Qualitative Analysis of the Effects of Sequence of Physical and Virtual Activities on Student Conceptual Understanding in Mechanics

Adrian Carmichael[1], Jacquelyn J. Chini[1], N. Sanjay Rebello[1], and Sadhana Puntambekar[2]

[1]Kansas State University

[2]University of Wisconsin

Abstract

This study investigates the effects of the sequence of virtual and physical activities on student learning of physics concepts related to pulleys. We analyzed open-ended student responses to worksheet questions and compared these to pre-, mid- and post-test data and progression of student understanding as they worked through the different sequences of activities. Overall there was no statistically significant difference between the Physical-Virtual and Virtual-Physical sequence groups, but there were significant differences on individual questions based on the concept tested. The most dramatic differences occurred on questions dealing with the concept of work. To build a more complete picture of student understanding of work based on sequence, we analyzed students' answers to the worksheet questions completed during the activities as well as the changes in students' answers from pre- to mid test and mid to post- test. The students who completed the Virtual-Physical sequence with prior beliefs that which agreed with the data from the virtual experiment did not change from pre- to mid-test. Students with prior beliefs that did not agree with the data in the virtual experiment changed from pre- to mid-test. A majority of these students did not show a change in understanding of the work concept after completing the physical activity. The physical data presented after the virtual data did not seem to promote change in their concept of work. We speculate that the ambiguity in the data collected by students in the physical experiment may be a factor inhibiting conceptual change, while the lack of ambiguity in the virtual data promoted conceptual change. Future work investigating the strength of student prior beliefs and student epistemological views would shed light on other factors inhibiting conceptual change.

Introduction

The affordances and limitations of physical laboratory experiments and computer simulation activities have increasingly been described in science education research (De Jong & Van Joolingen, 1998; Finkelstein, et al., 2005; Klahr, Triona, & Williams, 2007; Triona & Klahr, 2003; Zacharia & Constantinou, 2008; Zacharia, Olympiou, & Papaevripidou, 2008). This growing body of research has yet to reach a clear consensus on the relative effectiveness of simulations and physical activities on student learning. Finkelstein et al. (2005) looked at how students learned about circuits differently with virtual or physical manipulatives. The simulations used by the students were similar to the physical materials, except that the simulations showed electron flow within the circuit, which the physical materials could not. Finkelstein reported that students who had used the virtual manipulatives, i.e. the simulations, scored better on an exam and were able to build physical circuits more quickly than students who had used the physical manipulatives. Triona, Klahr and Williams (2007) investigated how physical and virtual manipulatives affect student learning about mouse-trap cars. Students used either physical or virtual manipulatives to design their cars. The physical and virtual treatments showed the same effectiveness in helping students design cars. Zacharia, Olympiou, & Papaevipidou (2008) looked at physical and virtual manipulatives in the context of heat and temperature. One group of students used only physical manipulatives, while another group of students used physical manipulatives followed by virtual manipulatives. Students who used the physical and virtual manipulatives performed better on a conceptual test than students who used just the physical manipulatives. The time required for manipulating each type of equipment may have led to this result. The authors concluded that the simulation could be manipulated more quickly than the physical manipulative, increasing student learning. In another study, Zacharia

& Constantinou (2008) once again used heat and temperature as a context to study physical and virtual manipulatives. In this study, they kept all factors equivalent for the physical and virtual conditions except the mode in which the experiment was performed. They found that the physical and virtual manipulatives were equally effective in helping students gain conceptual understanding.

In light of these studies, there is potential that the combination of physical and virtual manipulatives will greatly enhance student learning. There are many aspects of integrating physical and virtual activities that are worthy of investigation. The sequence of activities performed is of particular interest to us. In our study, we investigate the effects of sequence of physical and virtual activities on student learning in the context of pulleys. Our goal is to understand the affordances and limitations of each sequence of activities and to investigate the physics concepts that are most affected by sequence. We are also interested in investigating student responses to anomalous data as they progress through the physical and virtual experiments.

Theoretical Framework

Chinn and Brewer suggest a framework to understand how students deal with contradictory information encountered while learning science (Chinn & Brewer, 1993). Students approach science learning with a set of pre-instructional beliefs gathered from their previous experiences. When faced with experimental data contradictory to prior beliefs (anomalous data), students often resist changing their beliefs to fit the new data. According to Chinn and Brewer there are seven possible responses of the student to this anomalous data. They are as follows: (a) ignore the anomalous data, (b) reject the data, (c) exclude the data from the domain of theory A, (d) hold the data in abeyance, (e) reinterpret the data while retaining theory A, (f) reinterpret the data and make peripheral changes to theory A, and (g) accept the data and change theory A, possibly in favor of theory B.

Chinn and Brewer provide insight into the factors that influence each of these possible responses. First, the characteristics of the students' prior knowledge have a strong influence on their response to anomalous data. Four of these characteristics include: (a) the entrenchment of the individual's current theory, (b) the individual's ontological beliefs, (c) the individual's epistemological commitments, and (d) the individual's background knowledge. Next, the characteristics of the data collected can determine how students respond when faced with anomalous data. More specifically, the data must have a strong credibility and low ambiguity to influence theory change. Credibility can be enhanced by increasing the credibility of the source of the data, using accepted methods of data collection and analysis, replicating data and observing the experimental results directly. Further, unambiguous data is needed to influence theory change. If the student is faced with ambiguous data, they may be able to interpret their data to be compatible with several different theories and are more likely to find an interpretation of the data that fits their pre-instructional beliefs. Finally, processing strategy can influence theory change. When students process anomalous data deeply, theory change is more likely.

Method

This study took place in a conceptual physics laboratory. Conceptual physics is a non-mathematical physics course designed to introduce students to basic physics phenomenon. Students are typically non-science majors. This conceptual physics course consists of three 50-minute lectures each week accompanied by a 110-minute lab. Students performed the activities of this study as a part of their regular lab meeting. They were assigned completion credit for the all parts of the activities except the post-test, for which they received a portion of their lab grade based on correctness. Students had not previously studied pulleys in the lecture portion of the course, though they had been exposed to the underlying concepts used to describe pulleys.

The activities students completed are part of CoMPASS (Concept Mapped Project-based Activity Scaffolding System), a design-based curriculum that integrates concept maps, hypertext, and physical and virtual experiments (Puntambekar, Stylianou, & Hübscher, 2003 and Puntambekar & Stylianou, 2002). This curriculum consists of several important parts. Learning is framed by a design challenge, which asks them to design the best pulley setup to lift a pool table into a moving van. To activate prior knowledge, students are given opportunity to make individual and group predictions and brainstorm questions they would like to know more about. To gather information related to the challenge, students navigate through the CoMPASS website where they are presented with interactive concept maps accompanied by textual descriptions of concepts related to pulleys as shown in Figure 1. Students also learn about pulleys using both a physical pulley setup (physical manipulative) and an interactive computer simulation (virtual manipulative) as shown in Figure 2. The temporal order in which students completed the physical and virtual activities was varied by lab section.

The sequence of activities performed by the students is as follows. All students began by individually completing a pre- test then made individual and group predictions. Following this, they learned more about pulleys using the CoMPASS website as shown in Figure 2. After gathering information, they performed the physical or virtual activity based on which lab section they were in. There were five lab sections in total. Three lab sections were randomly assigned the Physical-Virtual sequence (PV), while the other two lab sections were assigned the Virtual-Physical (VP) sequence. Next, students took a mid-test. Finally they performed either the physical or virtual experiment and completed a post-test. This sequence is illustrated in Figure 3.

The pre-, mid- and post-tests were identical and consisted of multiple-choice conceptual questions. While students completed the physical and virtual activities, they recorded data and answered open-ended questions on a worksheet. We coded and analyzed students' responses to the worksheet questions using a phenomenograhic approach (Marton, 1986). As per this approach student responses to the worksheet questions are categorized based on the meanings expressed in these responses. The categories are not predefined by the researcher. Rather, the categories naturally emerge from the data.

We analyzed the pre-, mid- and post- test data statistically on basis of the overall scores as well as scores on particular questions. In addition we examined how student responses changed between the pre- and mid- test and how they changed between the mid- and post- test. The instructions given to each lab section and the data gathered from each section was identical. Except for the type of activity (physical or virtual), we controlled for all conditions, such as the time on task and interaction with the instructor. Students spent about 30 minutes on each activity, although students spent a few extra minutes on the activity when working with the real pulleys. This extra time was mainly due to time required to set up various pulley systems.

Results

Overall Scores on Pre-, Mid- and Post-Test

Based on the pre-test, mid-test and post-test scores, there is no statistically significant difference between the Physical-Virtual and Virtual-Physical sequence groups. The students who completed either sequence showed improvement between the pre-test and the mid-test, but only the students who completed the Physical-Virtual sequence showed a statistically significant improvement between the mid-test and the post-test as illustrated in Table 1.

Test Questions Categorized by Concept Tested

The questions from the pre-, mid- and post-test were categorized based on the physics concept being tested. These concepts included effort force, distance pulled, work, potential energy and mechanical advantage. The trends in the test scores were different depending on sequence and concept tested. A t-test (assuming unequal variances) was performed to find statistical differences between treatments on each of the assessments. Paired-sample t-tests were performed to determine differences between assessments within a particular treatment.

The questions dealing with effort force showed the same response pattern regardless of sequence, while those dealing with work showed significant differences in student performance based on sequence. These results are shown in Figure 4. In the questions which dealt with effort force, (Questions 1, 2b, 3 and 6a), there was no statistical difference between the pre- test scores of the PV and VP sequence [($t = 0.78$, $p=0.43$)], though there was a significant difference between the mid- and post- test scores [$t_{mid}(121) = 3.03$, $p_{mid}<0.01$; $t_{post}(113) = 3.07$, $p_{mid}<0.01$]. This indicates that students in the Physical-Virtual sequence were better able to answer test questions about effort force concept than students in the Virtual-Physical sequence. We then compared the pre-, mid- and post- test scores within each treatment. The Physical-Virtual treatment showed an increase in scores between the pre- and mid-test [$t(70)= 12.44$, $p<0.001$] but no significant difference between the mid- and post-test [$t(70)=1.18$, $p=0.24$]. The Virtual-Physical treatment showed the same trends [$t_{tpre-mid}(60)=10.69$, $p<0.001$; $t_{mid-post}(60)=-0.76$, $p=0.45$]. This indicates that students' scores on questions assessing the effort force concept improved after the first activity performed, regardless of whether this activity utilized physical or virtual manipulatives, but did not improve after the second activity.

In the questions which dealt with the work concept, (Questions 6b, 7, 8, 9, 13), there was no statistically significant difference between the pre- test and post- test scores of the PV and VP sequences [$t_{pre}(130)=0.54$, $p_{pre}=0.59$; $t_{post}(129)=0.79$, $p_{post}=0.43$]. These results are shown in Figure 4. There was a significant difference on the mid- test scores with the VP group doing significantly better than the PV group [$t(60)=5.01$, $p<0.001$]. The students in the PV sequence showed no significant difference between the pre- and mid-test [$t(70)=1.19$, $p=0.24$] but there was a significant increase in scores between the mid- and post-test [$t(70)=-6.79$, $p<0.001$]. The students in the VP sequence showed a significant increase in scores between the pre- and mid-test [$t(60)=5.62$, $p<0.001$] but there was no significant difference between the mid- and post-test scores [$t(60)=0.15$, $p=0.88$]. Thus, the test scores on questions assessing the work concept improved when students used the virtual manipulatives.

Thus, two distinct patterns emerged from the quantitative data that guide further analysis. When asked about more concrete quantities such as effort force, students showed significant improvement on test questions directly after the experiment they performed first, regardless of whether this experiment utilized physical or virtual manipulatives. Sequence did not affect the trend in test answers on questions assessing the force concept. On the other hand, when asked about abstract quantities such as work, students showed significant improvement on test questions after they performed the virtual experiment. This difference in test answers about work was based on the sequence in which the activities were performed. To understand more about this result, we investigated the students' responses to the worksheet questions dealing with the concept of work. Further, we analyzed the changes in students' pre-, mid- and post- test responses to questions assessing the work concept after having completed each type of activity.

Worksheet Responses to Questions about Work

While completing the physical or virtual experiments students recorded data and answered open-ended questions on a worksheet. To better understand how students interpreted their data while performing the physical and virtual experiments, we look at the worksheet responses to Questions 4 and 5 (WSQ4 and WSQ5) which both ask students about the work concept as shown in Figures 5 and 6.

WSQ4 read, "Based on your data, when you changed the pulley setup, how did it affect the work required to lift the object? Why do you think that is?" On WSQ4, students in both the Physical –Virtual (PV) and Virtual-Physical (VP) sequence interpreted the data from the virtual experiment as showing the work being the same for different pulley setups. In contrast, students in different sequences did not interpret the data from the physical experiment in the same way. Students in the PV sequence stated that the work changed when they changed the pulley setups while students in the VP sequence were split, with nearly half of the students stating that the work stayed the same across pulley setups and the other half stating that the work changed across pulley setups.

A similar trend was seen on WSQ5. WSQ5 read "Based on your data, how does work compare to potential energy for a given pulley system? Why do you think that is?" Students in both the PV and VP sequences interpreted the data from the virtual experiment as showing that the work was equal to the potential energy for a given pulley setup. In contrast, students in the PV and VP sequences interpreted the data from the physical experiment differently. The students in the PV sequence did not come to a consensus about how the work was related to the potential energy while students in the VP sequence were more likely to say the work was equal to the potential energy.

Changes To Test Question Responses Based on Sequence

The differences between the worksheet responses of students completing the PV and VP sequence motivated the analysis of how pre-, mid- and post- test answers change over time as students progress through each sequence. The worksheet responses themselves give us insight into how the students interpret their data during the activity. The changes in student responses between pre- and mid- or mid- and post- test may provide insight into how students' interaction with a specific activity may have affected students' conceptual understanding of work.

Question 9 on the pre-, mid- and post- test (TQ9) assesses the same concept as WSQ4. Both questions ask students to compare the work needed to lift a load in three different pulley systems, if friction were not a factor. Question 13 on the pre-, mid- and post- test (TQ13) assesses the same concept as WSQ5. Both questions ask students how the work done to lift a load compares to the potential energy of the load once lifted. In the following section, we will examine the way that TQ9 and TQ13 change from pre- to mid- to post- test. Later, we will combine this information with students' responses to WSQ4 and WSQ5 to sketch out a possible story discussing student learning through both sequences.

Test Question 9 (TQ9).

Figure 7 shows how student responses to TQ9 changed from pre- to mid- test. A vast majority of students on the pre-test responded (incorrectly) that work needed to lift a load a certain height was different across pulley systems. Among students who completed the physical experiment between the pre- and mid-test, 62% of students responded that work needed to lift a load was different across pulley systems on the pre-test and continued to provide this response on the mid-test. Apparently, completing the physical experiment did not seem to change most students' preconception that work needed to lift a load remains unchanged across pulley systems, but rather reinforced this preconception. Among students who completed the virtual experiment between the pre- and mid-test, 47% of students responded that work needed to lift a load was different across pulley systems on the pre- test and then, after completing the virtual experiment responded on the mid-test that work needed to lift a load was the same across different pulley systems. Doing the virtual activity seemed to correct many students' conceptions that work needed to lift a load remains unchanged across pulley systems in a frictionless situation.

Figure 8 shows the trends in student responses between mid- and post-test on TQ9. After completing the virtual experiment, 45% of students persisted with their response that the work needed to lift a load was different across pulley systems while 38% changed their answer to indicate that work needed to lift a load was the same across pulley systems. In contrast, after completing the physical activity the majority of students (58%) maintained that work needed to lift the load was the same after completing the physical experiment.

It seems that if the physical experiment is performed first, it reinforces the student belief that work should be different between pulley systems, even in the absence of friction. Then, the virtual experiment, which presents a frictionless situation demonstrating that the work is the same across pulleys, is only somewhat effective at helping students come to the correct answer.

Thus, if the virtual experiment is performed first, most students move from responding on the pre-test that work needed to lift a load is different across pulley systems to responding on the mid-test that work needed to lift a load is the same across pulley systems. Then, when they complete the physical experiment, they persist with their correct response on the post-test that work needed to lift a load is the same across pulley systems, despite the fact that the physical experimental data do not clearly show this trend.

Test Question 13 (TQ13).

Figure 9 shows how student responses to TQ13 changed from pre- to mid- test. For those who completed the physical experiment between the pre- and mid- test, there is no well defined pattern in student responses. Among these students, 36% responded that work needed to lift the load was equal to potential energy gained by the load on the pre- test and then responded that work needed to lift the load is not equal to potential energy gained by the load on the mid- test. In this same group, 30% of students responded that work to lift the load is not equal to potential energy gained on both the pre- and mid- tests. Apparently, completing the physical experiment teaches many students that work to lift a load is not equal to the potential energy gained, regardless of prior belief. Among students who completed the virtual experiment between the pre- and mid- test, a majority responded that work to lift the load was equal to potential energy gained on the pre- test and continued to provide this response on the mid- test. Doing the virtual activity seemed to reinforce many students' prior understanding that work to lift a load is equal to the potential energy gained by the load in a frictionless situation.

Figure 10 shows how student responses to TQ13 changed between the mid- and post-test. Among students who completed the virtual experiment between the mid- and post- test, there was an equal spread of responses for three of the response patterns, and no students responded for the fourth response pattern. The fourth response pattern not chosen by any student was that work was equal to potential energy on the mid- test and work was not equal to potential energy on the post- test. After completing the virtual experiment, students' understanding of the relationship between work and potential energy was extremely varied. In contrast, among students who completed the physical experiment between mid- and post-, a majority of students responded that work was equal to potential energy on the mid- test and continued to provide this response on the post- test. This result is similar to the trend seen in TQ9.

Discussion and Conclusion

By synthesizing the data from the pre-, mid- and post-tests as well as the worksheet responses we can describe the progression of conceptual change as students proceeded through the sequence of activities.

We first consider the PV sequence with WSQ4 and TQ9 (Figure 5, Figure 7 and Figure 8). On the pre- test, 72% of students indicate that work changes across different pulley systems. The responses to WSQ4 indicate that most students interpreted the data from the physical experiment to mean that work needed to lift a load changes across different pulley systems. On TQ9 most students indicate that work needed to lift a load is different across pulley systems on both the pre-test and the mid-test. It seems that the data from the physical experiment, which is subject to measurement error and friction, reinforces the students' prior ideas that work is different across pulleys. Then the students complete the virtual experiment. Prior to the virtual experiment, 86% of students indicated that work was different across pulley systems. Responses from WSQ4 tell us that students interpret the data from the virtual experiment, which does not include the effects of friction, to indicate that work is the same across pulley systems. This virtual data can be considered anomalous data for most students because it is likely to contradict their prior conception that work is different across different pulley systems. In looking at the change from mid- to post-test on TQ9, we see that about 45% of students who said work was different on the mid-test, still answered the post-test question with a response, which is consistent with the misconception that work needed to lift a load is different across different pulley systems. In addition, 38% of students said that work was different across pulley systems on the mid- test and later responded that work was the same across pulley systems on the post-test. Here we see a small increase in students who move from the wrong answer (work is different) to the correct answer (work is the same) after completing the virtual experiment, though many persist in their wrong understanding.

Next we consider the VP sequence with WSQ4 and TQ9 (Figure 5, Figure 7 and Figure 8). On the pre-test 77% of students indicate that work is different across pulley systems. Students' responses to WSQ4 appear to indicate that most students have the conception that work is the same across pulley systems after doing the virtual experiment. The results of this virtual experiment can be considered anomalous data for most students. A majority of students change their answer about work from pre- test to mid- test. On the pre-test TQ9 they indicate that work is different, then after performing the virtual experiment, on the mid-test TQ9, they indicate that work is the same. Next they perform the physical experiment. On the mid test, 65% of students indicate that work is the same across pulley systems. Students' responses to WSQ4 demonstrate that there is no clear consensus on work in pulley systems. These responses may indicate that the physical data has a high amount of ambiguity from the student perspective. Finally, we analyze the change in responses to TQ9 from mid- to post-test. The majority of students indicated in response to TQ9 that the work was the same on the mid and post- tests.

We next consider the PV sequence with WSQ5 and TQ13 as shown in Figures 6, 9 and 10). On the pre-test 54% of students indicated that work to lift the load was equal to the potential energy gained, while 44% indicated that these two quantities were not equal. Student responses to WSQ5 demonstrate that after performing the physical experiment, there was no consensus on the relationship between work done and potential energy gained by load lifted with a pulley system. The results of the physical experiment presented anomalous data to some of students and agreed with other students' prior beliefs. On TQ13 there also appeared to be no consensus on the relationship between work and potential energy as changes in student responses from the pre- test to mid test are somewhat random. It seems that the data from the physical experiment, which is subject to measurement error and friction, does not give the students clear ideas about work and potential energy. Then the students complete the virtual experiment. On the mid test 66% of students indicated that work was not equal to potential energy gained, while 32% indicated these quantities were equal. Responses from WSQ5 tell us that after the virtual experiment, in which there is no friction , a majority of students now correctly indicate that work is equal to potential energy for a given frictionless pulley system. In looking at the change from mid to post- test on TQ13, there is still no clear response pattern shown on the mid- and post- tests. It seems that the physical experiment may have confused students. Then, though the students clearly saw that work was equal to potential energy in the simulation, they continued to answer the post- test question in an inconsistent way.

Next we consider the VP sequence with WSQ5 and TQ13 (Figure 6, Figure 9 and Figure 10). On the pre- test, 54% of students indicate that work to lift the load is equal to potential energy gained and 47% indicate that work is not equal to potential energy. In the responses to WSQ5, most students indicate work is equal to potential energy after completing the virtual experiment. By comparing the pre-test scores and the worksheet questions, we conclude that the data gathered from the virtual experiment was anomalous data for about half the students. On TQ13, 48% of students indicate that work is equal to potential energy on both the pre-test and the mid-test, while 29% indicate work was not equal to potential energy on the pre- test and work was equal to potential energy on the mid test. It seems that there is a distribution of beliefs about work and potential energy before the virtual experiment, then the virtual experiment presents anomalous data to about half the students. After the virtual experiment, most students respond that work is equal to potential energy. The virtual experiment either reinforces prior conceptions that were correct or influences change in conceptions that were incorrect. Next the students perform the physical experiment. On the mid test, 77% of students respond that work is equal to potential energy. In the responses to WSQ5, most students indicate that work is equal to potential energy. On TQ13 most students indicate that work needed to lift a load is equal to the potential energy gained on both the mid-test and the post-test. Thus, even though the data from the physical experiment is subject to friction and measurement error, students persist in the belief that work is equal to potential energy.

The clearest trend in the data with regards to the work questions is seen in the VP sequence. Completion of the virtual experiment at the beginning of the sequence can either induce conceptual change if the results of the virtual experiment are not consistent with prior beliefs or can strengthen conceptual understanding if the results are in agreement with prior beliefs. The conceptual understanding achieved after the virtual experiment seems somewhat robust as evidenced by the trend that most student responses stay consistent between mid- and post-test. This aspect of our data is particularly interesting, because students have likely seen data from the physical experiment which conflicts with their conclusions drawn from the virtual experiment. This result can be explored further with a theoretical framework dealing with anomalous data and theory change suggested by Chinn and Brewer. According to Chinn and Brewer there are seven possible responses of the student to anomalous data. They are as follows: (a) ignore the anomalous data, (b) reject the data, (c) exclude the data from the domain of theory A, (d) hold the data in abeyance, (e) reinterpret the data while retaining theory A, (f) reinterpret the data and make peripheral changes to theory A, and (g) accept the data and change theory A, possibly in favor of theory B.

In the VP sequence, students had initial conceptions that were either consistent or inconsistent with the data then encountered the virtual data. The virtual data was anomalous data to some students while it agreed with other students' prior conceptions. Further, the virtual data was interpreted as having low ambiguity as compared to the physical data, which had high ambiguity. The completion of the virtual experiment served to change incorrect initial conceptions or reinforce correct initial conceptions. Following this, students encountered the physical data. The students generally viewed the physical data as ambiguous. The conceptions held by most students did not change as a result of the physical experiment. In summary, the virtual experiment influenced a change in conceptions in about half the students while the physical experiment did not.

The nature of the student responses gathered in this study does not allow us to determine exactly which of the seven possible responses suggested by Chinn and Brewer students use when they encounter the physical data, though we propose that most students respond to the anomalous data as described by options (a) through (f) suggested by Chinn and Brewer. The physical data does not seem to induce theory change as described by option (g) suggested by Chinn and Brewer. This result may stem from the characteristics of the anomalous data gathered in the physical experiment. Chinn and Brewer cite credibility and ambiguity of the data as important factors that affect theory change. They explain that unambiguous data is more likely to influence theory change, while ambiguous data allows for multiple interpretations of the same data. Students form a conception of work based on unambiguous data from the virtual experiment, then encounter ambiguous data from the physical experiment. The ambiguity allows students to interpret the physical data in a way that is consistent with their correct understanding constructed after completing the virtual experiment.

In order to more deeply understand the differences in PV or VP sequences in light of the Chinn and Brewer framework on theory change, further research is necessary. In future work, we plan to gather students' confidence ratings on pre-, mid and post- test questions to explore students' entrenchment of prior conceptions and how this entrenchment changes after completing either the physical or virtual experiments. An epistemological survey will be used to gather data on students' views on the credibility of the data from both types of experiments, as well as the students overall epistemological commitments. Further, we will ask students to explain their reasoning for each answer on the pre-, mid and post- test. With this information we hope to more fully trace the development of the students' conceptions of physics concepts through the sequence. These additional data will allow us to draw stronger conclusions on how the sequence of physical and virtual experiments affects or inhibits conceptual change in students, in particular with regard to their overall understanding of pulleys.

References

Chinn, C., & Brewer, W. (1993). The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Review of Educational Research , 63*, 1-49.

De Jong, T., & Van Joolingen, W. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research , 68*, 179-201.

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefskey, N., et al. (2005). When Learning about the Real World is Better Done Virtually: A Study of Substituting Simulations for Laboratory Equipment. *Physical Review Special Topics-Physics Education Research , 1*, 010103-1--010103-8.

Klahr, D., Triona, L. M., & Williams, C. (2007). Hands on What? The Relative Effectiveness of Physical Versus Virtual Materials in an Engineering Design Project by Middle School Children. *Journal of Research in Science Teaching , 44* (1), 183-203.

Marton, F. (1986). Phenomenography—A Research Approach to Investigating Different Understandings of Reality. *Journal of Thought , 21*, 28-49.

Puntambekar, S., & Stylianou, A. (2002). CoMPASS: Students' Use of External Representations in Science Learning. In P. Bell, R. Stevens, & T. Satwicz (Ed.), *Keeping Learning Complex: The Proceedings of the Fifth International Conference of the Learning Sciences (ICLS)* (pp. 352-358). Mahwah, NJ: Erlbaum.

Puntambekar, S., Stylianou, A., & Hübscher, R. (2003). Improving Navigation and Learning in Hypertext Environments with Navigable Concept Maps. *Human-Computer Interaction , 18*, 395-428.

Triona, L., & Klahr, D. (2003). Point and Click or Grab and Heft: Comparing the Influence of Physical and Virtual Instructional Materials on Elementary School Students' Ability to Design Experiments. *Cognition and Instruction , 21* (2), 149-173.

Zacharia, Z. C., & Constantinou, C. P. (2008). Comparing the Influence of Physical and Virtual Manipulatives in the Context of the Physics by Inquiry Curriculum: The Case of Undergraduate Students' Conceptual Understanding of Heat and Temperature. *American Journal of Physics , 76* (4), 425-430.

Zacharia, Z. C., Olympiou, G., & Papaevripidou, M. (2008). Effects of Experimenting with Physical and Virtual Manipulatives on Students' Conceptual Understanding in Heat and Temperature. *Journal of Research in Science Teaching , 45* (9), 1021-1035.

Table 1

Overall pre-, mid-, and post-test scores on the conceptual assessment. Uncertainties are the standard error of the mean.

Treatment	N	Pre-test %	Mid-test %	Post-test %	Pre to Mid t-test (2-tailed, paired sample)	Mid to Post t-test (2-tailed, paired sample)
Physical-Virtual	71	**37 ± 2**	**58 ± 2**	**66 ± 3**	t(70) = -9.79, p < 0.001	t(70) = -5.64, p< 0.001
Virtual-Physical	61	**33 ± 2**	**60 ± 3**	**61 ± 3**	t(60) = -8.75, p < 0.001	t(60) = 0.94, p = 0.35
t-test (2 tailed, unequal variances)		t(125)=1.18, p=0.24	t(113)=-0.41, p=0.68	t(121)=1.36, p=0.18		

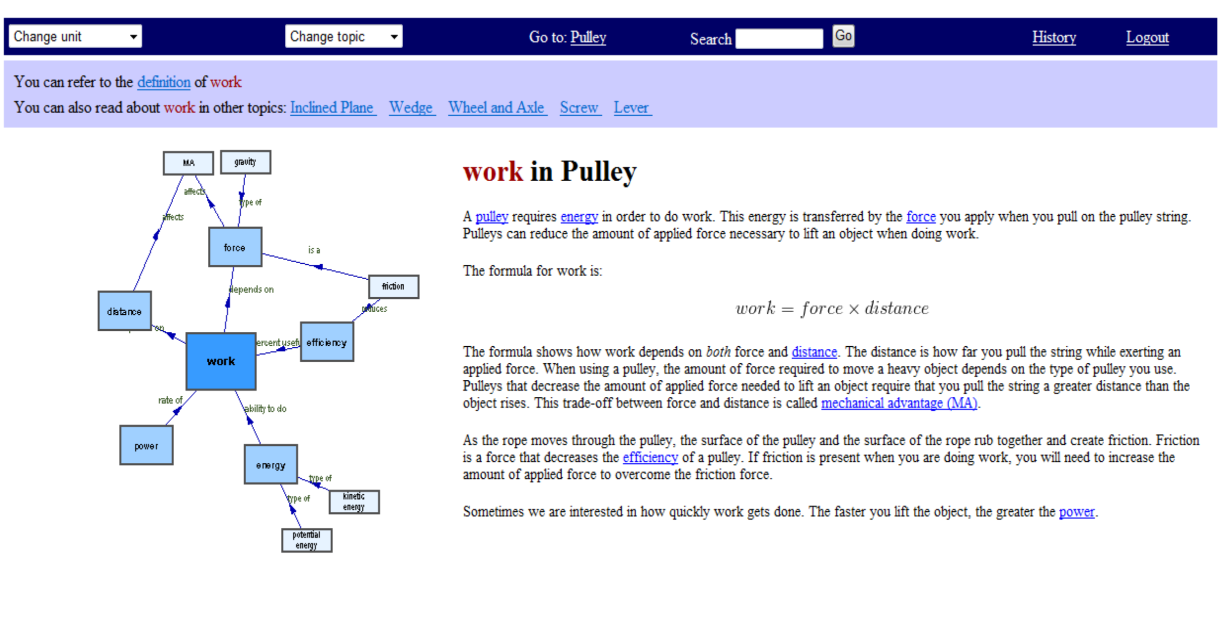


Figure 1. CoMPASS, dynamic concept maps and hypertext-based environment.

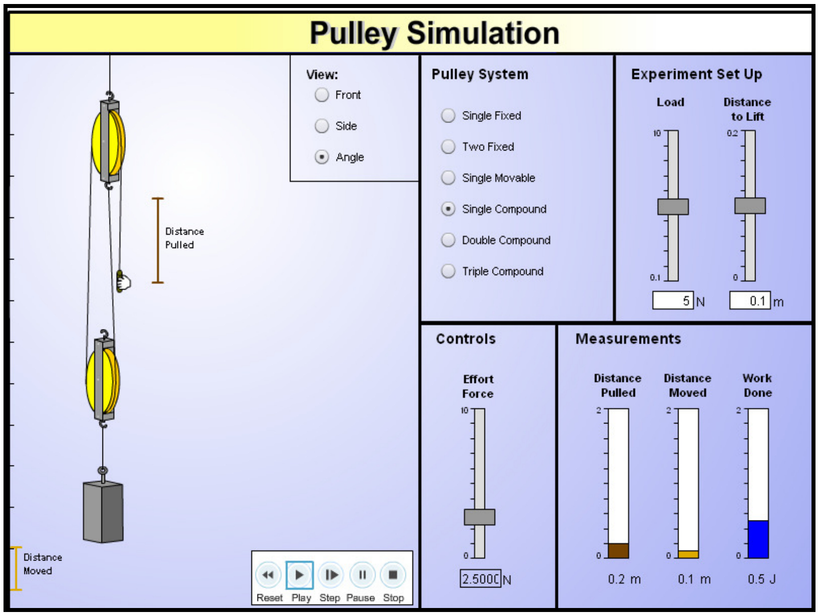


Figure 2. Screen shot of simulation.

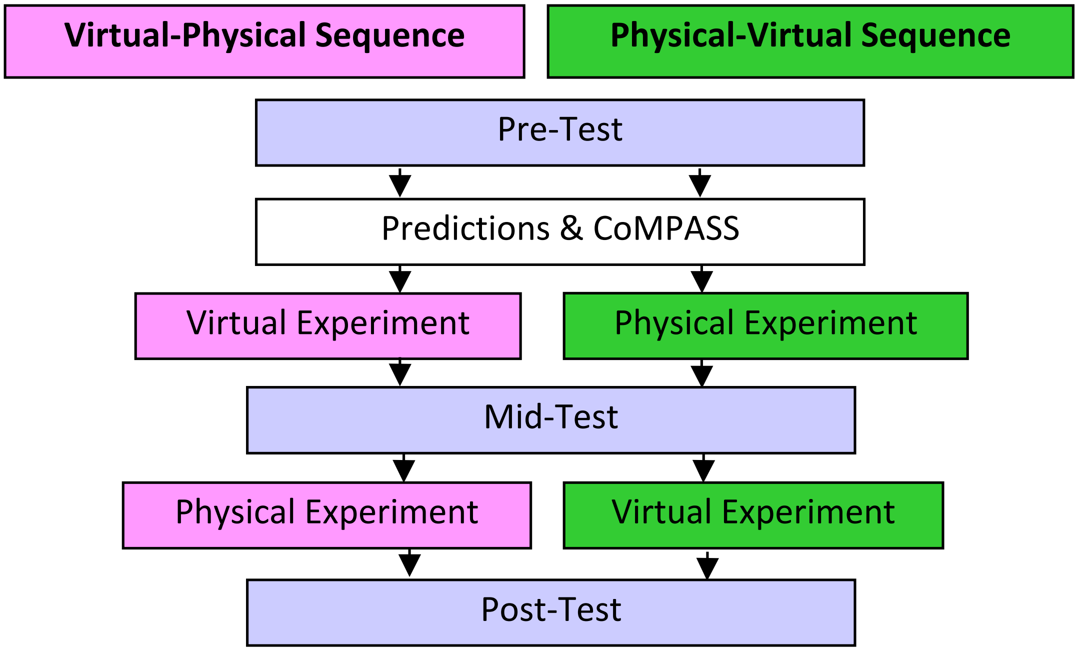


Figure 3. Sequence of activities performed by students.

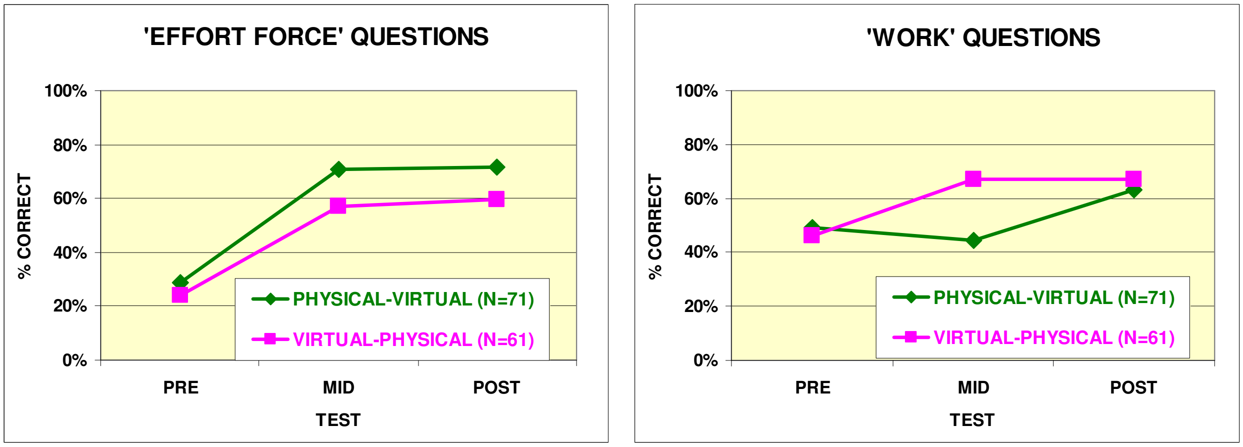
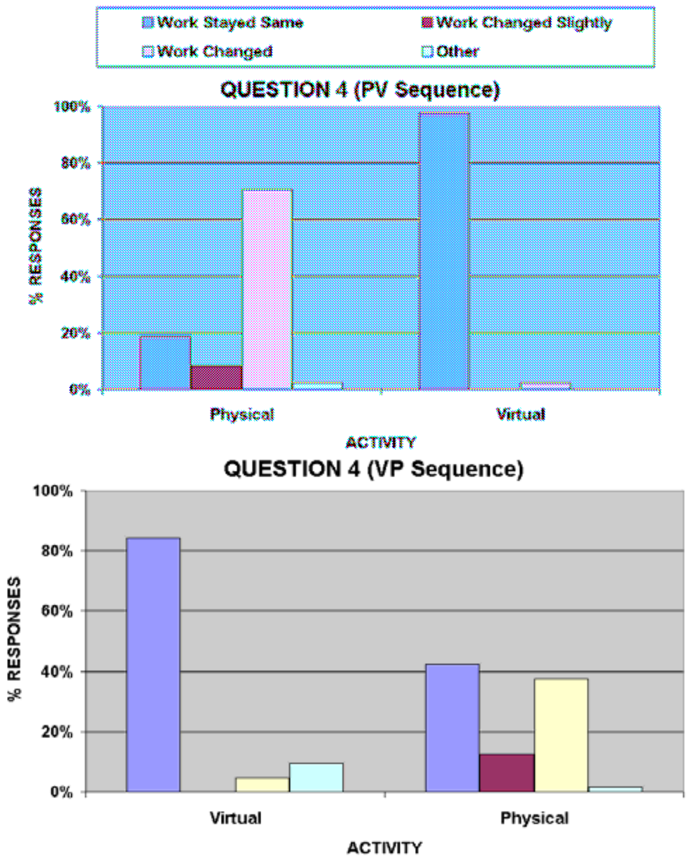
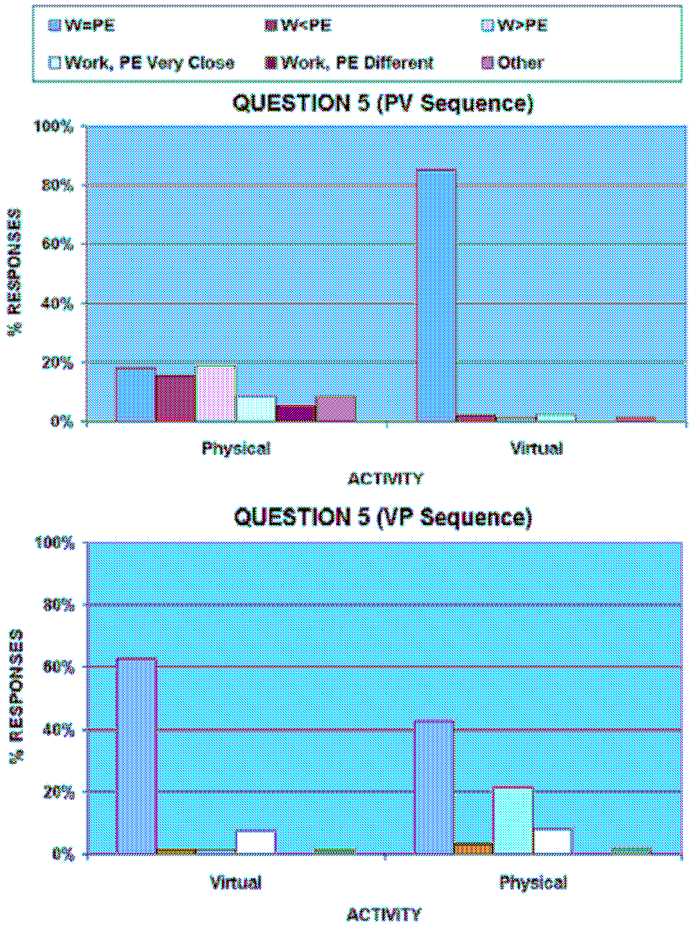


Figure 4. Average scores by category on the conceptual assessments.



Question WSQ4: “Based on your data, when you changed the pulley setup, how did it affect the work required to lift the object? Why do you think that is?”

Figure 5. Student responses to Question 4 on the activity worksheet.



Question WSQ5: "Based on your data, how does work compare to potential energy for a given pulley system? Why do you think that is?

Figure 6. Student responses to Question 5 on the activity worksheet.

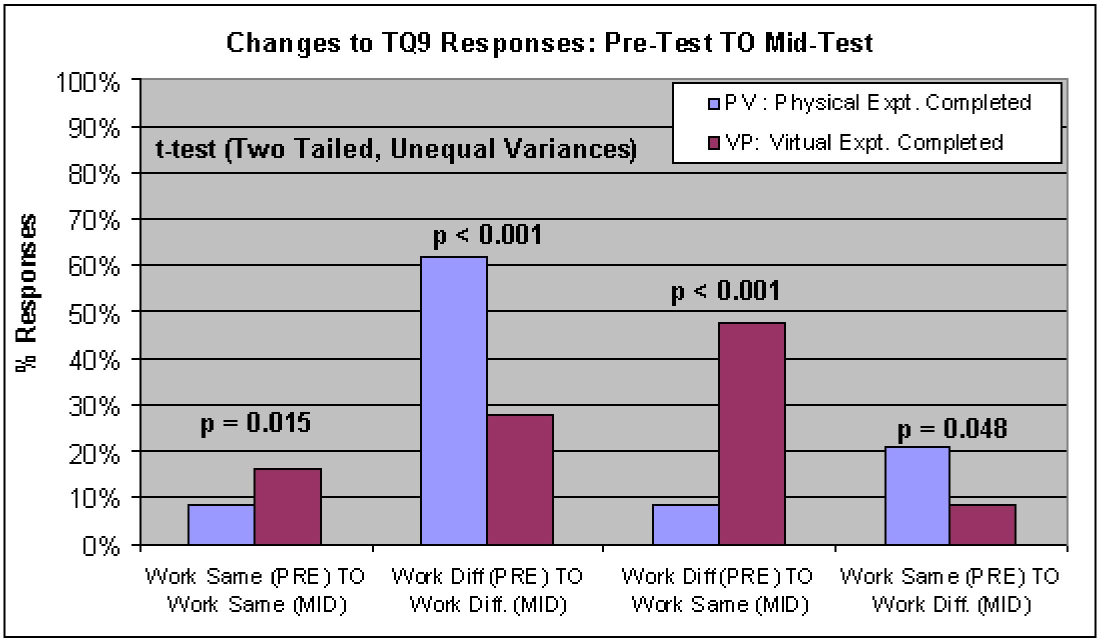


Figure 7. Changes in student answers from pre- to mid test on Question 9.

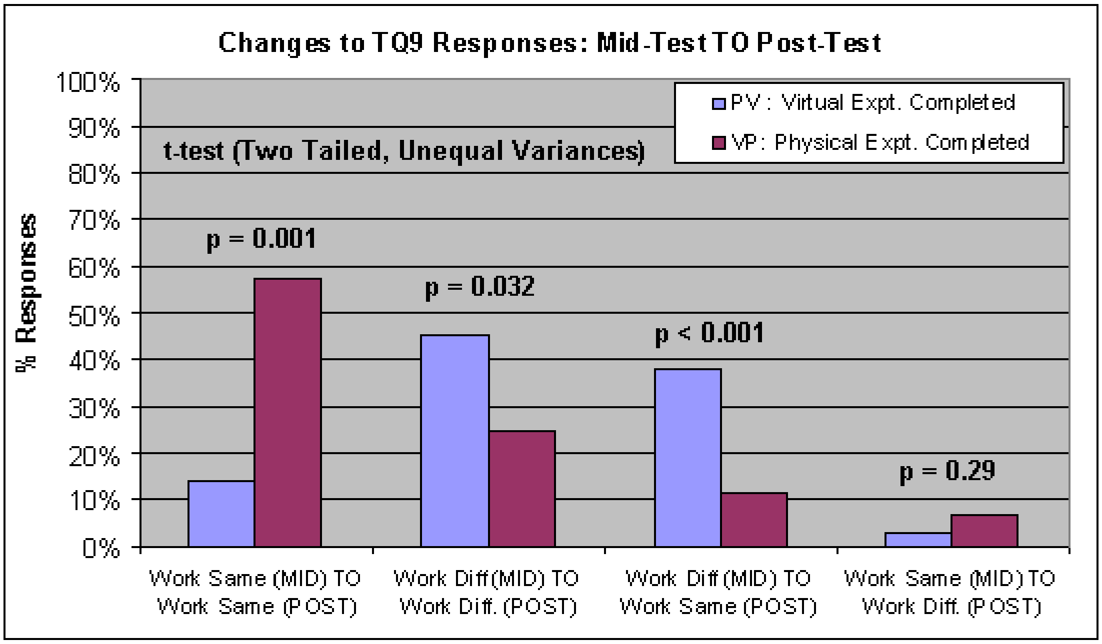


Figure 8. Changes in student answers from mid to post- test on Question 9.

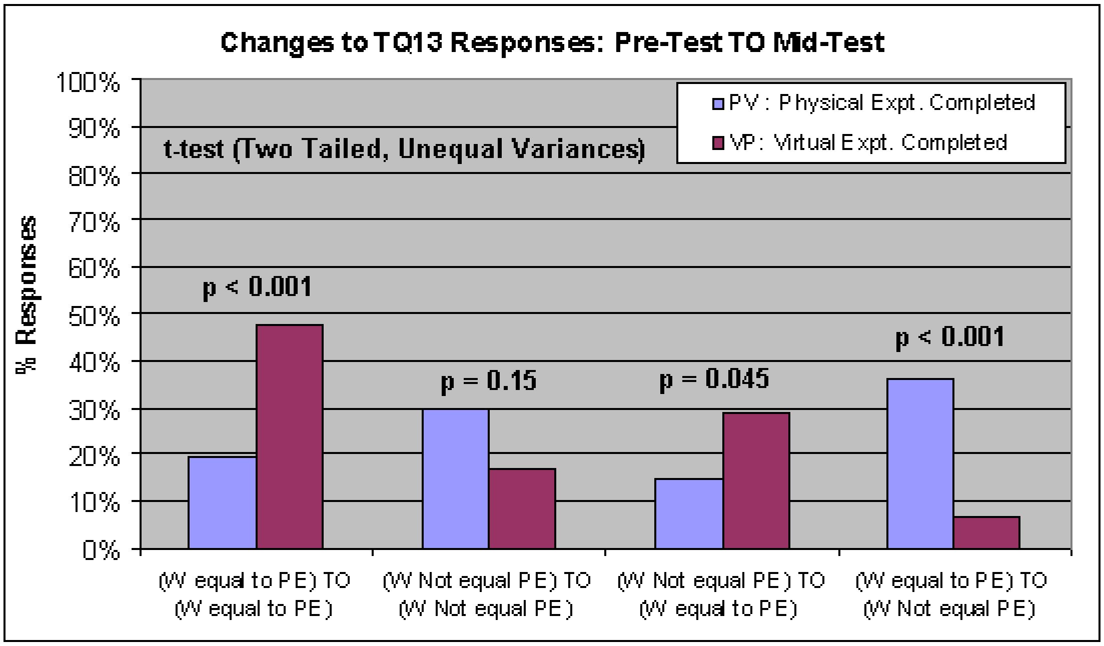


Figure 9. Changes in student answers from pre- to mid test on Question 13.

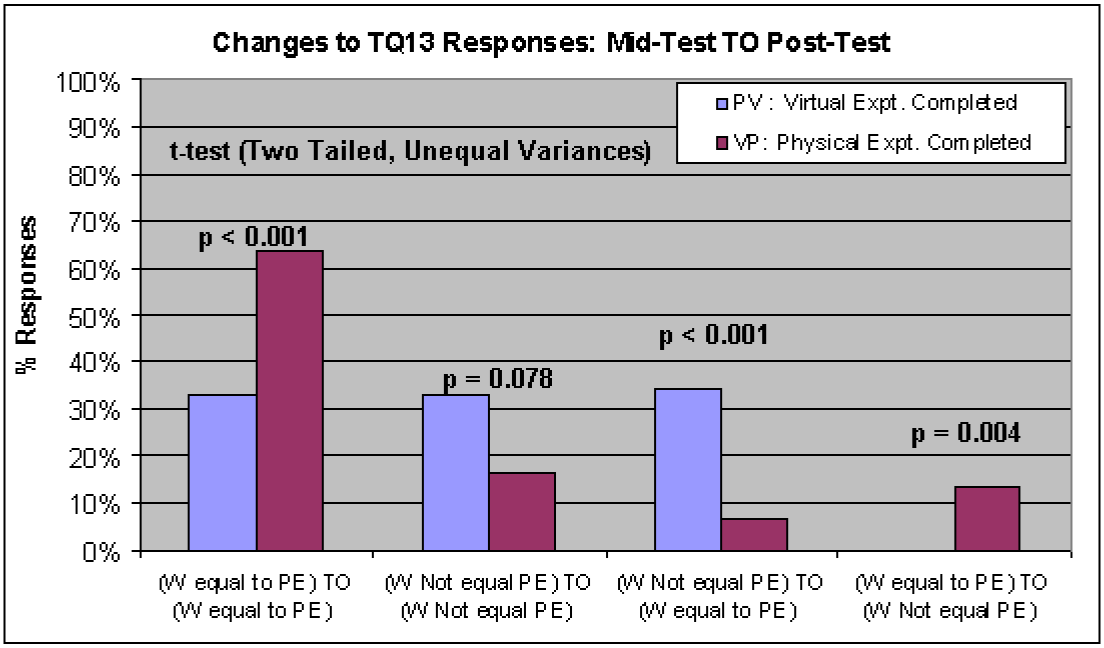


Figure 10. Changes in student answers from mid to post- test on Question 13.