The Effects of Physical and Virtual Manipulatives on Students' Conceptual Learning About Pulleys

Elizabeth Gire, Adrian Carmichael, Jacquelyn J. Chini, Amy Rouinfar & Sanjay Rebello, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506
egire@phys.ksu.edu, adrianc@phys.ksu.edu, haynicz@phys.ksu.edu, amy.rouinfar@gmail.com, srebello@phys.ksu.edu
Garrett Smith and Sadhana Puntambekar, University of Wisconsin, 1025 West Johnson Street, Suite 785, Madison, WI 53705
gwsmith@wisc.edu, puntambekar@education.wisc.edu

Abstract: With computers becoming more ubiquitous in our daily lives and in our classrooms, questions of how students interact and learn with physical experiments and computer simulations are central in science education. We investigated how students' ideas about pulleys were influenced by the use of physical and virtual manipulatives. We found that there were advantages for each type of manipulative, and that virtual and physical manipulatives helped students develop correct understandings of different concepts. We also found that the order the manipulatives were used affected student learning, with students who used real pulleys before the simulation achieving higher scores on questions having to do with effort force, the distance the rope is pulled, and mechanical advantage.

Introduction & Background

Laboratory experiments play a critical role in furthering scientists' understandings of how the universe works, and in light of this importance, it is no wonder that educators have historically placed high value on laboratory experiences in science classrooms. However, due to practical concerns of procuring laboratory equipment, safety concerns, and time constraints, computer simulated experiments are becoming an attractive alternative to laboratory experiments. In light of this trend, recent research in science education has explored whether computer simulations (virtual manipulatives) can be as effective for learning as experiments involving real objects (physical manipulatives) and researchers have begun looking at the circumstances in which these two alternatives may be best employed.

Finkelstein *et al.* (2005) investigated how physical versus virtual manipulatives supported students' learning about circuits. Students used either physical materials or simulations to examine combinations of resistors, build simple circuits, predict the behavior of specific elements and develop a method for measuring resistance. The simulations were similar to the set-up with physical materials, except that the simulations represented electron flow within the circuit, an aspect of the physical materials that cannot not directly be perceived. After these experiences, students who had used the virtual manipulatives were able to build physical circuits quicker than students who had previously used the physical manipulatives. In addition, the students in the virtual conditions were able to provide better explanations of circuit behavior and scored better on a related exam question. Therefore, Finkelstein *et al.* suggest properly designed simulations can be beneficial to student learning when applied in the appropriate contexts.

Triona, Klahr and Williams (2007) investigated how physical and virtual manipulatives support students' learning about the factors affecting how far a mouse trap car will travel. Students explored these factors by designing cars to be used for an experiment. Students used either physical or virtual manipulatives and were allowed to design either a certain number of cars or were allowed to design cars for a certain length of time, creating four treatment groups. All treatments were equally effective at increasing students' knowledge about causal factors for travel distance, supporting students' ability to design cars, and students' confidence in their knowledge. Based on these findings, the researchers suggest that simulations may be preferred due to their other pragmatic advantages.

Zacharia, Olympiou, & Papaevipidou (2008) studied physical and virtual manipulatives used in combination to learn about heat and temperature. Students in the control group used only physical manipulatives, while students in the experimental condition used physical manipulatives followed by virtual manipulatives. The researchers aimed to limit the differences between the physical and virtual manipulatives to speed of manipulation. On a conceptual test, students in the experimental group outperformed students in the control group. The researchers suggest this difference may be a result of virtual manipulatives being manipulated faster than physical manipulatives.

In a similar study (Zacharia & Constantinou, 2008), the researchers controlled for all differences between the physical and virtual conditions except for the mode in which experiments were performed. In particular, the simulations did not model any aspects of the phenomena that could not be perceived with the physical manipulatives (in contrast to Finkelstein, et al). In this case, the physical and virtual manipulatives equally supported students' conceptual understanding.

In our study, we also controlled for all conditions (curriculum, mode of instruction and resource capabilities) except for the mode of the activities (physical or virtual). Students spent approximately 30 minutes on each activity, although working with the real pulleys typically took a few minutes longer than working with the simulation. The design of this study replicates the study performed by Zacharia & Constantinou in a new domain (pulleys rather than heat & temperature). Furthermore, we not only looked at overall learning during these activities, but isolated particular concepts and looked at the effect of manipulative type and ordering of manipulatives on students' understandings of these concepts.

Context and Data Collection

Students in a university-level conceptual physics lab performed two activities to learn about pulleys. One activity involved working with real pulleys (physical manipulatives) while the other activity involved an interactive computer simulation of pulleys (virtual manipulatives) (Figure 1). The activities are part of CoMPASS, a design-based curriculum that integrates concept maps and hypertext that students explore prior to performing physical or virtual experiments (Puntambekar, Stylianou & Hübscher, 2003 and Puntambekar & Stylianou, 2002). During each activity, students answered questions on a worksheet. These worksheet questions were the same for both activities. However, the temporal order of the activities was varied creating two treatment groups. Three sections (N=71) used the physical pulleys first (the Physical-Virtual treatment), while two sections (N=61) began with using the virtual pulleys (the Virtual-Physical treatment).

Students answered a set of conceptual assessment questions before the activities (pre-test), after the first activity (mid-test), and after the second activity (post-test). The assessment questions on the pre-, mid-, and post-tests were identical. The mid-test scores allowed for comparisons to be made between the effects of physical manipulatives (PM) and the effects of virtual manipulatives (VM) only, while mid-test and post-test scores indicated ordering effects.

The assessment contains 13 multiple-choice questions, with each question weighted equally in the total score. The assessment questions were developed locally to probe students' conceptual understanding of pulley concepts, including effort force, work, mechanical advantage, the distance the rope is pulled and the potential energy of the load. The assessment contained more questions about effort force and work concepts than other concepts because these are the most central to the topics of pulleys and the most applicable in other science topics. Figure 2 indicates the distribution of questions for each concept. A question was considered to be related to a concept when the concept is explicitly mentioned in the problem statement. For example, the question "If we ignore friction, what will require *less effort* (force) to lift a box to a height of 1 meter – using the pulley shown or lifting the box straight up?" is considered to be an effort force question.

A reliability analysis was conducted on effort force and work questions. For the effort force questions Cronbach's $\alpha = .70$. For the work questions, Cronbach's $\alpha = .51$. The lower reliability for the work questions may indicate that students have a harder time constructing correct understandings about the concept of work.

Open-ended worksheet questions were coded and analyzed using a phenomenographic approach (Marton, 1986). The conceptual assessments were analyzed statistically. For these assessments, categories of questions were created based on the physics concept probed by each question (as indicated explicitly in the question statement). The analysis included comparisons of overall scores and category scores that were made using a Repeated Measures Analysis of Variance with a between subjects factor of treatment type. P-values less than 0.05 were interpreted to indicate a statistically significant difference.

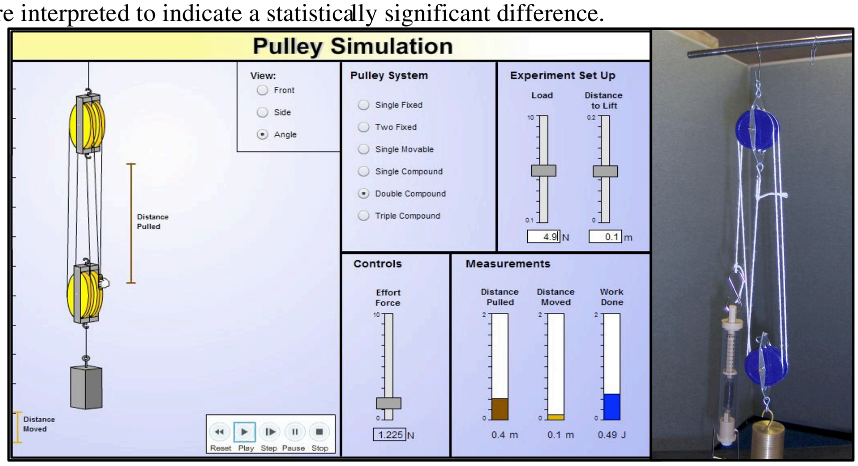


Figure 1. Virtual (left) and physical (right) manipulatives.

Data Analysis

Conceptual Assessment

The overall scores for the pre-, mid- and post-tests are shown in Table 1. Table 2 contains the results of a Repeated Measures Analysis of Variance for all the students. Mauchly's test indicated that the assumption of sphericity had been violated on all the comparisons made, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity. Table 2 also contains information of Mauchy's test and sphericity estimates.

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 %
Physical-Virtual	71	**37 ± 2**	**58 ± 2**	**66 ± 3**
Virtual-Physical	61	**33 ± 2**	**60 ± 3**	**61 ± 3**

Table 2: Mauchly's Test, Greenhouse-Gessier Estimates of Sphericity, and Repeated Measures ANOVA for Overall Score, Force Questions, and Work Questions

Effect	Mauchly's Test		Sphericity	Repeated Measures Analysis of Variance	
	$\chi^2(2)$	p	ε	F	p
Total Score	67.28	<.001	.71	F(1.42,184.87) = 173.57	<.001
Total Score*Treatment				F(1.42,184.87) = 2.33	0.12
Effort Force Questions	61.40	<.001	.73	F(1.45,188.58) = 167.24	<.001
Effort Force*Treatment				F(1.45,188.58) = 4.89	0.02
Work Questions	25.81	<.001	.85	F(1.69, 220.09) = 27.69	<.001
Work*Treatment				F(1.69, 220.09) = 15.28	<.001

The Repeated Measures analysis shows that students' total scores changed significantly between tests. However, the insignificance of the interaction between total score and treatment condition indicates that the changes in scores for the two treatments were similar. The scores for the effort force questions also show a significant change between tests and a significant interaction between the effort force questions and the treatment conditions indicating that the changes in effort force score were different for the two conditions. Similarly, the scores for the work questions show a significant change between tests and a significant interaction between work score and treatment condition.

Table 3 shows the results of contrast comparisons for each effect described in Table 2. These comparisons help locate when significant changes in score occurred. Both treatment conditions resulted in a change of total score between the pre- and mid-test, but the scores of the treatment conditions changed differently between the mid- and post-test. Figure 2 shows plots of average score on each exam. The plot shows that students in the physical-virtual condition had a significantly greater change in score between the mid- and post-test. Therefore, the students who used the physical manipulatives first continued to progress on the conceptual assessment after the second activity while the students who used the virtual manipulatives for showed little further progression, although students in both treatment conditions benefitted from the activities on the whole.

On Effort Force questions, both conditions showed a change in score between the pre- and mid-test, but no change in score was observed between the mid- and post-test. The change in score for the treatment conditions was significantly different between the pre- and mid-test, with students who initially used the physical manipulatives showing a larger gain in Effort Force score than students who initially used the virtual manipulatives.

On the Work questions, an overall change in score was observed for between pre- and post-test. However, the changes were significantly different for the different treatments. Students in the virtual-physical condition show a large gain in Work score between pre- and mid-test while the students in the physical-virtual condition do not show a gain. However, this trend switches between the mid- and post-test, with the virtual-physical students showing no gain and the physical-virtual students showing a large gain. It seems that students show a large increase in score on work questions after they have used the pulley simulation.

For the other questions dealing with mechanical advantage, distance pulled, and potential energy, students in both conditions showed an increased of score between pre- and post-tests but no interaction effect was observed for these questions.

Table 3: Contrast Comparisons for Repeated Measures ANOVA

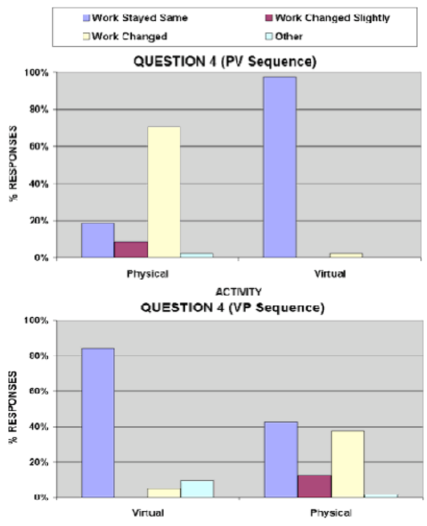
Effect	Comparison	F(1,130)	p
Total Score	Pre-Mid	170.94	<.001
	Mid-Post	22.33	<.001
Total Score*Treatment	Pre-Mid	1.71	.19
	Mid-Post	12.04	.001
Effort Force Questions	Pre-Mid	181.30	<.001
	Mid-Post	1.74	.19
Effort Force*Treatment	Pre-Mid	5.56	.02
	Mid-Post	<.001	.98
Work Questions	Pre-Mid	10.83	.001
	Mid-Post	26.27	<.001
Work*Treatment	Pre-Mid	24.18	<.001
	Mid-Post	28.14	<.001

Figure 2. Average scores by category on the conceptual assessments. Error bars indicate standard error.

Questions on Activity Worksheet

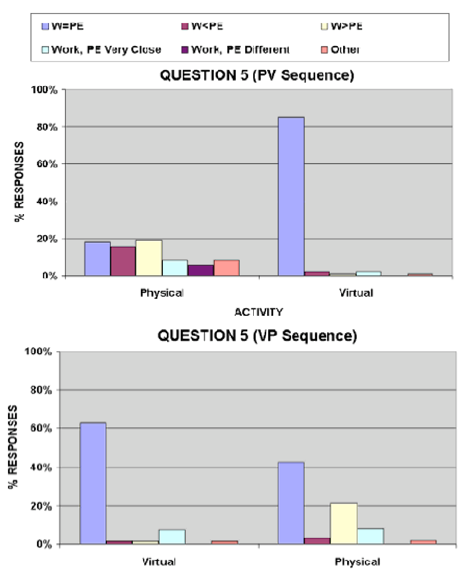
While doing the activities (with physical pulleys and virtual pulleys), students responded to questions on a worksheet. We report students' responses on some of the worksheet questions to aid interpreting the above assessment results. Question 4 on the worksheet had to do with the concept of work (Figure 3). Students in both treatments interpreted the data from the simulation as showing the work being the same for different pulley-setups. In contrast, students in the different treatments disagreed on how to interpret the data from the real pulley: students in the Physical-Virtual treatment said the work changed when you changed set-ups while students in the Virtual-Physical treatment were split, with nearly half of the students claiming the work stayed the same while the other half said the work changed across pulley set-ups.

A similar trend was seen on Question 5 on the worksheet (Figure 4). Question 5 had to do with comparing work and potential energy for a given pulley system. Students in both treatments interpreted the data from the simulation as showing the work was equal to the potential energy for a given pulley set-up. In contrast, students in the different treatments disagreed on how to interpret the data from the real pulley: students in the Physical-Virtual treatment did not come to a consensus about how the work was related to the potential energy while students in the Virtual-Physical treatment were more likely to say the work was equal to the potential energy.



Question WS4: "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 3. Student responses to Question 4 on the activity worksheet.



Question WS5: "Based on your data, how does work compare to potential energy for a given pulley system? Why do you think that is?

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

Discussion

The analysis of overall score indicates that there is no overall preferred manipulative for learning about pulleys; the students' mid-test scores (which isolate each type of manipulative) are the same. Looking at changes between the mid-test and post-test scores, however, indicates that there may be an ordering effect for using both types of manipulatives, in that the scores of students in the Physical-Virtual treatment continued to increase after the second activity, while the scores of students who used virtual pulleys first did not increase when the students switched to physical pulleys.

In looking at categories of questions, the data show that the different types of manipulative (physical or virtual) affected different pulley concepts differently. In the Effort Force category, although an increase in score was observed on mid-test for both treatments, students who used real pulleys scored higher. In contrast, students showed no increase in score on questions about work after using real pulleys, while students who used the pulley simulation showed significant improvement on work questions.

Additionally, the Effort Force category showed an ordering effect on category score, while the Work category did not. Students who used the physical pulleys first showed a greater increase in score between the pre- and mid-tests, while neither treatment group showed an increase in score between the mid- and post-tests. In contrast, students who used virtual pulleys more often answered questions about work correctly than students who only used physical pulleys, regardless of which order they encountered the manipulatives. In light of these data, it seems the ordering effect seen in the overall score may be explained by the ordering effect of questions having to do with effort force.

In short, students obtain a better understanding as measured by the conceptual assessment of the concept of work with the computer simulation and a better understanding of effort force with the real pulleys.

Why should the type of manipulative affect different pulley concepts differently? One possibility is that this result is due to a difference in the salience of the physics concepts. The concepts of effort force and distance pulled are more salient for real pulleys than with virtual pulleys. The real pulleys give students a kinesthetic experience with effort force and distance pulled; they feel the force they need to exert and how far their arms move in relation to the pulley, while students who use the virtual pulleys read-off values from a screen. On the other hand, work is less salient than effort force for real pulleys. Work depends on the combination of effort force and distance pulled, and this makes the concept of work more removed from the kinesthetic experience. In order to reason about work correctly, students need to coordinate the experiences that if less force is needed, the distance pulled is longer in exactly the right proportion so that work is constant across pulley set-ups (assuming friction is small enough to be neglected). Alternately, students might perceive the energy expended in lifting a load a certain distance with different pulley set-ups. Both of these types of reasoning are difficult to achieve through kinesthetic experience. However, all concepts are equally salient in the simulation – all quantities must be read-off the screen (Figure 1).

Capacity theories of attention suggest that people can only attend to a limited amount of information (Kahneman, 1973), and it is also known that attention can be influenced by the salience of cues, with high salience cues naturally attracting more attention (Denton & Kruschke, 2006). With physical manipulatives, effort force and distance pulled have a relatively high salience and probably dominate the attention of the learner, while less salient concepts, like work, receive less attention. This explanation is consistent with our data that Physical-Virtual students have relatively high scores on Effort Force questions and relatively low scores on Work questions. Students who use the virtual manipulatives probably divide their attention more evenly among the concepts due to their equivalence in salience, resulting in less relative attention to effort force, distance pulled and mechanical advantage concepts, resulting in lower scores in these categories and higher scores on work questions than students who used physical manipulatives. Furthermore, if the simulation is done first, the initial equivalence in salience among concepts may lessen the impact of the subsequent kinesthetic experience with the real pulleys, resulting in the continued suppression, or blocking (Kamin, 1968), observed in the Effort Force, Distance Pulled scores, and Mechanical Advantage categories.

Furthermore, the students in this study did not receive instruction on how to interpret their data in light of friction effects and experimental uncertainty. The data show that they did not interpret trends in their real pulley work data in a consistent way. This point is supported by the students' responses to Question 4 and Question 5 on the worksheet. Also, Physical-Virtual students had trouble reasoning about work in frictionless situations on the mid-test. However, after the virtual experiment, all students were much more successful in reasoning about work in frictionless environments, regardless of the order the activities.

This study suggests that curricula that include pulleys might ideally use both experiments with real pulleys and simulations. The data suggest that an ideal ordering might be to have students begin with real pulley experiments, focusing on effort force, distance pulled and mechanical advantage. Then students might perform a set of experiments with the simulation, reinforcing trends found with the real pulleys and exploring the concepts of work and potential energy in a frictionless simulated environment. Finally, students might finish with a set of experiments with real pulleys to explore how work and potential energy are related in a situation where friction is not negligible.

Conclusion

Previous research has demonstrated that virtual manipulatives can be as effective as physical manipulatives in some circumstances. Our data extends this result to the domain of pulleys and addresses how manipulative types affect different pulley concepts. In looking at the effect of physical and virtual manipulatives on students' understandings of pulleys, we find that although there is no difference in overall score between types of manipulative, each manipulative has advantages for different pulley concepts: physical manipulatives better address the concepts of effort force, distance pulled and mechanical advantage, while virtual manipulatives better address the concept of work. The order the manipulatives are used by the students affected conceptual gains for the concepts of effort force, distance pulled and mechanical advantage. We suggest that these differences may be attributed to differences in concept salience.

Endnotes

(1) The assessment and worksheets can be found online at http://www.phys.ksu.edu/personal/egire/Resources.html

References

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

Denton, S.E. & Kruschke, J.K. (2006) Attention and salience in associative blocking. *Learning & Behavior,* 34(3), 285-304.

Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Kohl, K. K., Podolefskey, N.S., 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.

Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior: Aversive stimulation* (pp. 9-33). Coral Gables, FL: University of Miami Press.

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 understanding of reality. *Journal of Thought, 21*, 29-39.

Puntambekar, S., Stylianou, A., & Hübscher, R. (2003). Improving navigation and learning in hypertext environments with navigable concept maps. *Human-Computer Interaction, 18*, 395-428.

Puntambekar, S. & Stylianou, A. (2002). CoMPASS: Students' use of external representations in science learning. In P. Bell, R. Stevens & T. Satwicz (Eds.), *Keeping Learning Complex: The Proceedings of the Fifth International Conference of the Learning Sciences (ICLS)* (pp.352-358). Mahwah, NJ: Erlbaum.

Triona, L. M. & 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&5), 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.

Acknowledgments

Foremost, we would like to thank the students and staff at CoMPASS Group, Wiconsin Center for Educational Research for their cooperation. We would also like to thank the students who participated in this study, particularly those who volunteered to be interviewed. Their cooperation and candor was critical to this work. Additionally, we'd like to thank the members of the Physics Education Research Group at Kansas State University for fruitful discussions and their thoughtful suggestions. This work was supported in part by the National Science Foundation under DUE grant DGE 0841414 and U.S. Department of Education IES Award R305A080507. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation or the U.S Department of Education.