# Comparing the Effects of Sequencing of Physical and Virtual Manipulatives on Student Learning and Confidence

Adrian Carmichael[1], Jacquelyn J. Chini[1], Elizabeth Gire[2], N. Sanjay Rebello[1], and Sadhana Puntambekar[3]

[1]Kansas State University

[2]University of Memphis

[3]University of Wisconsin

## Abstract

Previous research has shown that depending upon context, virtual manipulatives either outperform or perform equally well as physical manipulatives for promoting student learning. Here we focus on which sequence of using both manipulatives better supports learning, which manipulative provides for more retention of information learned and how confidence ratings by students of answers on a conceptual assessment depend on manipulative. Our research is conducted in the context of a conceptual physics course for non-science majors at a large Midwestern university. Students completed activities in two, two-hour laboratory sessions. Some students completed the physical activity before the virtual activity, while others completed the virtual activity before the physical activity. The activities were spaced one week apart. We present the results of student performance on a conceptual test and Likert-scale confidence ratings as well as implications for learning and retention of different science concepts as well as students' confidence on the assessments. We find that the virtual experiment is more beneficial to students' total learning and their learning about work and energy from pre-test to mid-test1 and mid-test2 to post, though we find no overall difference between the physical-virtual sequence and the virtual-physical sequence between pre-test and post-test. Further, we find no difference in students self-reported confidence ratings based on type of manipulative used.

## Introduction and Background

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, Zacharia & Olympiou, 2011). 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 & Olympiou (2011) 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 investigated five conditions: physical only, virtual only, physical then virtual, virtual then physical and a control condition which consisted of traditional demonstrations and lectures. They found that the first four conditions were equally effective in helping students gain conceptual understanding and all were more effective then the control condition.

In light of these studies, there is potential that the combination of physical and virtual manipulatives will greatly enhance student learning. In our study, we investigate the effects of sequence of physical and virtual activities on student learning in the context of pulleys. Our first 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.

Conceptual understanding is one of many factors that can be studied when comparing physical and virtual manipulatives. Another is students' perceptions of the usefulness of physical and virtual manipulatives in different contexts, studied by Chini (2010). In this study, future elementary school teachers were asked to discuss which manipulative they would prefer to use in each of three contexts: on a test, in a rental store and for a laboratory make-up. They were specifically asked about the use of pulleys. It was found that most students prefer using the virtual manipulatives on a test, the physical manipulatives in a rental store and both types for a laboratory make up. Further, in the laboratory makeup context, students also preferred to use virtual manipulatives over physical. In our study, we add to the investigation of students' perceptions of their learning from each type of manipulative in the context of pulleys. Our second goal is to probe students perceptions less directly by asking students to rate their confidence on test answers before and after using each manipulative and look for the same pattern observed by Chini using direct questioning.

We are also interested in determining which of the two manipulatives provides better support for retention of learning. The study of retention is among the oldest areas of formal study in the science of learning, dating back to Ebbinghaus's study of spacing effects in the 1880's (Ebbinghaus, 1885). In order to retain, it is helpful if the information is organized in a meaningful way for the subjects. Research has shown that organizing information into a schema or the use of various sorts of "organizers" to prime the students as they process new material improves retention. (Lawton & Wasanka, 1977; Moore & Readance, 1984) Our third goal of this study is to investigate how learning with virtual and physical manipulatives affects retention of information learned.

## 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. . The participants performed activities with physical and virtual manipulatives over two two-hour labs sessions spread across two weeks. This laboratory experience was the first experience for these students with the topic of pulleys. 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.

The CoMPASS (Concept Mapped Project-based Activity Scaffolding System) curriculum, (Puntambekar, Stylianou, & Hübscher, 2003) which integrates hypertext with hands-on activities and simulations, formed the pedagogical framework for the study. This curriculum consists of several important parts. Learning is framed by a design challenge, which asks the students 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 Fig. 2. The temporal order in which students completed the physical and virtual activities was varied by where students sat in the lab room.

The pulley unit in the CoMPASS curriculum was used over two consecutive weeks. Each week, half of the class used the physical manipulatives while the other half used the virtual manipulatives. The physical and virtual experiments were closely matched, such that students performed the same sorts of trials, collected the same types of data, and answered the same analysis questions. During the first two-hour laboratory session, students completed an activity with one of the manipulatives (physical or virtual) and during the second two-hour session they completed a similar activity with a different manipulative (i.e. virtual or physical). The study design is shown in Fig. 3. In the first week, students took a pre-test before beginning instruction and a mid-test (mid-test1) after completing the first activity. In the second week, students took a second mid-test (mid-test2) before beginning the second activity and a post-test after completing the second activity. The pre-test, mid-test (1 & 2) and the post-test all had conceptual multiple-choice questions that assessed students' understanding of the concepts of force and work/energy. Multiple-choice questions were typically followed with questions that asked students to provide a reason for their answer. After each question, students were also asked to rate the confidence of their answer on a 1-5 Likert Scale, where 5 designated absolute certainty and 1 designated a random guess. The total score on the test comprised the score on each multiple-choice question combined with the score on the reasoning questions. The latter were rated on a scale of 0-2 using a rubric designed by the authors shown in Fig. 4. The inter-rater reliability of over 80% was established on the rubric using two independent raters.

We analyzed the pre-, mid- and post- test data statistically on basis of the overall scores as well as scores by concept tested. In addition we examined how student responses changed between the pre- and mid- test1, mid-test1 to mid-test2, mid-test2 to and post- test and pre-test to 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 90 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

### Total Scores

Fifty-eight (58) students completed the activities in the physical-virtual (PV) sequence, and 63 students completed the activities in the virtual-physical (VP) sequence. Students could possibly score 39 points on the pulley test. Students' total scores are shown in Fig. 5. In the PV sequence, students had a mean total score of 8.48 on the pre-test. After completing the physical activity, they had a mean total score of 13.84 on mid-test1. When they came back the second week, students in the PV sequence had a mean total score of 13.10 on mid-test2. After completing the virtual activity, they had a mean total score of 17.26. In the VP sequence, students had a mean total score of 10.20 on the pre-test. After completing the virtual activity, they had a mean total score of 20.41 on mid-test1. When they came back the second week, students in the VP sequence had a mean total score of 17.73 on mid-test2. After completing the physical activity, they had a mean total score of 20.00 on the post-test. In addition to analyzing the trends in total score show in Fig. 5, we also analyzed the sub scores on the sets of questions that cover two most important concepts: force and work & energy

**Force Scores**

Students could score up to 14 points for questions about force on the pulley test. In the PV sequence, students had a mean force score of 4.43 on the pre-test, 8.53 on mid-test1, 8.14 on mid-test2 and 9.57 on the post-test. In the VP sequence, students had a mean force score of 5.17 on the pre-test, 10.17 on mid-test1, 8.6 on mid-test2 and 10.59 on the post-test. Students' mean scores on the questions pertaining to force are shown in Fig. 6.

**Work/Energy Scores**

Students could possibly score 17 points for questions related to work and energy on the pulley test. In the PV sequence, students had a work/energy score of 6.29 on the pre-test, 6.97 on mid-test1, 6.88 on mid-test2 and 8.74 on the post-test. In the VP sequence, students had a work/energy score of 7.48 on the pre-test, 11.42 on mid-test1, 10.35 on mid-test2 and 10.41 on the post-test. Students' mean scores on the questions pertaining to work and energy are shown in Fig. 7.

**Confidence Ratings**

The total confidence rating summed over all 20 multiple-choice questions is shown in Fig. 8. The trends show that students' confidence in their answers significantly ($p<.05$) improves from pre-test to mid-test1, then significantly declines from mid-test1 to mid-test2 (when students return after 1 week) and then significantly increases again from mid-test2 to post-test after they complete their second activity. There is no significant difference in trends between the PV and VP sequences.

**Statistical Analysis**

To examine the trends across the test scores more clearly, we completed a repeated measures ANOVA for the total scores as well as the sub-scores for force and work/energy. We also completed an analysis of the contrasts between various tests to examine if there was a statistically significant difference between scores on pairs of tests. Further, we calculated the effect size (R) for each of the contrasts. Table 1 and Table 2 show the results of the statistical analysis. The columns 'Repeated Measures Analysis of Variance' from Table 1 are repeated in Table 2 for clarity.

As seen in Table 1 and Table 2 main effects of Total, Force and Work/Energy (WE) indicate whether students' scores changed significantly during the sequence by averaging both sequences. Students' total scores changed significantly during the sequence, $F(2.7, 317.7)=162.7$, $p<.001$. The contrasts show students' total scores increased significantly from pre-test to mid-test1 ($F(1, 119)=239.4$, $p<.001$, $r=.817$), mid-test2 to post-test ($F(1, 119)=71.1$, $p<.001$, $r=.612$) and decreased significantly from mid-test1 to mid-test2 ($F(1,119)=19.5$, $p<.001$, $r=.38$). Students' force sub-scores changed significantly during the sequence, F(2.8, 334.0)=161.3, p<.001. The contrasts showed students' force sub-scores increased significantly from pre-test to mid-test1 (F(1, 119)= 287.1, p<.001, r=.841), mid-test2 to post-test (F(1, 119)=51.2, p<.001, r=.548), pre-test to post-test (F(1,119)=323.6,p<.001, r=.85) and decreased from mid-test1 to mid-test2 (F(1,119)=17.2, p<.001, r=.36) . Students' work/energy sub-scores changed significantly during the sequence, F(2.8, 329.0)= 34.1, p<.001. The contrasts showed students' scores increased significantly from pre-test to mid-test1 (F(1, 119)= 51.3, p<.001, r=.549), from mid-test2 to post-test (F(1, 119)=14.4, p<.001, r=.329), pre-test to post-test (F(1,119)=70.4, p<.001, r=.61) and decreased from mid-test1 to mid-test2 (F(1,119)=5.3, p=.023, r=21). These results indicate that students learned from both the first and second activities and the learning activity as a whole. They also indicate that students did not retain all the knowledge gained in week 1 of the lab as measured at the beginning of week 2.

The Tot*Seq, For*Seq and WE*Seq effects indicated whether students' total scores, force scores and work/energy scores respectively differed between the PV and VP sequences. Students' total scores changed differently between the two sequences, F(2.6, 309.2)=11.2, p<.001. Students' total score changed differently from pre-test to mid-test1 (F(1, 119)= 23.3, p<.001, r=.405), with the VP sequence outscoring the PV sequence. Students' scores also changed differently from mid-test2 to post-test (F(1, 119)=6.1, p=.015, r=.221), with the PV sequence making a steeper increase than the VP sequence. Further, total scores changed differently from mid-test 1 to mid-test 2 (F(1,119)=6.3, p=.014,r=.22) with the VP sequence making a greater decline in score than the PV. However, students' scores did not change differently from pre-test to post-test (F(1, 119)= 1.0, p=.706, r=.091), which indicates that the two sequences resulted in the same overall learning.

Students' force scores did not change differently between the two sequences, $F(2.8, 334.0)=1.86$. However, students' work/energy scores changed differently between the two sequences, $F(2.8, 329.0)= 14.2$, $p<.001$. Students' scores changed differently from pre-test to mid-test1 ($F(1, 119)= 25.8$, $p<.001$, $r=.422$), with the VP sequence outscoring the PV sequence. Students' scores changed differently from mid-test2 to post-test ($F(1, 119)=12.5$, $p=.001$, $r=.308$), with the PV sequence making a steeper increase than the VP sequence. However, students' scores did not change differently from pre- to post-test ($F(1, 119)= .579$, $p=.448$, $r=.070$), which indicated that the two sequences resulted in the same overall learning about work/energy. The work/energy sub-scores were nearly significant from mid-test1 to mid-test2 ($F(1,119)=3.8$, $p=.052$, $r=.18$) suggesting that the VP scores decreased more steeply then the PV sequence.

## Discussion and Conclusion

The total score averaged across both sequences increased significantly from pre-test to mid-test1 and from mid-test2 to post-test, suggesting that both experiments benefited students' learning. Students learning about the physics concepts that govern pulleys did not saturate after using one type of manipulative during the first laboratory session. More time on task while using a different manipulative was found to improve students conceptual understanding of pulleys. Further, the average force sub-score and the average work-energy sub-score also increased significantly from pre-test to mid-test1, and mid-test2 to post-test, again indicating that both experiments benefited student learning. There was a significant interaction between sequence and total score and sequence and work/energy sub-score. These interactions were significant from pre-test to mid-test1 and mid-test2 to post-test. Combined with the plots of average scores, these results indicate that the virtual experiment is more beneficial to students' total learning and their learning about work and energy. Previous research has looked at students overall conceptual learning with physical and virtual manipulatives. (Finkelstein, et al., 2005; Klahr, Triona, & Williams, 2007; Triona & Klahr, 2003; Zacharia & Constantinou, 2008; Zacharia, Olympiou, & Papaevripidou, 2008, Zacharia & Olympiou, 2011) In this study, we show that virtual manipulatives are more beneficial for learning particular concepts, namely work and energy. Further testing of the concept dependent advantage of virtual manipulatives should be conducted in other contexts to determine which concepts are best learned virtually and why. The simulation used in this study provided students with data for a frictionless environment and also displayed dynamically changing bar charts of the data. Further, the simulation could be manipulated very quickly so that many different combinations of data could be tested. These factors may have lead to the advantage of the virtual manipulatives in learning about work and energy from pre-test to mid-test1 and mid-test2 to post-test. From pre-test to post-test, there was no significant difference in scores based on sequence of activities performed. This indicates that the order of manipulative use does not affect the final score, thus neither the PV nor the VP sequence is preferred.

Interestingly, students in the VP sequence show a steeper decline in score between mid-test1 and mid-test2. The students who has used the virtual manipulative during week one demonstrated more conceptual understanding at the end of week one (as measured by mid-test1) but also retained less of this information at the beginning of week 2 (as measured by mid-test2) According to this finding, there may be an effect of which manipulative was used for learning and the retention of physics concepts one week later. Previous research has shown retention can be improved by increasing organization of material learned. (Lawton & Wasanka, 1977; Moore & Readance, 1984) The virtual manipulative used in this study displayed the initial parameters of the experiment, controlled variables and results in clearly marked subsections of the computer screen and included dynamically changing bar charts for each variable measured. Students collected data from the computer simulation and recorded it in a separate data table. When students took physical data, they read values from a spring scale or meter stick and recorded them in a data table identical to that used in the virtual experiment. The virtual data was collected from a source with a much higher degree of visual organization than the physical data. According to the findings of previous research, the higher level of organization when using the virtual manipulative should produce a higher level of retention one week later. The findings of our study do not support this hypothesis. Investigation of the relationship between retention and type of manipulative used should be pursued in future studies.

Students' confidence in their answers followed a somewhat predictable pattern. There was a statistically significant increase ($p < 0.05$) from pre-test to mid-test1 when students completed the first activity. This was followed by a statistically significant decline from mid-test1 to mid-test2 when students returned after one week. Finally, there was a statistically significant increase from mid-test2 to post-test when students completed the second activity. There was no statistically significant difference based either or manipulative used or on the sequence in which the manipulatives were used. These findings support Chini's (2010) survey result where students reported that they would choose to use both physical and virtual manipulatives to learn in a laboratory setting. In our study, students learning in an actual laboratory setting, as opposed to being asked about a theoretical one, reported no difference in their confidence when answering test questions based on manipulative type. This indicates that students self-perception is that they learn equally well from physical and virtual manipulatives. These self-reported perceptions of confidence in learning contrast with the actual learning gains between pre-test and mid-test1 and mid-test2 and post-test, which show an advantage for those who used the virtual manipulative. Students perceive that both are equally beneficial, though the virtual manipulative is preferred when learning about certain concepts.

Future studies investigating the use of physical and virtual manipulatives should address the issues of retention of learning after using each type of manipulative over a larger time scale. It is important to determine not only how students score on a conceptual test immediately after the intervention, but track learning over time. Further, future studies should investigate other context related advantages of physical and virtual manipulatives to determine where each is most beneficial.

References

Chini, J. J. (2010). Comparing the scaffolding provided by physical and virtual manipulatives for students' understanding of simple machines. (Doctoral dissertation). Retrieved from ProQuest. (UMI No. 343498)

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

Ebbinghaus H. (1885). Memory: A Contribution To Experimental Psychology.

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.

Lawton J. and S. Wanska (1977). Advance Organizers As A Teaching Strategy: A Reply to Barnes and Clawson. *Review of Educational Research*, *47* (1), 233-244.

Moore D. and J. Readence (1984). A Quantitative and Qualitative Review of Graphic Organizer Research. *The Journal of Educational Research, 78* (1), 11-17.

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.

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.

Zacharia, Z. C., & Olympiou, G. (2011). Physical versus virtual manipulative experimentation in physics learning. *Learning and Instruction, 21*(3), 317-331.

Table 1

*Repeated measures ANOVA and contrasts from pre- to mid-test1 and mid-test2 to post-test2.*

| Effect | Repeated Measures ANOVA | | Pretest/Midtest1 Contrast | | | Midtest2/Posttest2 Contrast | | |
|---|---|---|---|---|---|---|---|---|
| | F | p | F | p | R | F | p | R |
| Total | *F*(2.7, 317.7)=162.7 | <.001 | *F*(1, 119)=239.4 | <.001 | .82 | *F*(1, 119)=71.1 | <.001 | .612 |
| Tot* Seq | *F*(2.6, 309.2)=11.2 | <.001 | *F*(1, 119)= 23.3 | <.001 | .41 | *F*(1, 119)=6.1 | .015 | .221 |
| Force | *F*(2.8, 334.0)=161.3 | <.001 | *F*(1, 119)= 287.1 | <.001 | .84 | *F*(1, 119)=51.2 | <.001 | .548 |
| For*Seq | *F*(2.8, 334.0)=1.86 | .136 | -- | -- | -- | -- | -- | -- |
| WE | *F*(2.8, 329.0)= 34.1 | <.001 | *F*(1, 119)= 51.3 | <.001 | .55 | *F*(1, 119)=14.4 | <.001 | .329 |
| WE*Seq | *F*(2.8, 329.0)= 14.2 | <.001 | *F*(1, 119)= 25.8 | <.001 | .42 | *F*(1, 119)=12.5 | .001 | .308 |

Table 2

Repeated measures ANOVA and contrasts from mid-test1- to mid-test2 and pre-test to post-test.

| Effect | Repeated Measures ANOVA | | Midtest1/Midtest2 Contrast | | | Pretest/Posttest Contrast | | |
|---|---|---|---|---|---|---|---|---|
| | F | p | F | p | R | F | p | R |
| Total | $F(2.7, 317.7)=162.7$ | <.001 | $F(1, 119)=19.5$ | <.001 | .38 | $F(1, 119)=313.9$ | <.001 | .85 |
| Tot*Seq | $F(2.6, 309.2)=11.2$ | <.001 | $F(1, 119)=6.3$ | .014 | .22 | $F(1, 119)= 1.0$ | .706 | .09 |
| Force | $F(2.8, 334.0)=161.3$ | <.001 | $F(1, 119)=17.2$ | <.001 | .36 | $F(1, 119)= 323.6$ | <.001 | .85 |
| For*Seq | $F(2.8, 334.0)=1.86$ | .136 | -- | -- | -- | -- | -- | -- |
| WE | $F(2.8, 329.0)= 34.1$ | <.001 | $F(1, 119)=5.3$ | .023 | .21 | $F(1, 119)= 70.4$ | <.001 | .61 |
| WE*Seq | $F(2.8, 329.0)= 14.2$ | <.001 | $F(1, 119)=3.8$ | .052 | .18 | $F(1, 119)= .579$ | .448 | .07 |

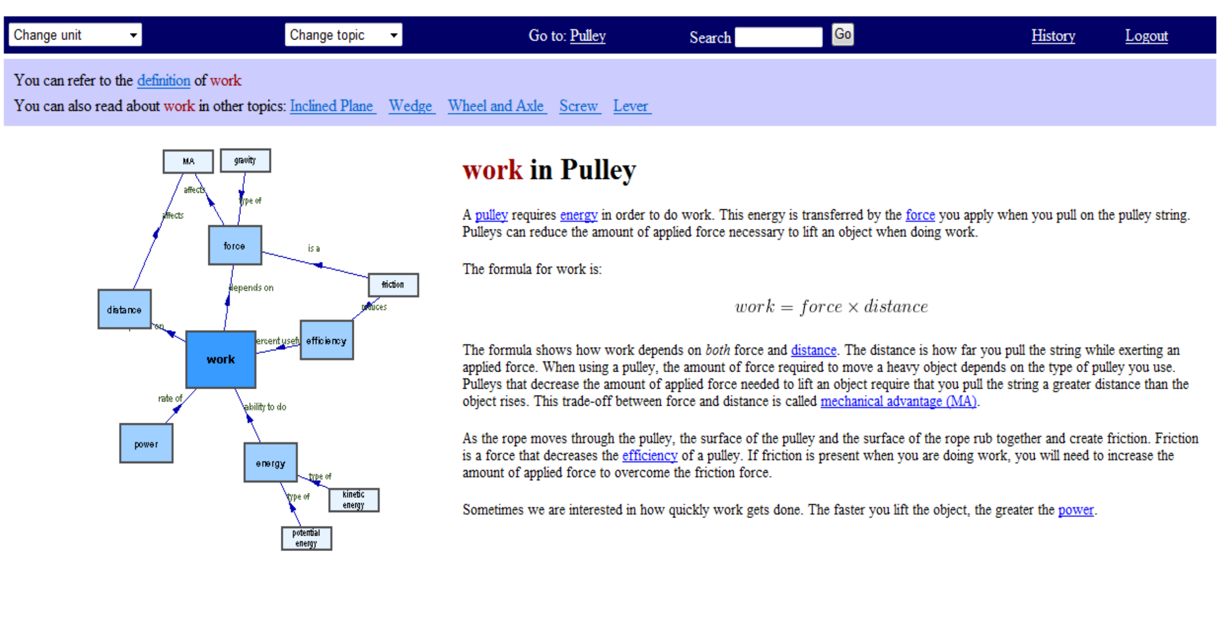



*Figure 1*. CoMPASS, dynamic concept maps and hypertext-based environment.

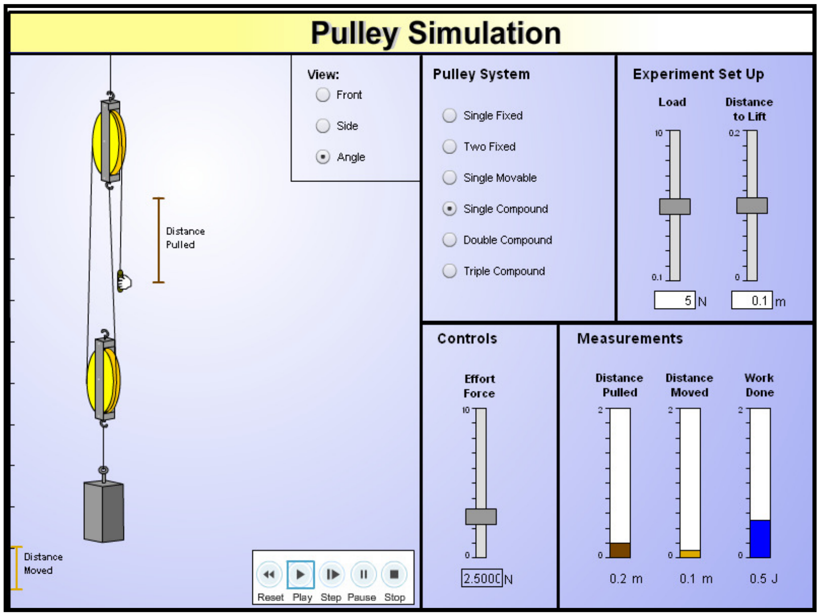



*Figure 2***.** Screen shot of simulation.

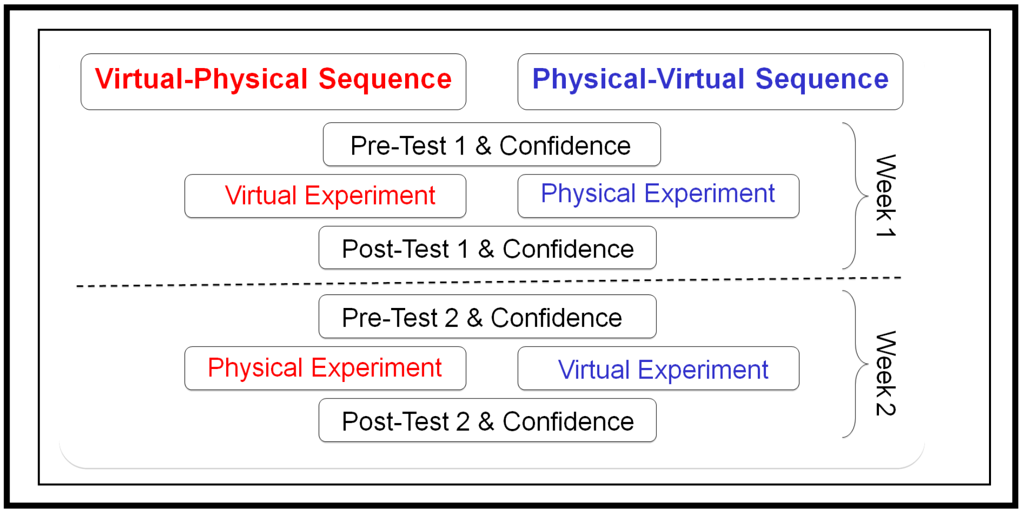
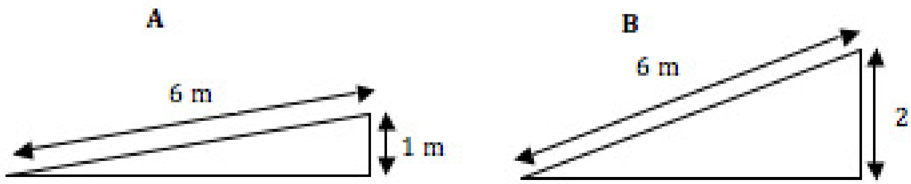



*Figure 3.* Sequence of activities performed by students.

12) Jacob is using ramps to lift two boxes of the same size and mass up to two different heights. He lifts one box to the top of Ramp A and then lifts the second box to the top of Ramp B. Ignoring friction, when lifting the box to the top of Ramp B Jacob is doing ______ *work* as/than when lifting the box to the top of ramp A?




A.) More

B.) Less

C.) Same amount of

D.) Not enough information to decide

12a) Explain your reasoning about *work* in this question.

| Work | | |
|---|---|---|
| **Points** | **Description** | **Example** |
| **0** | Incorrect response | "If the ramp was 10 meters long the work would increase." |
| **1** | Saying that the height for Ramp B is greater than for Ramp A<br><br>Talk about the idea that Ramp B is steeper than Ramp A. Focus on the angle or steepness rather than height. Relate this to the increased amount of applied force and therefore increased work.<br>OR<br>Use lay terms in their explanations (e.g. makes work harder or easier) | "Because in a it is only 1m tall and in b it is 2m tall"<br><br>"He's doing more because ramp B is steeper then ramp A. so then you would have to work harder to get the load up the ramp." |
| **2** | Must talk about work in terms of how the change in the height affects the amount of applied force needed to move an object over the same distance<br>OR<br><br>Discussion of the relationship between energy and height or energy and work | "the height will cause the force to increase and they are both the same length. "<br><br>"He's doing more because ramp B is steeper then ramp A. so then you would have to apply more force over the same distance to get the load up the ramp" |

*Figure 4*. Example of rubric used to score "explain your reasoning" questions on conceptual tests.

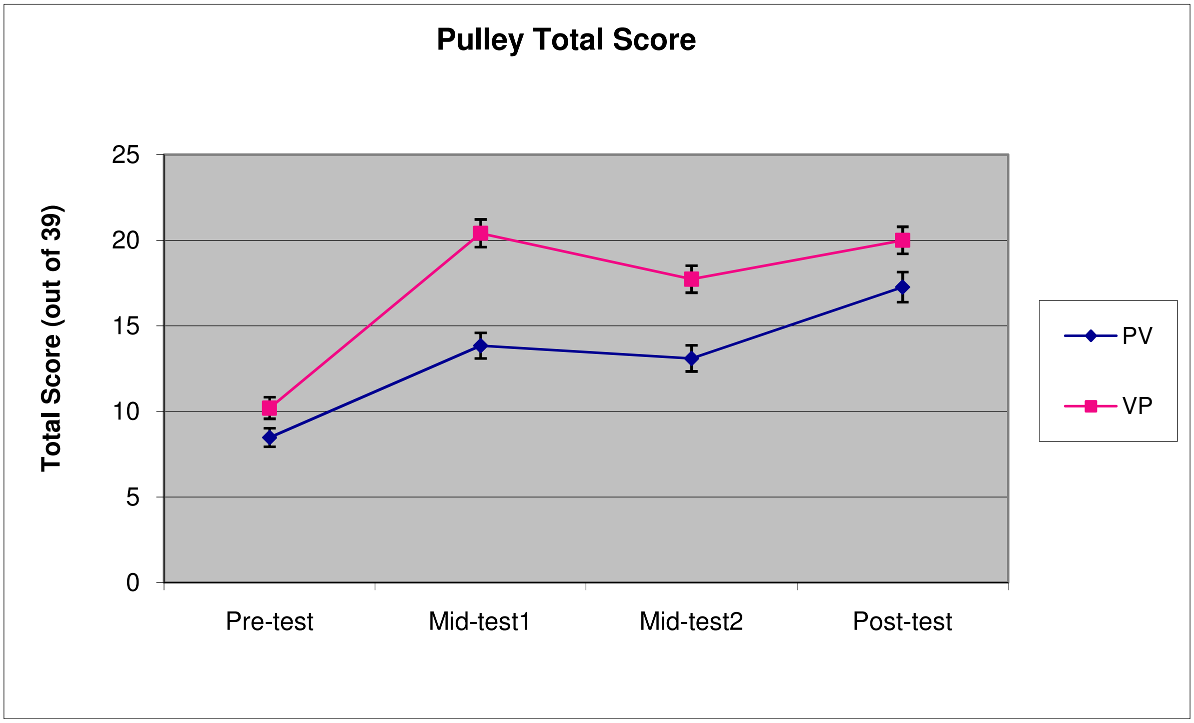



*Figure 5*. Mean total score on tests. Error bars show standard error.

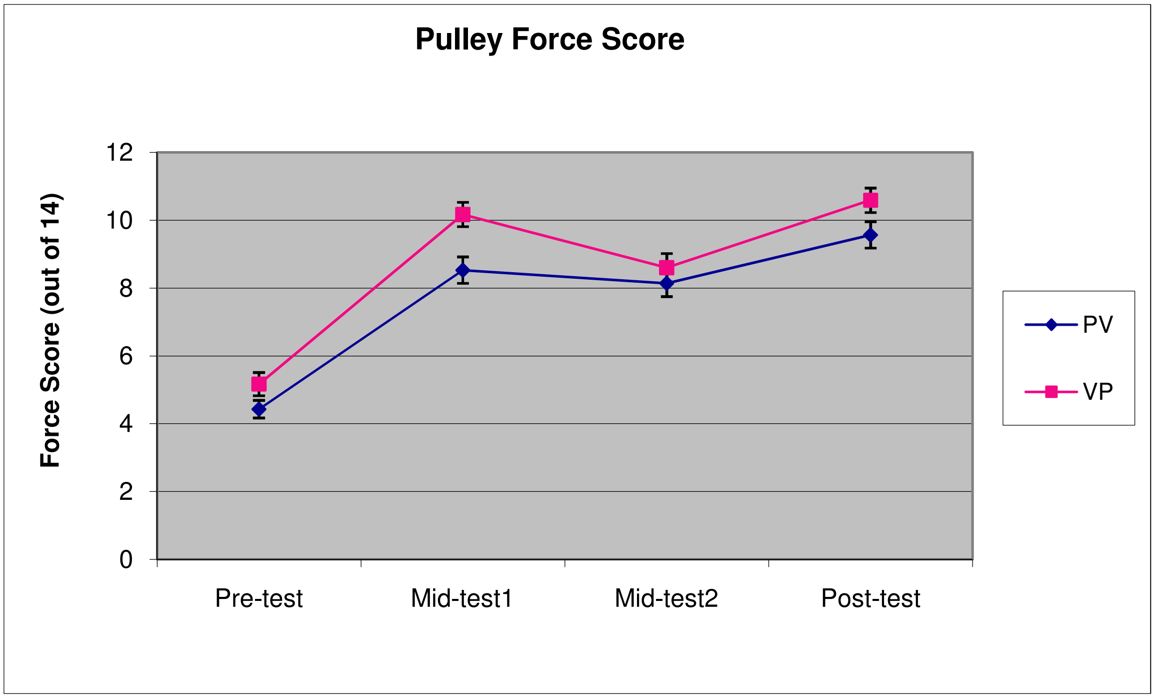



*Figure 6*. Mean force score on tests. Error bars show standard error.

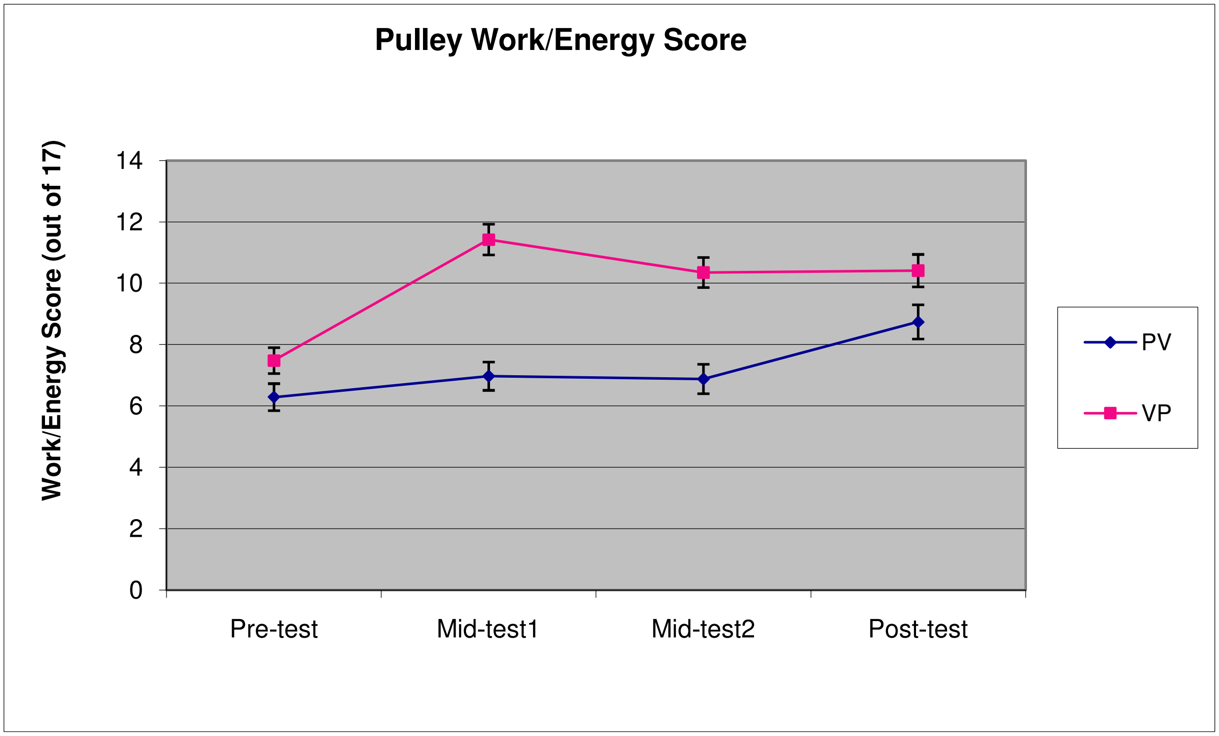



*Figure 7.* Mean work/energy score on tests. Error bars show standard error.

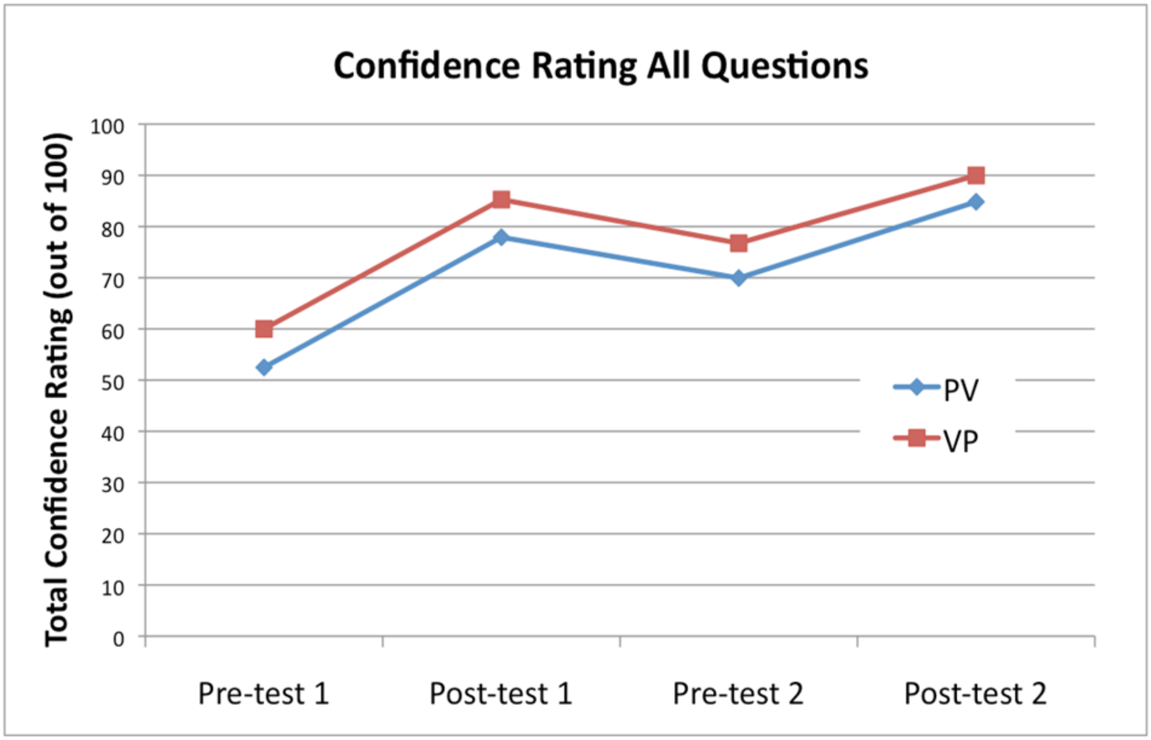



*Figure 8*. Total confidence rating on all questions.