

Comparing Physical and Virtual Manipulatives for Retention and Preparation for Future

Learning of Science Concepts

Amy Rouinfar¹, Adrian M. Madsen¹, Jacquelyn J. Chini², Elizabeth Gire³, N. Sanjay Rebello¹

and Sadhana Puntambekar⁴

¹Kansas State University

²University of Central Florida

³University of Memphis

⁴University of Wisconsin, Madison

Abstract

Previous research on the effects of physical and virtual experimentation has focused on comparing the two manipulatives for learning. In this study we focus on which of the two manipulatives better support retention of knowledge learned and preparation for future learning. 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 over two consecutive weeks. 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 implications for retention and preparation for future learning. Our results seem to indicate that the physical manipulative better supports retention than the virtual manipulative. Further, learning with the physical manipulative better prepares students for future learning with the virtual manipulative.

Comparing Physical and Virtual Manipulatives for Retention and Preparation for Future
Learning of Science Concepts

Background and Introduction

With the advent of computers there has been a movement to replace hands-on experiments with computer simulations. Several researchers have studied the comparative benefits afforded by computer simulations over hands-on activities (e.g. Finkelstein, et. al. 2005; Klahr, Triona & Williams, 2007; Zacharia & Constantinou, 2008; Zacharia, Olympiou, & Papaevripidou, 2008). Taken together, these studies appear to suggest that depending upon the context of the study, virtual manipulatives either outperform or perform equally well as physical manipulatives.

In our own research we recognize that physical and virtual manipulatives may have different affordances and therefore we have not focused merely on comparing the effectiveness of the two manipulatives, rather we have focused on how students learn from *both* manipulatives in different sequences (Gire, *et al.*, 2010),

The focus of the present study is to investigate which of the two manipulatives (physical or virtual) provide better support for retention of learning. Also, which of the two manipulatives (physical or virtual) better prepare students for future learning with the other manipulative (virtual or physical). The research question is: Which of the two manipulatives – physical or virtual—when used in learning about pulleys, is more likely to enable students to:

- retain the science concepts they have learned one week after their learning episode?
- prepare for future (re)learning of these concepts using the second manipulative one week after their first learning episode?

Theoretical Underpinnings

We hold a constructivist view of learning which acknowledges that students construct their own understanding. While constructivist theory emphasizes that students should be actively involved in the learning process, Triona and Klahr (2003) have suggested that active involvement does not require physical manipulation of the learning resource. While prior studies have focused on investigating which of these two manipulatives better supports learning, there have been few studies, if any on examining whether one of these manipulatives better supports retention and transfer of learning.

Retention is the ability to remember or recall what one has learned previously in a new situation. The study of retention is among the oldest areas of formal study in the science of learning, dating back to Ebbinghaus's (1885) study of spacing effects in the 1880s. In order to retain what one has learned 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). Both physical and virtual manipulatives, each with their own affordances or constraints may better support retention. For instance, physical manipulatives have been shown to facilitate retention in younger learners (Reed, 2005). Alternatively, virtual manipulatives might better facilitate organization of information into a schema and hence support retention. Thus, in this study we examine whether learning using a physical or a virtual manipulative better supports retention one week after the study is completed.

While retention focuses on recalling information, transfer focuses on the ability to apply the information. The literature is replete with evidence that transfer is rare (Detterman & Sternberg, 1993). Some researchers have argued for a different perspective on transfer. Rather

than characterize transfer as a process in which knowledge learned in one situation is utilized intact in the new situation, researchers have expanded their view to include transfer as the process of constructing knowledge in the new situation. Among these contemporary perspectives of transfer, the perspective most relevant to this study is the one proposed by Bransford and Schwartz (1999). They contend that an appropriate measure of transfer is one that focuses on the extent to which the learner becomes better prepared for learning in the future i.e., the ‘preparation for future learning’ (PFL). To assess transfer through this prism we must assess whether students are able to *learn* in a new situation. They contrast their PFL model of transfer with the more traditional model of transfer which they refer to as ‘sequestered problems solving’ (SPS). As per the SPS model, researchers examine whether learners can apply what they have learned in one situation to a new situation. However, as per the PFL model, the question researchers ask is whether students’ experiences in a learning situation prepare them for future learning in a new situation. Thus, it is not as important whether learners can perform a novel task without any scaffolding, rather whether learners have become better prepared to utilize the scaffolding provided to learn in the new situation.

The PFL view of transfer has particular implications for comparing physical and virtual manipulatives for learning. Rather than ask which of these two manipulatives better supports learning, we instead ask whether one of these manipulatives better prepares students for future learning with the other manipulative. In other words, based on the PFL view, we are interested in investigating whether students who use the physical manipulative show evidence of increased learning by using the virtual manipulative later, or whether the learners who use the virtual manipulative show evidence of increased learning by using the physical manipulative later.

Methodology

Non-science majors (N=121) enrolled in a conceptual-based physics laboratory at a large Midwestern U.S. university performed physical and virtual activities about pulleys over two two-hour labs spread across two weeks. The topic of pulleys was not discussed in lecture, so this laboratory experience was the first experience for these students with the topic of pulleys.

Students used the pulley unit in the CoMPASS curriculum (Puntambekar, Stylianou & Hübsher, 2003) and performed experiments with the physical and virtual manipulatives shown in Figure 1. The CoMPASS curriculum integrates hypertext with hands-on activities and simulations.

The pulley curriculum was used for two consecutive weeks. In 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. Students in both treatments 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 summarized in Figure 2. Our experiment differs from the traditional PFL design in that we do not have a “common learning resource.” We instead investigated how the use of one manipulative (physical or virtual) prepares students to learn with the other manipulative (virtual or physical).

At the beginning of the first week, students took a pre-test before using the manipulatives. Following the activity (physical or virtual) students took a mid-test (mid-test 1). At the start of the second week, students took a second mid-test (mid-test 2) before completing

the second activity. After completing the second activity, students took a post-test. The pre-test, mid-tests (1 & 2) and the post-test all had the same questions. These were conceptual questions that assessed students' understanding of the concepts of force, work, potential energy, and mechanical advantage in the context of a pulley. Multiple-choice questions were typically followed with questions that asked students to provide a reason for their answer. 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 provided in Table 1. The inter-rater reliability of over 80% was established on the rubric using two independent raters. Students could score a maximum of 39 points on the pulley test.

Students in each lab section were given the same instructions and the same amount of time to complete the activities. Both the physical and virtual activities took about 90 minutes. Those performing physical activity sometimes took a few minutes longer due to the time required to set up the various pulley systems.

Results

The scores of students who completed both the physical and virtual activities were analyzed. Fifty-eight students completed the activities in the physical-virtual (PV) sequence, and 63 students completed the activities in the virtual-physical (VP) sequence. 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 mean scores of students in the two sequences (PV and VP) on each of the four tests (Pre-, Mid1-, Mid2-, and Post) is shown in Figure 3. The repeated measures ANOVA on the total score and the interaction between total score and treatment is shown in Table 2. The

contrast between consecutive pairs of tests before and after the first activity (Pre-test1 versus Mid-test1) are shown in Table 3. The contrast between consecutive pairs of tests between the two activities (Mid-test1 vs. Mid-test2) are shown in Table 4. The contrast between consecutive pairs of tests before and after the second activity (Mid-test2 versus Post-test) are shown in Table 5. Finally, the contrast between the pre-test and post-test total scores are shown in Table 6.

From Table 2 through Table 6 we can see that students' total scores changed significantly during the treatment, $F(2.7, 317.7) = 162.7, p < .001$. The contrasts show students' total scores increased significantly from pre-test to mid-test1 ($F(1, 119) = 239.4, p < .001, r = .817$) and mid-test2 to post-test ($F(1, 119) = 71.1, p < .001, r = .612$). There was also a significant decline from Mid-test1 to Mid-test2, which indicates the lack of complete retention in the one-week interval between mid-test1 and mid-test2.

The Total*Treatment effects shown in Table 2 through Table 6 indicate whether students' scores differed between the PV and VP treatments. From Table 2 we see that students' total scores changed differently between the two treatments across the entire unit ($F(2.6, 309.2) = 11.2, p < .001$). However, from Table 6, we see that students' scores did not change differently between the two treatments from pre-test to post-test ($F(1, 119) = 1.0, p = .706, r = .091$), which indicates that the two treatments resulted in the same overall learning. This result is consistent with our previous result (Gire, *et al.*, 2010).

From Table 3 we see that students' scores changed differently between the two treatments after completing the first activity i.e. from pre-test to mid-test1 ($F(1, 119) = 23.3, p < .001, r = .405$), with the VP sequence outscoring the PV sequence. This indicates that students appeared to have learned more from the virtual activity than the physical activity in the first activity.

From Table 4 we see that students' scores changed differently between the two treatments from after completing the first activity to before starting the second activity i.e., from mid-test1 to mid-test2 ($F(1, 119) = 6.3, p = .014, r = .224$), with the VP sequence showing a greater decline than the PV sequence. Examining this result through the lens of retention of learning, it appears to indicate that students appeared to have retained more of what they had learned from the physical activity than the virtual activity. We speculate that one possible reason for this result might be that the physical activity afforded students the opportunity to learn kinesthetically and this experience of embodied cognition facilitated greater retention of the science concepts learned compared to the virtual activity in which students had no opportunities for kinesthetic learning (Lakoff & Johnson, 1999).

From Table 5 we see that students' scores also changed differently between the two treatments from before to after the second activity i.e. from Mid-test2 to Post-test ($F(1, 119) = 6.1, p = .015, r = .221$), with the PV sequence making a greater increase than the VP sequence. This appears to indicate that students learned more from the second activity in the PV sequence (i.e., the virtual activity) than from the second activity in the VP sequence (i.e., the physical activity). Examining this result through the lens of preparation for future learning, it appears to indicate that students who completed the PV sequence i.e., those who completed the physical activity first were better prepared to learn in the second activity i.e., in the virtual activity, as evidenced by the larger increase in scores due to the second activity. Comparatively, students who completed the VP sequence i.e., those who completed the virtual activity first were not as prepared to learn in the second activity i.e., in the physical activity, as evidenced by their smaller increase in scores due to the second activity.

We speculate, based on evidence gathered through observing the students perform the experiment, that one possible reason for this result might be that students who completed the physical activity first had to contend with the ambiguity of physical data which motivated them to explore these issues further with the virtual manipulative (Chinn & Brewer, 1993). However, students who completed the virtual activity first did not have to contend with ambiguous physical data. Their knowledge was thus firmly constructed through their experiences with unambiguous data and therefore they were not motivated to continue to learn in the subsequent physical activity.

Conclusions and Implications

In conclusion and in response to our research questions, we can state that based on our data, students who completed the physical activity first (PV) appeared to be outperformed by the students who completed the virtual activity first (VP). However, our data also appears to provide evidence to support the conclusion that students who completed the physical activity first (PV) were better able to retain what they had learned than those who completed the virtual activity first (VP). This finding is consistent with Reed (2005). Thus, based on our data it appears that the physical manipulative better supports retention than the virtual manipulative.

Similarly, we also find evidence in our data that students who completed the physical activity first (PV) were better prepared to learn in the future from the second activity (i.e., virtual) than the VP students. The same amount of scaffolding was provided to students in the second activity, however students who first used the physical manipulative were better prepared to utilize the scaffolding provided to them. The PV group could then make better use of the scaffolding provided to them while performing the virtual experiment leading to a greater increase in their learning as evidenced by the gain in scores from mid-test 2 to the post-test.

Based on our data it appears that learning with the physical manipulative may better prepare students for future learning with the virtual manipulative.

This study contributes to the growing body of literature in this area of assessing the benefits of computer-supported learning. The insights provided by this study are relevant to educators and educational researchers who wrestle with decisions pertaining to the relative merits of physical and virtual manipulatives for science learning. However, rather than previous studies that ask which manipulative produces better learning gains, this study investigates which manipulative is better at helping students retain their knowledge and which manipulative better prepares them for future learning. We believe that examining any instructional tool or manipulative from this perspective can provide useful insights and have important implications for instruction.

References

- Bransford, J. D., & Schwartz, D. L. (1999). Rethinking Transfer: A Simple Proposal with Multiple Implications. *Review of Research in Education, 24*, 61-100
- Chinn, C.A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research, 63*, 1-49.
- Detterman, D. K., & Sternberg, R. J. (1993). *Transfer on Trial: Intelligence, Cognition and Instruction*. Norwood, NJ: Ablex Publishing Company,
- Ebbinghaus, H. (1885). *Memory: A Contribution To Experimental Psychology*. New York, NY: Teachers College, Columbia University.
- Finkelstein, N. D., Adams, W. K., Keller, C. J., Kohl, P. B., Perkins, K. K., Podolefskey, N., & Reid, S. (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.
- Gire, E., Carmichael, A., Chini, J. J., Rouinfar, A., Rebello, S., & Puntambekar, S. (2010). The effects of physical and virtual manipulatives on students' conceptual learning about pulleys. In *Proceedings of the 10th International Conference of the Learning Sciences*, June 28- July 2, 2010, Chicago, IL, ICLS 2010 Vol. 1, 937, 2010.
- 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., & Wanska. S. (1977), "Advance Organizers as a Teaching Strategy: A Reply to Barnes and Clawson," *Review of Educational Research, 47* (1), 233-244.

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

Puntambekar, S., Stylianou, A., & Hübscher, R. (2003). Improving navigation and learning in hypertext environments with navigable concept maps. *Human Computer Interaction*, 18 (4), 395-426.

Reed, J. L. (2005). *The mind-body connection : using bodily-kinesthetic activities to increase vocabulary acquisition and retention in third-grade students*. M.S., Gratz College, Melrose Park, PA. Retrieved from <http://worldcat.org/oclc/61199805>

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

Rubric for Scoring Reasoning Questions

Score	Standard
0	Incorrect Response
1	Uses basic, everyday language without scientific explanation or provides scientific explanation that is not entirely accurate
2	Provides scientific explanation using appropriate terminology

Table 2

Repeated Measures ANOVA on Total Score

Effect	Repeated measures ANOVA	
	F	p-value
Total	$F(2.7, 317.7) = 162.7$	$< .001$
Total*Treatment	$F(2.6, 309.2) = 11.2$	$< .001$

Table 3

Contrast between Tests Before and After the First Activity

Effect	Pre-Test/Mid-Test 1 Contrast		
	F	p-value	Effect Size (r)
Total	F(1, 119) = 239.4	< .001	.817
Total*Treatment	F(1, 119) = 23.3	< .001	.405

Table 4

Contrast between Tests After First and Before Second Activity

Effect	Mid-Test 1/Mid-Test 2 Contrast		
	F	p-value	Effect Size (r)
Total	F(1, 119) = 19.5	< .001	.375
Total*Treatment	F(1, 119) = 6.3	.014	.224

Table 5

Contrast between Tests Before and After the Second Activity

Effect	Mid-Test 2/Post-Test 2 Contrast		
	F	p-value	Effect Size (r)
Total	F(1, 119) = 71.1	< .001	.612
Total*Treatment	F(1, 119) = 6.1	.015	.221

Table 6

Contrast between Pre-test and Post-test

Effect	Pre-Test/Post-Test Contrast		
	F	p-value	Effect Size (r)
Total	F(1, 119) = 313.9	< .001	.851
Total*Treatment	F(1, 119) = 1.0	.706	.091

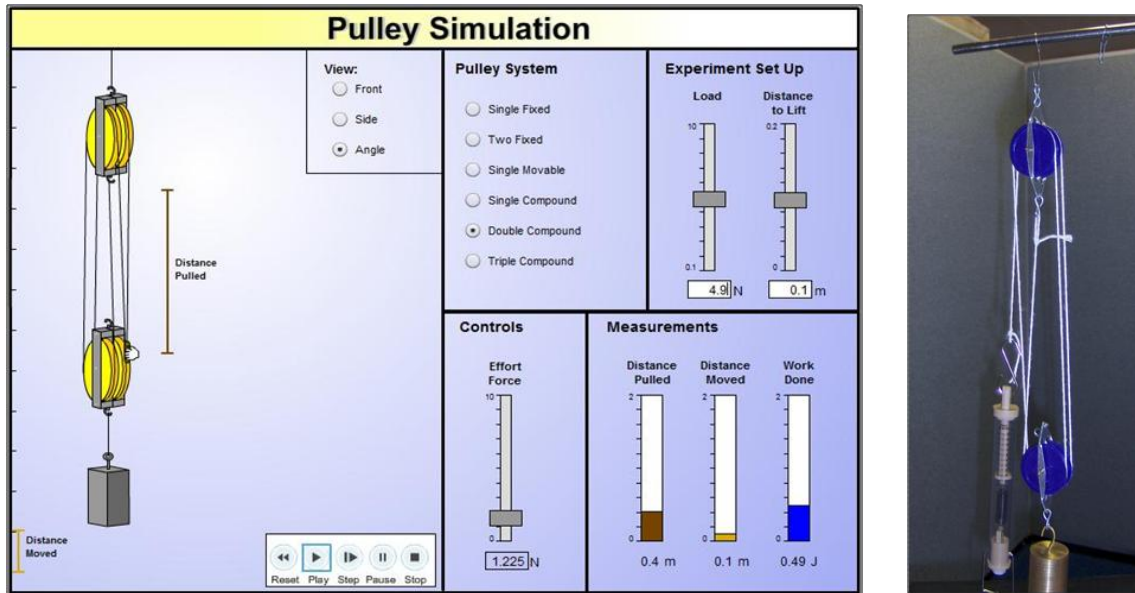


Figure 1. Virtual manipulative (left) and physical manipulative (right).

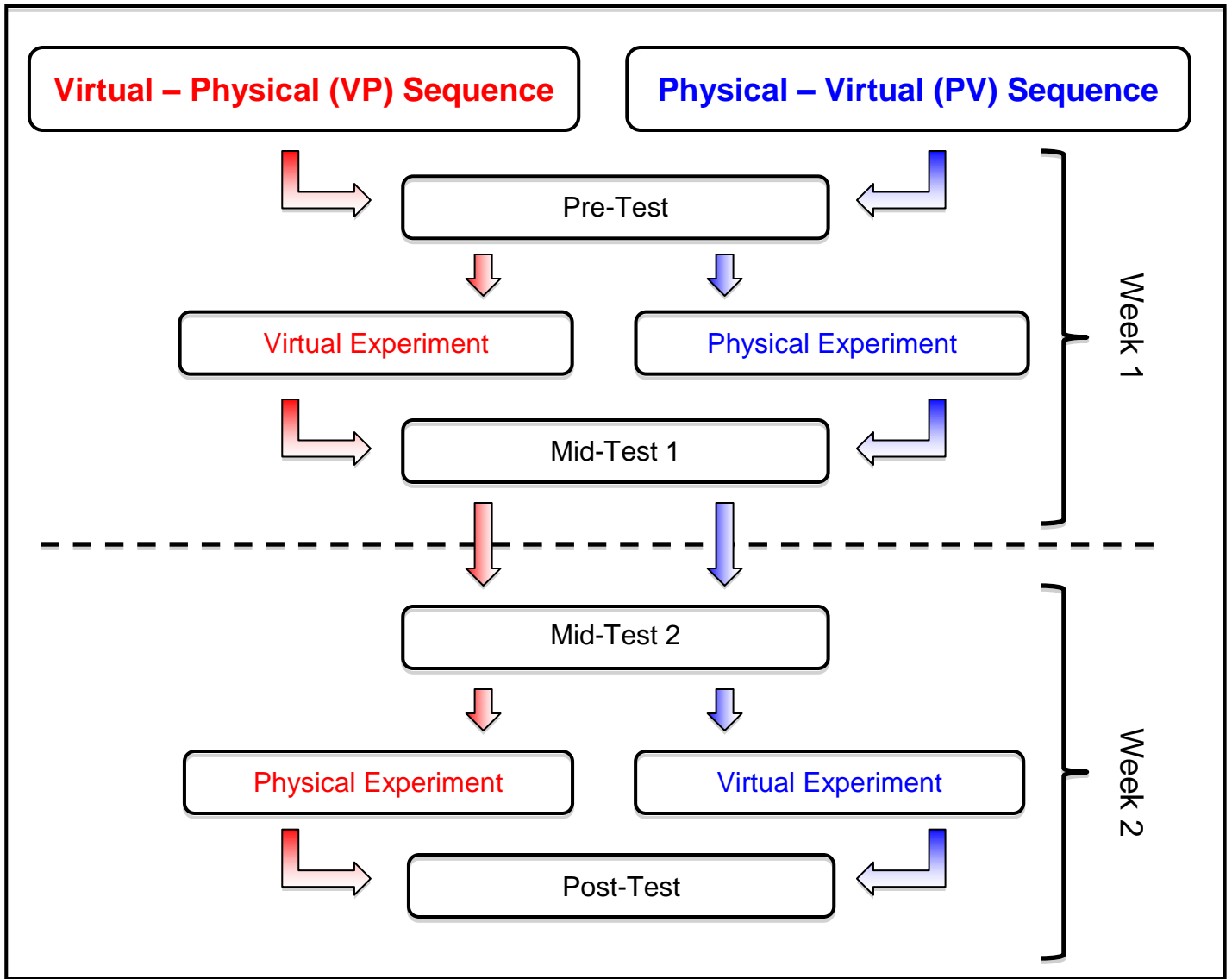


Figure 2. Sequence of activities performed by students.

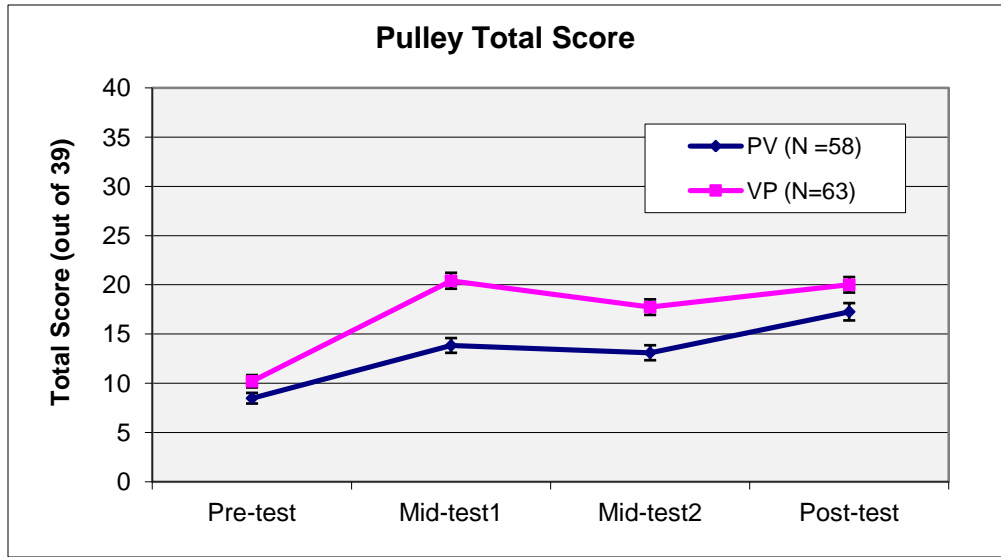


Figure 3. Average total score on each of the four tests for the two sequences. PV stands for Physical-Virtual sequence, while VP stands for Virtual-Physical sequence. The error bars represent the standard error.