

## Exploration of factors that affect the comparative effectiveness of physical and virtual manipulatives in an undergraduate laboratory

Jacquelyn J. Chini,<sup>1</sup> Adrian Madsen,<sup>2</sup> Elizabeth Gire,<sup>3</sup> N. Sanjay Rebello,<sup>2</sup> and Sadhana Puntambekar<sup>4</sup>

<sup>1</sup>*Department of Physics, University of Central Florida, 4000 Central Florida Boulevard, Orlando, Florida, 32816-2385, USA*

<sup>2</sup>*Department of Physics, 116 Cardwell Hall, Kansas State University, Manhattan, Kansas 66506-2601, USA*

<sup>3</sup>*Department of Physics, 216 Manning Hall, University of Memphis, Memphis, Tennessee, 38152, USA*

<sup>4</sup>*Department of Educational Psychology, University of Wisconsin, Madison, 1025 West Johnson Street, Madison, Wisconsin, 53706-1796, USA*

(Received 4 November 2011; published 4 April 2012)

Recent research results have failed to support the conventionally held belief that students learn physics best from hands-on experiences with physical equipment. Rather, studies have found that students who perform similar experiments with computer simulations perform as well or better on measures of conceptual understanding than their peers who used physical equipment. In this study, we explored how university-level nonscience majors' understanding of the physics concepts related to pulleys was supported by experimentation with real pulleys and a computer simulation of pulleys. We report that when students use one type of manipulative (physical or virtual), the comparison is influenced both by the concept studied and the timing of the post-test. Students performed similarly on questions related to force and mechanical advantage regardless of the type of equipment used. On the other hand, students who used the computer simulation performed better on questions related to work immediately after completing the activities; however, the two groups performed similarly on the work questions on a test given one week later. Additionally, both sequences of experimentation (physical-virtual and virtual-physical) equally supported students' understanding of all of the concepts. These results suggest that both the concept learned and the stability of learning gains should continue to be explored to improve educators' ability to select the best learning experience for a given topic.

DOI: [10.1103/PhysRevSTPER.8.010113](https://doi.org/10.1103/PhysRevSTPER.8.010113)

PACS numbers: 01.40.Fk, 01.40.gb, 01.50.Lc

### I. INTRODUCTION

The constructivist philosophies of learning, which span a diverse set of views ranging from Piaget [1] to Vygotsky [2], as well as many others, converge on the assertion that learners construct their own understanding. Some of these views are now established consensus in the physics education research community. Until recently, a conventional interpretation of these views has suggested that students learn physics best from hands-on experience. In physics laboratories, this has often translated to students performing experiments with physical equipment. However, contemporary research efforts have found little evidence to support this interpretation. Triona and Klahr [3] have pointed out that while constructivist theory suggests students must be actively involved in the process of learning, active involvement does not require physical manipulation.

Recently, researchers have reported finding that comparisons between students who perform similar experiments with physical equipment or computer simulations

indicate that the students who perform the virtual experiment learn the same as [4,5] or more than [6,7] students who perform the physical experiment. Zacharia and Olympiou [8] summarize the viewpoints of advocates for both physical and virtual experimentation. Advocates for physical experimentation cite evidence that physicality supports memory and learning and may ground their arguments in the working memory theoretical framework or cognitive load theory. These theories suggest that if physically touching and manipulating equipment represents a new channel for information to travel, then learning through physical manipulation will result in reduced cognitive load for the same amount of information (see Ref. [8] for an explanation of how these theories can explain learning through physical manipulation). However, the authors point out that no research has proven “touch” enhances students' learning; rather, the findings from studies comparing physical manipulation to more passive modes of instruction (involving lack of manipulation rather than lack of physicality, e.g., Ref. [9]) may have been inappropriately extended to the comparison of physical and virtual manipulation. On the other hand, in addition to summarizing many practical benefits of virtual experimentation (e.g., portability, safety, cost efficiency), Zacharia and Olympiou [8] suggest virtual experiments may be a better introduction than physical experiments

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

for particularly abstract concepts. The authors suggest that the fundamental difference between physical and virtual experimentation is physicality, and advocates of virtual experimentation suggest the sensory input provided by physicality may not be important for learning.

Zacharia and Olympiou [8] designed a study to explore whether physicality is important in university-level physics learning, focused specifically on heat and temperature. Four experimental groups used either physical, virtual, physical then virtual, or virtual then physical equipment to perform experiments, while one control group observed the experiments performed with physical equipment in lecture. Curriculum and time on task were controlled between the conditions. Students' performance suggested that while the control group performed significantly worse than the experimental groups, all four experimental groups performed similarly. These results suggest that, in this case, physicality or the type of sensory input was not important. However, the authors caution that more research is needed to explore whether their findings extend to other contexts and other student populations. For example, physicality could be unimportant in this study since university-level students already have real-life experience with heating objects.

Reviewing the literature on the use of computers in physics experiments, we have built a "master list" of the reasons why computers can potentially be effective learning tools. Thornton and Sokoloff [10], who successfully used microcomputer-based labs in a kinematics curriculum, suggested five important characteristics: students focused on the physical world, immediate feedback was available, collaboration was encouraged, tools reduced drudgery, and students moved from the specific and familiar to the more general and abstract. Redish, Saul, and Steinberg [11], who successfully used microcomputer-based labs in mechanics, agreed and added to the above list the conjecture that students were actively involved in exploring and constructing their own understanding. Finkelstein *et al.* [6] successfully used a simulation to replace a physical circuits lab. They noted that the simulation was successful because it made visible models that were useful for forming concepts and constrained students in productive ways.

## II. THEORETICAL PERSPECTIVE

A theoretical framework that encapsulates many of the issues highlighted by the research cited above is the model of learning as dynamic transfer, described by Schwartz, Varma, and Martin [12]. Dynamic transfer involves the application of component competencies in an environment to yield new concepts. An environment supporting dynamic transfer allows for distributed memory (i.e., a learner can off-load mental work on the environment), affords alternative interpretations and feedback (i.e., helps learners overcome initial misconceptions), offers candidate

structures by constraining and structuring actions (i.e., provides scaffolding), and provides a focal point for coordination (i.e., helps a learner to bring together different pockets of knowledge). It is beyond the scope of this paper to discuss the framework in its full detail. Rather we focus on how the framework is relevant to this specific study.

Table I demonstrates an alignment of properties of successful computer simulations with the characteristics of environments that support dynamic transfer as well as how the simulation used in our study could support dynamic transfer in these ways. We find that the characteristics for successful computer use overlap significantly with the characteristics of an environment supportive of dynamic transfer. The remaining three properties for successful computer use are more general views of learning.

While the dynamic transfer framework [12] provides a useful lens with which to examine the attributes of virtual and physical manipulatives, an alternative perspective is provided by sociocultural theory. Otero [13,14] describes learning in terms of sociocultural theory and mediated action, where students' intuitive ideas and canonical concepts may develop at the same time and mediate the development of each other. She explains, "Mediated action may be defined broadly as action in which individuals make opportunistic use, consciously or unconsciously, of features of the environment to transform symbols into meanings" ([14], p. 3). Otero applied this theory of learning to students' experiences with physical and virtual experiments designed to support the development of a model of charge transfer, and found that the simulation was more successful at mediating the scientific model of charge transfer than experiments with physical equipment. She described several ways that the simulation successfully mediated the development of the canonical model. Otero proposed that the simulation provided an image that allowed the student to make meaning out of the scientific words used by the teacher, whereas the physical experiment relied on student-generated images, which in the case discussed actually inhibited the development of meaning. Additionally, the simulation helped the student to become more aware of the model she was using and to evaluate that model. We find that these advantages of the simulation are similar to traits supporting dynamic transfer described in Table I, such as offering feedback, constraining actions, and facilitating coordination of knowledge. Thus, an environment that supports dynamic transfer may also be a successful mediating artifact, and simulations may be better than physical experiments at meeting the requirements of each.

In examining the large body of literature on studies comparing physical and virtual manipulatives as well as reflecting on the various theoretical perspectives that shed light on the affordances and constraints of each manipulative, we find that no clear consensus emerges regarding which of the two manipulatives better supports learning.

TABLE I. Alignment of characteristics for successful computer use and dynamic transfer with the features of the simulation used in this study.

|   | Theme 1  | Theme 2  | Theme 3  | Theme 4   | Unaligned  |
|---|--|--|--|---|--|
| Characteristics of environment for dynamic transfer | Offers alternative interpretations and feedback                              | Allows for distributed memory                            | Provides a focal point for coordination of different knowledge pockets   | Offers candidate structures by constraining and structuring actions |  |
| General properties of successful computer use       | Immediate feedback is available  | Powerful tools reduce drudgery                           | Understand the specific and familiar before moving to the more general and abstract<br>Useful models for forming concepts are made visible | Students are constrained in productive ways                         | Students are actively engaged in exploring and constructing their own understanding<br>Focus on the physical world |
| Features of simulation used in this study           | Students can immediately see the values of abstract quantities, such as work | Allows students to backtrack by quickly rerunning trials | Bar graphs help students to consider several variables at once   | Bar graphs provide structure for comparison                         | Collaboration is encouraged  |

Even within the same concept (e.g., circuits), some studies suggest that virtual experimentation is more successful than physical experimentation (e.g., [6]) while others suggest the two provide equal support and that other factors, such as the type of instruction, are more important than the physicality of the equipment (e.g., [15]). Our study also explores learning using both physical and virtual manipulatives, but we go beyond the either-or experimental paradigm. Rather than just compare the two manipulatives in terms of learning, we also investigate learning with the two manipulatives used together in different sequences.

### III. GOALS OF THE STUDY

Our study aims to move towards the goal of exploring how physical and virtual experimentation support students' learning in physics, in an attempt to help educators and curriculum designers predict when physical or virtual experimentation would be more appropriate, and how they may be used together to promote learning. In the present study, we explore how the particular concept studied and timing of the post-test affect the comparison between performance for students who used physical and virtual manipulatives in different sequences to learn about physics concepts related to pulleys. Pulleys offer an interesting context in which to explore these research questions. Students likely have little first-hand experience constructing pulley systems. (Anecdotally, students were often surprised about the behavior of the pulley systems.) Also, students can be guided to explore a variety of concepts related to pulley systems, such as force and work, which may be supported better by the physical or virtual manipulatives due to their varying level of abstraction.

Students were guided to use either physical equipment or a computer simulation to explore how different pulley systems affected the force and work needed to lift an object and the mechanical advantage of the system. Students completed a multiple-choice conceptual test before performing either experiment, immediately after performing the physical or virtual experiment, one week later, and immediately after performing the same experiment with the second type of equipment. Our research questions are as follows:

- (1) When students use either physical or virtual manipulatives to perform the experiments with pulleys: (a) Does one manipulative better support students' understanding of the science concepts related to pulleys than the other? (b) Does the comparative benefit of the physical and virtual manipulatives depend on the concept studied? (c) Does the benefit of the physical and virtual manipulatives, as measured by post-test performance, depend on the timing of the post-test?
- (2) When students perform experiments with both physical and virtual manipulatives: (a) Does one sequence (physical-virtual or virtual-physical)

better support students' understanding of the science concepts related to pulleys than the other? (b) Does the comparative benefit of the sequences depend on the concept studied?

One might anticipate that students would develop a better understanding of how force was related to the pulley system from the physical experiment where they actually feel the force needed to lift the object with each system. On the other hand, one might expect students to develop a better understanding of work from the virtual experiment where their data is not affected by frictional effects, difficulties with physical equipment (i.e., lifting a block at constant velocity), and measurement errors (i.e., mistakes using measuring instruments and systematic errors). In addition, the simulation is able to provide other types of potential scaffolding, such as visual representations of data, which may focus students' attention in more productive ways than the physical experiment. In fact, our results show both concept and post-test timing dependence.

## IV. EXPERIMENTAL DESIGN

### A. Participants and environment

The study was performed in Spring 2010 at a large research university in the Midwest. In the present study, our focus is on how interaction with physical and virtual manipulatives in a real physics classroom affected students' conceptual understanding. Thus, the curriculum and assessments in this study are directed at promoting

and measuring students' conceptual understanding at several stages of this interaction. The students were enrolled in a conceptual-based introductory physics laboratory for nonscience majors. The students were concurrently taking or had previously taken a conceptual-based introductory physics lecture, but there was no attempt to synchronize lecture and laboratory, nor were the lecture and laboratory restricted to cover the same topics. Typically, the laboratory proceeds as a separate 1 credit-hour course, and students attend one 2-hour laboratory session per week. The laboratory is taught by undergraduate or graduate teaching assistants.

The materials used in this study were developed for the CoMPASS (Concept Mapped Project-based Activity Scaffolding System) Simple Machines curriculum ([16,17]). The CoMPASS curriculum combines project-based learning with concept exploration in an online hypertext system (Fig. 1) and hands-on experimentation with both traditional physical equipment and computer simulations. In the full CoMPASS Simple Machines curriculum, students study science concepts such as the force-distance trade-off and conservation of energy in the context of several machines; throughout the curriculum, they are motivated by a final project where they are challenged to combine simple machines in such a way to lift a can with the least amount of input force. In this study, students completed the pulley unit, which asks them to conceive of the best pulley system to use to lift a heavy pool table into a van. The unit aims to help students build their

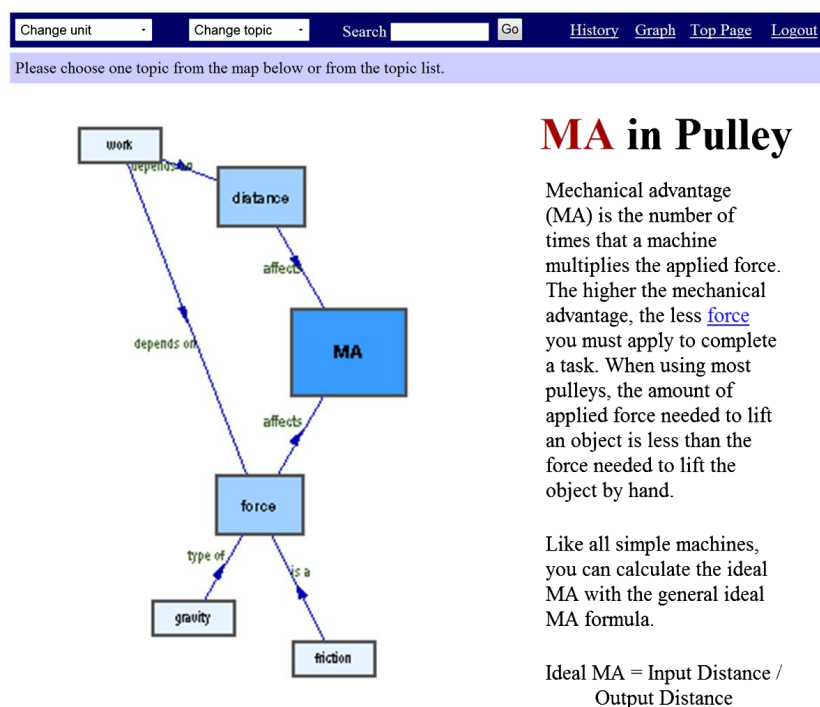


FIG. 1 (color online). The CoMPASS hypertext system. Students used the system to explore the physics concepts related to pulleys.



understanding of force, work, and mechanical advantage in the context of pulleys; specifically, students explore how changes to the pulley system affect the force and work needed to lift a given load a given distance. The physical equipment (Fig. 2) consisted of small pulleys, string, a stand, a spring scale, a meter stick, and a small mass. Students strung the pulleys to create single fixed, single movable, single compound, and double compound pulley systems, and measured the force needed to lift the load (at constant speed) and the distance the string had to be pulled to lift the object 10 cm. In the simulation (Fig. 3), students created the same pulley systems by selecting the pulley system with their mouse and setting the distance the object was to be lifted. Students supplied an input force by pulling up on the slider until the object began to move; the simulation capped the input force at the minimum force required to lift the load. Students could then view mea-



FIG. 2 (color online). The physical manipulatives for the pulley experiment included pulleys, string, a spring scale, a meter stick, and a mass.

surements of the input force, work needed, change in the object's potential energy, and mechanical advantage of the pulley system. Worksheets [17] guided students to record data about force, work, change in the load's potential energy, and mechanical advantage for each pulley system. Then students were asked to respond to analysis questions (e.g., "For which pulley system was the distance pulled the greatest?").

## B. Procedure

The students completed two 2-hour activities using pulleys, with a week-long separation between the lab meetings. In the first activity (week 1), students used the CoMPASS online hypertext system to explore the science concepts related to pulleys. Then, students either used physical or virtual manipulatives to perform experiments to explore how force, work, potential energy, and mechanical advantage were related to changes in the pulley system. A week later in the second activity (week 2), the students performed similar experiments with the other type of manipulative; students who had used the physical manipulatives now used the virtual (PV sequence), and students who had used the virtual manipulatives now used the physical (VP sequence). The same curriculum and method of instruction were used in both sequences so that differences in performance on questions related to the concepts studied can be attributed to differences in the manipulatives' affordances and ordering [4]. The data presented in this paper were collected in Spring 2010, and during that semester there were four sections of the laboratory. In each section, half of the students used the equipment in the PV sequence ( $N = 58$ ) and half in the VP sequence ( $N = 63$ ). The sequence experienced by each student was determined by which lab table the student chose to sit at before the beginning of the first activity. The students worked in pairs.

The students individually took the same 20-item multiple-choice test several times throughout the study, which allows us to explore how their knowledge of the science concepts related to pulleys evolved. Students completed the test at the beginning of week 1, before using any materials, and at the end of week 1, after using the hypertext system and one manipulative type. Students took the test again at the beginning of week 2, before using any additional materials, and at the end of week 2, after using the second manipulative type. The experimental design is summarized in Fig. 4.

Students' scores on the test were analyzed using a repeated measures model, both for the total score and the concept subscores for questions related to force, work and/or energy, and mechanical advantage. While retesting effects can affect the reliability of an instrument, we believe those effects can be ignored in the present analysis as our main focus is on comparisons between two groups who have been exposed to the test the same number of times.

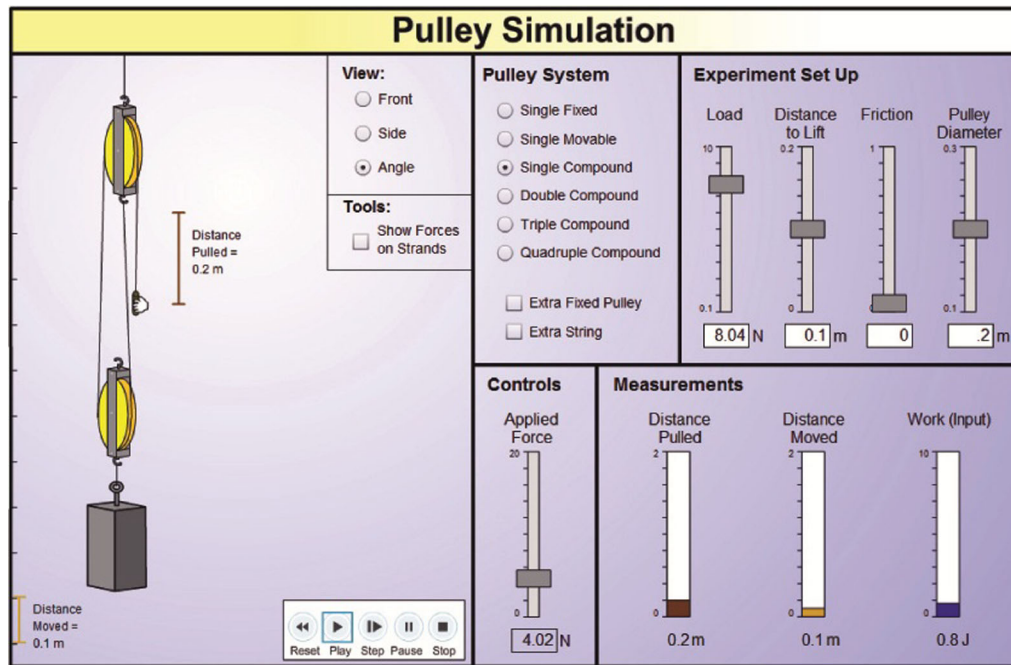


FIG. 3 (color online). The virtual manipulative for the pulley experiments was a computer simulation that allowed students to select a pulley system, set the load and distance lifted, apply an input force, and observe the distance the string was pulled and the work required to lift the load.

### C. Assessment

The assessment instrument [17] used in this study consisted of 20 multiple-choice questions, eight questions about force, nine questions about work and/or potential energy, and three questions about mechanical advantage. Half of the questions used diagrams of pulleys, and half of the questions were purely verbal. The majority of questions (16) explicitly stated that “friction” or “the effects of friction” could be ignored. Two questions asked about cases where friction could not be ignored. Additionally, two questions about the change in potential energy did not discuss friction. As measured by Cronbach’s alpha, the test is acceptably reliable for all administrations after the pre-test:  $\alpha_{\text{pre}} = 0.65$ ,  $\alpha_{\text{mid 1}} = 0.76$ ,  $\alpha_{\text{mid 2}} = 0.72$ , and  $\alpha_{\text{post}} = 0.77$ . Equivalent-forms reliability of the pulley test was previously explored in an interview study, and

students were found to give consistent responses to written pulley test questions and similar verbal questions that were restated with a different context [18]. The test questions were reviewed by middle school teachers, physics graduate students, and physics postdoctoral research associates to establish content validity. Reviewers were in agreement that the questions adequately sampled the domain of introductory physics concepts related to pulleys.

### V. RESULTS

The pre-test, mid-test, post-test design allows us to compare how students’ performance was supported by interaction with individual manipulatives (physical or virtual) as well as the sequence of manipulatives (physical-virtual or virtual-physical). The results will be discussed below by concept (total score, force subscore, work-energy

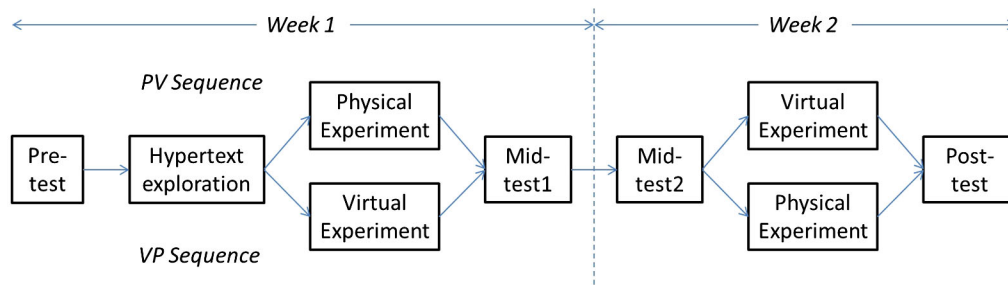


FIG. 4 (color online). The experimental design involved two sequences. In the PV sequence students performed the physical experiment followed by the virtual experiment. In the VP sequence students performed the virtual experiment followed by the physical experiment. All tests were identical.

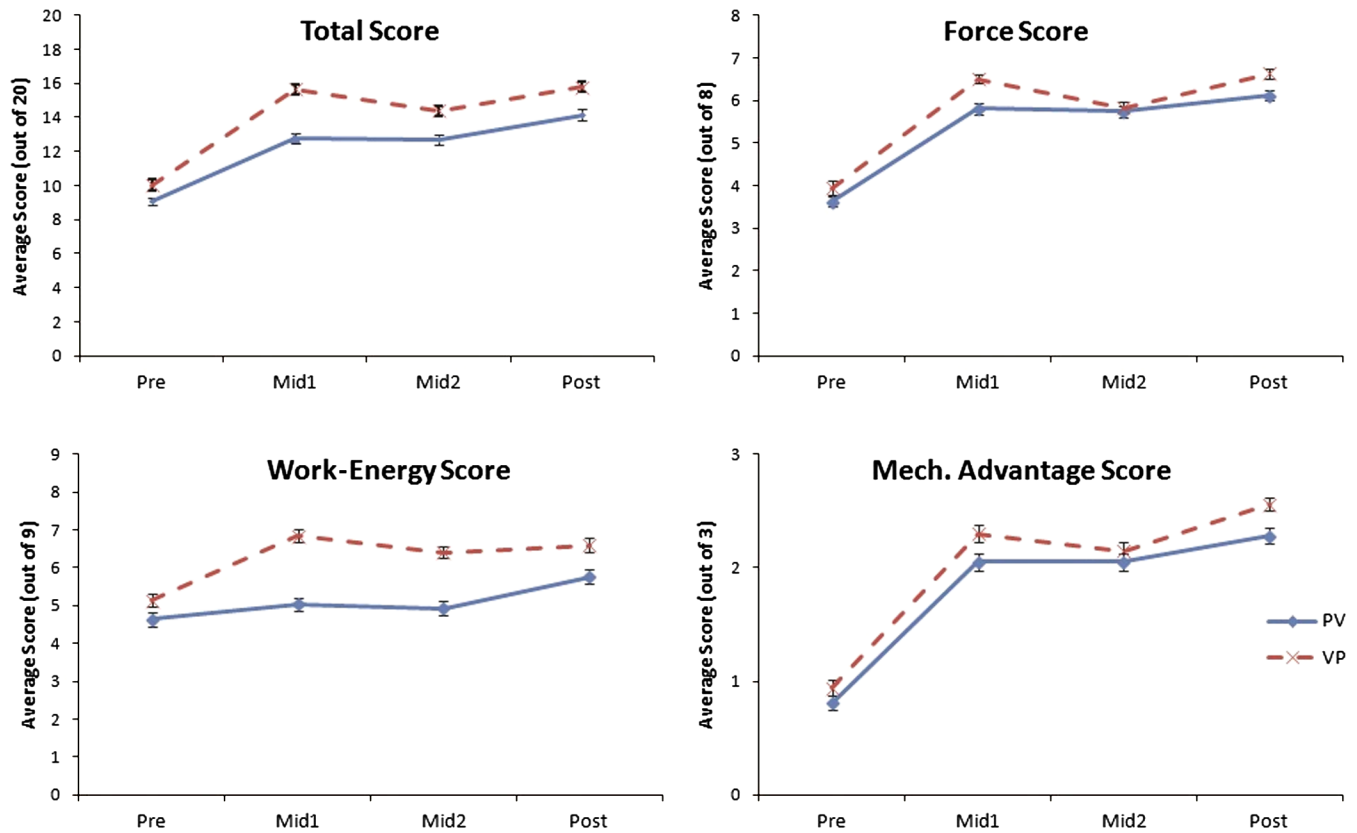


FIG. 5 (color online). Mean score by concept. Error bars represent standard error.

subscore, and mechanical advantage subscore). The average score for each test by concept is displayed in Fig. 5. A repeated measures analysis of variance test was used to compare students' scores across the tests. Mauchly's test was used to determine if sphericity could be assumed, and the Greenhouse-Geisser correction was used in cases where sphericity could not be assumed [19]. We examined both the main effect of score and the interaction of manipulative with score. When an effect was significant at the level  $p < 0.05$ , planned contrasts were examined to determine between which levels of test the effect was significant.

#### A. Total score

The total score changed significantly across the tests,  $F(2.7, 317.4) = 200.1$ ,  $p < 0.001$ . Planned contrasts

revealed the scores changed significantly between each test, as shown in Table II. Combining these results with the evolution of the average total score displayed in Fig. 5, we find that students' scores increased significantly from pre-test to mid-test1 while they used the first manipulative type, decreased significantly from mid-test1 to mid-test2 in the week between the labs, and increased significantly again from mid-test2 to post-test.

The interaction of total score with manipulative was also significant. The planned contrasts revealed the interaction was only significant from pre-test to mid-test1 and mid-test1 to mid-test2. Thus, while students who used the virtual manipulative first learned significantly more than students who used the physical manipulative in week 1, they also forgot significantly more between mid-test1 and mid-test2. Both groups' scores increased similarly after

TABLE II. Repeated measure analysis of variance: total score.

| Effect       | Main: $F(2.7, 317.4) = 200.1$ , $p < 0.001$ |        |      | Interaction: $F(2.67, 317.4) = 4.7$ , $p = 0.005$ |       |      |
|--------------|---|--------|------|---|-------|------|
|              | $F(1, 119)$                                 | $p$    | $r$  | $F(1, 119)$                                       | $P$   | $r$  |
| Pre vs mid1  | 312.7                                       | <0.001 | 0.85 | 11.4  | 0.001 | 0.30 |
| Pre vs mid2  | 245.3                                       | <0.001 | 0.82 | 1.5   | 0.225 | 0.11 |
| Mid1 vs mid2 | 13.4  | <0.001 | 0.32 | 7.7   | 0.006 | 0.25 |
| Mid2 vs post | 44.1  | <0.001 | 0.52 | 0.0   | 0.936 | 0.07 |
| Pre vs post  | 367.7                                       | <0.001 | 0.87 | 2.0   | 0.160 | 0.13 |

TABLE III. Repeated measure analysis of variance: force subscore. \* indicates that the planned contrasts could not be examined.

| Effect       | Main: $F(2.6, 310.4) = 154.1, p < 0.001$ |        |      | Interaction: $F(2.6, 310.4) = 1.9, p = 0.131$ |     |     |
|--------------|--|--------|------|---|-----|-----|
|              | $F(1, 119)$                              | $p$    | $r$  | $F(1, 119)$                                   | $p$ | $r$ |
| Pre vs mid1  | 248.1                                    | <0.001 | 0.82 | *   | *   | *   |
| Pre vs mid2  | 180.3                                    | <0.001 | 0.78 | *   | *   | *   |
| Mid1 vs mid2 | 9.3                                      | 0.003  | 0.27 | *   | *   | *   |
| Mid2 vs post | 32.3                                     | <0.001 | 0.46 | *   | *   | *   |
| Pre vs post  | 280.9                                    | <0.001 | 0.84 | *   | *   | *   |

using the second type of manipulative. Overall, the two sequences seem to be equivalent since students' total scores changed similarly from pre-test to post-test in both sequences.

### B. Force subscore

The force subscore changed significantly across the tests,  $F(2.6, 310.4) = 154.1, p < 0.001$ . The planned contrasts revealed the force subscore changed significantly between each test, as shown in Table III. The average force subscore followed the same trend as the total score, increasing significantly after the first activity, decreasing significantly during the week off, and increasing significantly again after the second activity. The interaction of force subscore with manipulative was not significant,  $F(2.6, 310.4) = 1.9, p = 0.131$ . Thus, students' scores did not change differently based on the manipulative used.

### C. Work-energy subscore

The work-energy subscore changed significantly across the tests,  $F(2.7, 322.8) = 26.4, p < 0.001$ . The planned

contrasts reveal that the change was significant between each test, as shown in Table IV, and the evolution of the average work-energy subscore shown in Fig. 5 shows that the scores followed the same trend as the total score and force subscore.

The interaction of work-energy subscore with manipulative was also significant,  $F(2.7, 322.8) = 7.4, p < 0.001$ . The interaction was significant between the pre-test and mid-test1 as well as pre-test and mid-test2; Fig. 5 indicates that the virtual manipulative appears to better support students' understanding of work and energy in the context of pulleys than the physical manipulative both immediately and one week later. The interaction was not significant between mid-test1 and mid-test2, which indicates that students retained the same amount of information during the week off, regardless of the manipulative used in week 1. The interaction effect is again significant between mid-test2 and post-test. Again, the virtual manipulative appears to offer better support for students' understanding of work and energy since now the scores in the PV sequence are improved more than the scores in the VP sequence (although the VP post-test score is higher than

TABLE IV. Repeated measure analysis of variance: work-energy subscore.

| Effect       | Main: $F(2.7, 322.8) = 26.4, p < 0.001$ |        |      | Interaction: $F(2.7, 322.8) = 7.4, p < 0.001$ |        |      |
|--------------|---|--------|------|---|--------|------|
|              | $F(1, 119)$                             | $p$    | $r$  | $F(1, 119)$                                   | $p$    | $r$  |
| Pre vs mid1  | 38.2                                    | <0.001 | 0.49 | 14.7  | <0.001 | 0.33 |
| Pre vs mid2  | 22.0                                    | <0.001 | 0.40 | 8.6   | 0.004  | 0.26 |
| Mid1 vs mid2 | 4.1                                     | 0.045  | 0.18 | 1.5   | 0.219  | 0.11 |
| Mid2 vs post | 16.2                                    | <0.001 | 0.35 | 6.3   | 0.013  | 0.22 |
| Pre vs post  | 58.7                                    | <0.001 | 0.57 | 1.0   | 0.315  | 0.09 |

TABLE V. Repeated measure analysis of variance: mechanical advantage score. \* indicates that the planned contrasts could not be examined.

| Effect       | Main: $F(2.7, 320.2) = 172.0, p < 0.001$ |        |      | Interaction: $F(2.7, 320.2) = 0.7, p = 0.568$ |     |     |
|--------------|--|--------|------|---|-----|-----|
|              | $F(1, 119)$                              | $p$    | $r$  | $F(1, 119)$                                   | $p$ | $r$ |
| Pre vs mid1  | 223.5                                    | <0.001 | 0.81 | *   | *   | *   |
| Pre vs mid2  | 233.2                                    | <0.001 | 0.81 | *   | *   | *   |
| Mid1 vs mid2 | 1.2                                      | 0.285  | 0.11 | *   | *   | *   |
| Mid2 vs post | 26.1                                     | <0.001 | 0.42 | *   | *   | *   |
| Pre vs post  | 382.9                                    | <0.001 | 0.87 | *   | *   | *   |



the PV post-test score, the change between mid-test2 and post-test is greater in the PV sequence). The interaction is not significant from pre-test to post-test, which indicates that overall the two sequences equally support students' understanding of work and energy.

#### D. Mechanical advantage subscore

The mechanical advantage subscore changed significantly across the tests,  $F(2.7, 320.2) = 172.0$ ,  $p < 0.001$ . The planned contrasts revealed the change was significant between all tests except mid-test1 and mid-test2, as shown in Table V. Comparing these results with Fig. 5, students' mechanical advantage subscores appear to improve significantly from interaction with the manipulative in both week 1 and week 2, and do not decrease significantly in the off week. The interaction of mechanical advantage subscore with manipulative was not significant,  $F(2.7, 320.2) = 0.7$ ,  $p = 0.568$ . Thus, both manipulatives appear to support students' understanding of mechanical advantage equally.

## VI. DISCUSSION

Below, we discuss how the results address each research question. First, we discuss our findings about the comparison of the physical and virtual manipulatives. Next, we discuss the overall physical-virtual and virtual-physical sequences.

### A. Physical manipulative versus virtual manipulative

When we compare how the individual manipulatives supported students' understanding of various concepts throughout the two week intervention, we observe differences in the comparison based on both the timing of the post-test and the concept studied.

#### 1. Timing of post-test

The experimental design allows us to isolate the effect of each manipulative by looking at the pre-test, mid-test1, and mid-test2 comparisons for the interaction of test and manipulative. In this case, we can consider mid-test1 as an immediate post-test and mid-test2 as a delayed post-test. When we analyzed the total score, the pre-test and mid-test1 comparison indicated that the virtual manipulative better supported students' understanding than the physical manipulative. However, the pre-test and mid-test2 comparison indicated that the physical and virtual manipulatives offered equal support. Thus, the timing of the post-test appears to affect the results [20,21], suggesting that students may have initially learned more from their experiments with the virtual manipulative, but that material may have had low salience and was forgotten in the time between mid-test1 and mid-test2. This result contrasts those of other researchers who have found that students who learned with a virtual manipulative continued to

outperform students who learned with a physical manipulative on delayed measurements of understanding [6].

#### 2. Effect of concept

In addition to analyzing results for total score, we also looked at subsets of questions that pertained to force, work and/or energy, and mechanical advantage. The physical and virtual manipulatives appear to differently support these concepts. For the force subscore and mechanical advantage subscore, both manipulatives appear to equally support students' understanding. On the other hand, the virtual manipulative appears to offer better support for the work-energy subscore than does the physical manipulative. In both the PV and VP sequences, students' scores made their largest increase after their experience with the virtual manipulative.

It is perhaps surprising that the virtual manipulative is equally effective as the physical manipulative at supporting students' understanding of force since force is something you can feel in the physical experiment. Zacharia, Olympiou, and Papaevripidou [7] propose that physical manipulation is likely to be more important in domains that involve tactile sensation. In other words, one might suspect that an affordance of the physical manipulatives is that they allow the user to feel the difference in input force needed for different pulley systems. In the domain of pulleys, our results suggest that haptic feedback is not important for students' understanding, even about force, which is a very "tactile" concept. It is important to note that our students are young adults and are typically familiar with working with technology. In our collaborators' work with younger students, we have found that some students do benefit more from the physical experiment than the virtual experiment or from completing the physical experiment before the virtual experiment [22]. Thus, students' familiarity with the concept and technology appears to mediate their ability to learn from the virtual manipulative. Another potential reason that the physical and virtual manipulatives equally supported students' understanding of force is the reduced time between trials with the virtual manipulative. It is possible that the ability to perform trials more rapidly with the virtual manipulative balances out any tactile advantage present with the physical manipulative for these university-level students. It is also possible that students approach the simulation differently than they approach the physical experiment, but this cannot be assessed from the present study. Some exploration of how students approach physical and virtual experiments is done in Ref. [23].

It may be less surprising that the virtual manipulative provided better support for students' understanding of work and energy since the simulation was able to create idealized frictional effects. However, the effects of friction are quite small in the physical experiment and students were guided to make comparisons between the work

required to lift the object and the change in the object's potential energy. These findings support the proposal by Zacharia and Olympiou [8] that virtual manipulatives may provide a better introduction to abstract concepts. The opportunity to control factors, such as friction, that may make the concept more difficult to observe is an advantage of virtual manipulatives that is hard (or impossible) to recreate in a physical environment. For example, in the present study students used low-friction pulleys and observed only small differences in the work needed to lift the object with different pulley setups. However, it is impossible to completely eliminate friction in the physical experiment, whereas the pulley can be modeled as frictionless in the simulation. Additionally, students in the VP sequence reached their maximum average work-energy subscore after using only the virtual manipulative and did not improve more after using the physical manipulative, suggesting that the physical manipulative did not provide any insight that the students could not gain from the virtual manipulative on the type of conceptual understanding assessed in this study. It is possible that students may have developed other types of skills from the physical experiment, such as troubleshooting, that should be explored in future studies.

We find that the mechanical advantage subscore follows the same trend as the force subscore (i.e., physical and virtual manipulatives provide equal support) rather than the trend observed for the work-energy subscore (i.e., the virtual manipulative provides more support than the physical manipulative). This is perhaps surprising, as an argument can be made that mechanical advantage is an abstract concept or that it is physically observable. Mechanical advantage is defined as the ratio of the force needed to lift the load to the weight of the load, so it is closely related to force. Also, students can feel the mechanical advantage when it is “easier” or “harder” to lift the load (although they may confuse this sensation with work). Therefore, it is not that surprising that the mechanical advantage and force subscores follow the same trend. However, like work, mechanical advantage is a more abstract concept, so one may have expected to observe similar trends in the mechanical advantage and work-energy subscores. More research is needed to study how various contexts and physics concepts influence the physical versus virtual comparison so that researchers and educators can more accurately make such predictions.

### **B. Physical-virtual sequence versus virtual-physical sequence**

In comparing the overall manipulative sequence (physical-virtual or virtual-physical), we find that both sequences equally support students' performance from the pre-test to the post-test for the total score and each concept subscore. This finding is consistent with prior research with university-level students [8], but not

consistent with our collaborator's research with younger students [22], who benefited more from the PV sequence. While there is no statistically significant difference between the PV and VP sequences, the average VP score is consistently higher than the average PV score, including at the final post-test. While not statistically significant, this difference may be educationally significant and should be explored further in future work.

## **VII. CONCLUSIONS AND IMPLICATIONS**

Our results suggest that which manipulative, physical or virtual, is more appropriate for a given experiment depends on the concept to be learned. While students' understanding of force and mechanical advantage in the domain of pulleys was supported equally by physical and virtual manipulatives, their understanding of work was supported more by the virtual manipulative. When we analyzed students' understanding across these three concepts, we found that students' performance on an immediate post-test was significantly higher after using the virtual manipulative rather than the physical manipulative, but statistically similar on a delayed post-test one week later. Thus, our results for individual manipulative varied both by concept and by timing of the post-test. However, our results fail to support the belief that haptic feedback is essential even for concepts closely related to “touch,” such as force, at least for university-level students. Our results for manipulative sequence, whether students performed a physical experiment followed by a virtual experiment or a virtual experiment followed by a physical experiment, showed no difference between sequences for any concept.

These results support our previous findings from three prior implementations of the CoMPASS pulley curriculum with nonscience major undergraduate students [23]. Across three previous studies, we observed that students who used a virtual manipulative scored higher on multiple-choice questions related to work than students who used a physical manipulative. Students scored similarly on force and mechanical advantage questions regardless of the manipulative used. Also, the sequence in which the manipulatives were used did not affect students' scores. This result is also consistent with prior research on university-level students' learning about heat and temperature [8], but not consistent with prior research on middle school-level students' learning about the same domain of pulleys [22].

It is possible that the physical and virtual manipulatives differently supported students' understanding of work for several reasons. As previously stated, the computer simulation may have focused students' attention in important ways by providing a frictionless environment and bar graph representations. These features of the virtual environment may better support dynamic transfer than the physical environment. However, more in-depth exploration of how students engage with the manipulatives is necessary to support this hypothesis, and is not possible with the present

study design. To explore this issue, a qualitative study examining how students interact with each of the manipulatives as they construct their knowledge would be necessary. Otero [13,14] has looked at such interactions through the lens of mediated action, and the ways in which the virtual manipulative mediated students' learning in her study appear to be consistent with the ways an environment can support dynamic transfer. Further study is needed to explore how these theories can be used to describe students' learning process while interacting with different manipulatives and different manipulative sequences.

The present study has several limitations. It is important to note that our study has focused solely on students' conceptual understanding, and not their ability to conduct future scientific explorations or to physically use pulleys. One must consider the type of knowledge students are to gain when selecting which type of manipulative to use. (See Refs. [6,24] for studies that involve additional forms of knowledge.)

These results have several implications for education. We have found that short-term conceptual performance for some concepts is significantly affected by manipulative. Therefore, educators should consider conceptual learning goals when choosing between pulley manipulatives. On the other hand, the sequence of manipulatives did not significantly affect students' performance on the conceptual test.

Therefore, educators should consider other potential advantages when planning the order of physical and virtual pulley manipulatives. More studies are needed to investigate whether these findings extend beyond the domain of pulleys.

Based on the concept dependence in our results, we echo other researchers (e.g., Ref. [8]) in suggesting continued research in additional contexts to enhance our ability to predict the domains in which physical or virtual manipulatives are more appropriate. Based on the post-test timing dependence, we suggest that future research measure students' understanding at multiple endpoints (i.e., both an immediate post-test and a delayed post-test) to add to our knowledge of the stability of any benefit derived from virtual manipulation over physical manipulation. The differences found between university-level and middle school-level students' learning in the same domain suggest that more research is needed to explore how prior experiences affect students' ability to learn from virtual manipulatives.

## ACKNOWLEDGMENTS

This work is supported in part by U.S. Department of Education, Institute of Education Sciences Grant No. R305A080507 and the U.S. National Science Foundation GK-12 Program, Grant No. DGE-0841414.

- 
- [1] J. Piaget, Development and learning, *J. Res. Sci. Teach.* **2**, 176 (1964).
  - [2] L. S. Vygotsky, *Mind in Society: Development of Psychological Processes* (Harvard University Press, Cambridge, MA, 1978).
  - [3] L. M. Triona and D. Klahr, Point and click or drag and heft: Comparing the influence of physics and virtual instructional materials on elementary students' ability to design experiments, *Cogn. Instr.* **21**, 149 (2003).
  - [4] D. Klahr, L. M. Triona, and C. Williams, Hands on what? The relative effectiveness of physical versus virtual materials in an engineering design project by middle school children, *J. Res. Sci. Teach.* **44**, 183 (2007).
  - [5] Z. C. Zacharia and C. P. Constantinou, 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, *Am. J. Phys.* **76**, 425 (2008).
  - [6] N. D. Finkelstein, W. K. Adams, C. J. Keller, P. B. Kohl, K. K. Perkins, N. S. Podolefsky, S. Reid, and R. LeMaster, When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment, *Phys. Rev. ST Phys. Educ. Res.* **1**, 010103 (2005).
  - [7] Z. C. Zacharia, G. Olympiou, and M. Papaevripidou, Effects of experimenting with physical and virtual manipulatives on students' conceptual understanding in heat and temperature, *J. Res. Sci. Teach.* **45**, 1021 (2008).
  - [8] Z. C. Zacharia and G. Olympiou, Physical versus virtual manipulative experimentation in physics learning, *Learn. Instr.* **21**, 317 (2011).
  - [9] G. E. Glasson, The effects of hands-on and teacher demonstration laboratory methods on science achievement in relation to reasoning ability and prior knowledge, *J. Res. Sci. Teach.* **26**, 121 (1989).
  - [10] R. K. Thornton and D. R. Sokoloff, Learning motion concepts using real-time microcomputer-based laboratory tools, *Am. J. Phys.* **58**, 858 (1990).
  - [11] E. F. Redish, J. M. Saul, and R. N. Steinberg, On the effectiveness of active-engagement microcomputer-based laboratories, *Am. J. Phys.* **65**, 45 (1997).
  - [12] D. L. Schwartz, S. Varma, and L. Martin, Dynamic transfer and innovation, in *International Handbook of Research on Conceptual Change* (Routledge, New York, 2008), p. 497.
  - [13] V. Otero, in *Cognitive processes and the learning of physics, Part I: The evolution of knowledge from a Vygotskian perspective*, Proceedings of the International School of Physics "Enrico Fermi," Course CLVI, edited by M. Vicentini and E. F. Redish (IOS Press, Amsterdam, 2004), Vol. 156, p. 409 [[http://spot.colorado.edu/%7Eotero/OTERO\\_1\\_COGNITIVE\\_PROCESSES.pdf](http://spot.colorado.edu/%7Eotero/OTERO_1_COGNITIVE_PROCESSES.pdf)].

- [14] V. Otero, in *Cognitive processes and the learning of physics, Part II: Mediated action* (Ref. [13]), p. 447 [[http://spot.colorado.edu/%7Eotero/OTERO\\_2\\_MEDIATED\\_ACTION.pdf](http://spot.colorado.edu/%7Eotero/OTERO_2_MEDIATED_ACTION.pdf)].
- [15] T. Jaakkola, S. Nurmi, and K. Veermans, A comparison of students' conceptual understanding of electric circuits in simulation only and simulation-laboratory contexts, *J. Res. Sci. Teach.* **48**, 71 (2011).
- [16] S. Puntambekar, A. Stylianou, and J. Goldstein, Comparing classroom enactments of an inquiry curriculum: Lessons learned from two teachers, *J. Learn. Sci.* **16**, 81 (2007).
- [17] Materials are available at <http://web.phys.ksu.edu/CoMPASS-curriculum/>.
- [18] A. Rouinfar, J. J. Chini, A. Carmichael, S. Puntambekar, and N. S. Rebello, Investigating change and consistency in introductory college students' understanding about pulleys, in Proceedings of the Annual Meeting of the National Association for Research in Science Teaching, Philadelphia, 2010 (unpublished).
- [19] A. P. Field, *Discovering Statistics Using SPSS* (Sage Publications, Thousand Oaks, CA, 2005), p. 779.
- [20] L. Ding, N. W. Reay, A. Lee, and L. Bao, Effects of testing conditions on conceptual survey results, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010112 (2008).
- [21] E. Sayre and A. F. Heckler, Peaks and decays of student knowledge in an introductory E&M course, *Phys. Rev. ST Phys. Educ. Res.* **5**, 013101 (2009).
- [22] G. W. Smith and S. Puntambekar, Examining the combination of physical and virtual experiments in an inquiry science classroom, in Proceedings of the Conference on Computer Based Learning in Science, Warsaw, Poland, 2010 (unpublished).
- [23] J. J. Chini, Ph.D. thesis, Kansas State University, 2010.
- [24] E. E. Toth, D. Klahr, and Z. Chen, Bridging research and practice: A cognitively based classroom intervention for teaching experimentation skills to elementary school children, *Cogn. Instr.* **18**, 423 (2000).