the final two labs indicating that there may have simply been something unique to that lab.

This does bring up the concern that the quiz scores and survey ratings reflect the efficacy of individual labs rather than truly comparing the two types. It should also be noted that in this study, it can only be stated that PhET simulations were compared to certain hands-on labs designed to mimic them. The degree to which the labs were similar in their objectives and tasks also has an effect on the validity of the comparison.

One final limitation is that the type of knowledge analyzed in this study was specific. The quizzes only measured content knowledge – knowledge of physics principles, ability to use the equations, and ability to apply the principles in new situations. Many would say, though, that this is not the only type of knowledge students should be gaining. The ability to collect data, problem solve when equipment or set-up is difficult, and instrumentation skills are also important learning outcomes to prepare students for careers. Hands-on labs would seem to have an inherent advantage here, but this was not addressed by the study and is an important consideration. Future studies could address whether some combination of these lab types could be even more effective. Having students perform both the hands-on and simulation versions of the lab could lead to greater gains in their understanding.

6. Recommendations

From the data collected in this study, it can be concluded that, in general, students enjoyed these hands-on lab experiences more than the PhET simulation counterparts. Student learning outcomes, however, are similar between lab types. Simulation labs are just as effective as hands-on labs for student learning. Some simulation labs even appear to have a slight advantage but the benefit may not outweigh other factors. In that light, it would seem that the question of which to use is up to the interpretation and needs of the instructor. Simulation labs are generally easier to run, less expensive, and faster than hands-on labs. Thus, if those are limitations an instructor must deal with, using simulation labs is a viable option that has demonstrated positive and equivalent outcomes for student learning. This also reflects well on the viability of online courses and learning experiences. However, if an instructor has sufficient time and resources, it may be beneficial to conduct hands-on labs - knowing that students will learn effectively and have a more enjoyable experience.

These considerations can be taken in light of the clear preference students showed for hands-on labs through their ratings on the survey questions. Further, their quiz scores demonstrated that student learning occurs in similar amounts from both types, but there is a slight advantage to simulation labs. A comment a student made after completing a simulation lab summarizes the findings succinctly: "It was super boring, but it showed the concept clearly." The use of simulations will likely only increase as society relies more heavily on digital communication and modeling. It appears that learning and productivity will likely suffer no ill effects from this transition, but it does seem that physical manipulation and interaction will always play an important role in students feeling connected to the learning experience.

References

AP Physics 1 and 2 Inquiry-Based Lab Investigations Teachers Manual. College Board, 2015. Retrieved from https://secure-media.collegeboard.org/digitalServices/pdf/ap/ap-physics-inquiry-based-lab-manual.pdf

Cowan, Nelson. (2008). What Are the Differences between Long-term, Short-term, and Working Memory? *Progress in Brain Research Essence of Memory*, 323-38. https://doi.org/10.1016/S0079-6123(07)00020-9

Ellis, P. D. (2009). Thresholds for interpreting effect sizes. Retrieved from http://www.polyu.edu.hk/mm/effectsizefaqs/thresholds_for_interpreting_effect_sizes2.html

Gibson, D., Aldrich, C., & Prensky, M. (2007). *Games and simulations in online learning: Research and development frameworks*. Hershey PA: Information Science Pub. https://doi.org/10.4018/978-1-59904-304-3

Hensberry, K., Moore, E., & Perkins, K. (2015). Effective Student Learning of Fractions with an Interactive Simulation. *Journal of Computers in Mathematics and Science Teaching 34*(3), 273-298.

Iris Center. "What Is Instructional Scaffolding." IRIS | Page 1: What Is Instructional Scaffolding? 2016. Retrieved from https://iris.peabody.vanderbilt.edu/module/sca/cresource/q1/p01/http://iris.peabody.vanderbilt.edu/module/sca/cresource/q1/p01/

Lateef, F. (2010). Simulation-based learning: Just like the real thing. *Journal of Emergencies, Trauma, and Shock, 3*(4), 348-352. https://doi.org/10.4103/0974-2700.70743

National Research Council. (2005). *America's Lab Report: Investigations in High School Science*. Washington, D.C.: The National Academies Press. https://doi.org/10.17226/11311

Naylor, K. (2015). Depth of Knowledge-Utah. Retrieved from http://www.schools.utah.gov/CURR/educatoreffectiveness/Meetings/2015/DOKPresentation.asp

Perkins, D. N., & Salomon, G. (1988). Teaching for transfer. *Educational Leadership, 46*(1), 22-32.

Perkins, K. K., Adams W. K., Pollock S. J., Finkelstein N. D., & Wieman C. E. (2005). Correlating Student Beliefs With Student Learning Using The Colorado Learning Attitudes about Science Survey AIP Conference Proceedings 790, 61. https://doi.org/10.1063/1.2084701

PhET: Free online physics, chemistry, biology, earth science and math simulations. (2002). Retrieved from https://phet.colorado.edu/

Rutten, N. P. G., Joolingen, W., & der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & education, 58*(1), 136-153. https://doi.org/10.1016/j.compedu.2011.07.017

Smetana, L. K., & Bell, R. L. (2012). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education, 34*(9), 1337-1370. https://doi.org/10.1080/09500693.2011.605182