

Student learning outcomes with hybrid computer simulations and hands-on labs

Sheila Saguear

*Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA and
Department of Astronomy, Boston University, 725 Commonwealth Avenue, Boston, MA, 02215, USA*

Emily Allen

*Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA and
Science Department, The Governor's Academy, 1 Elm St, Byfield, MA, 01922, USA*

Andrew Duffy and Manher Jariwala

*Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA and
Wheelock College of Education, Boston University, Two Silber Way, Boston, MA, 02215, USA*

Computer simulations for physics labs may be combined with hands-on lab equipment to boost student understanding and make labs more accessible. Hybrid labs of HTML5-based computer simulations and hands-on lab equipment for topics in mechanics were investigated in a large, algebra-based, studio physics course for life science students at a private, research-intensive institution. Computer simulations were combined with hands-on equipment and compared to traditional hands-on labs using an A/B testing protocol. Learning outcomes were measured for the specific topic of momentum conservation by comparing student scores on post-lab exercises, related quiz and exam questions, and a subset of questions on the Energy and Momentum Conceptual Survey (EMCS) administered before and after instruction for both groups. We find that students who completed a hands-on lab vs. a hybrid lab showed no difference in performance on momentum assessments.

I. Introduction

Laboratory instruction is an integral part of science classes, as it allows students to engage with physical phenomena and deepen their knowledge of course content [1]. Using virtual tools in physics labs may be imperative in online courses, and can encourage content mastery as part of in-person courses [2]. Interactive lab simulations allow students to easily explore physical concepts [3], encourage self-guided exploration [4], and may be more accessible, inexpensive, and less time-consuming during setup and tear-down [5, 6]. Alternatively, hands-on labs provide opportunities for students to interact directly with physical phenomena, learn how to use tools, and collect real data [1, 7], while simulation-based labs may present over-simplified and idealized contexts [8]. Lab simulations may therefore be most effective as a supplement to hands-on labs, especially for courses being adapted to a blended online/in-person learning environment. Labs that include both hands-on and simulated sections may combine the advantages of lower cost and reduced setup and tear-down time along with the added benefits of hands-on labs [9].

Previous studies have largely focused on student performance with simulated labs, and have found that simulated labs are at least as effective as hands-on labs in supporting conceptual learning objectives for labs [2, 10–15]. As a continuation of our previous study comparing simulated and hands-on labs [15], we applied a "hybrid" approach to simulation implementation to further investigate student learning outcomes. The "hybrid" lab studied in this work was a combination of simulated and hands-on lab activities for two-cart collisions.

Here we investigate the research question: Do hybrid labs or hands-on labs affect student understanding, content knowledge gain, and course performance? Specifically, this study adds to the literature concerning the impact of simulated two-cart collision labs on student performance on momentum conservation assessments.

II. Methods

A. Target Population

We conducted this study through an introductory, algebra-based mechanics class at a large, private, research university. This was a studio-style class geared towards life science majors and ran for 15 weeks in the fall semester. The class initially enrolled a total of 418 students over 5 studio sections, of whom 374 completed the course. Studio classes combine lecture, discussion, and lab sections, and are designed to encourage collaboration and discussion by seating students at round tables [16]. The class was composed of approximately 65% female students, 25% male students (10% who did not report their gender), with approximately 20% URM (under-represented minority).

Each class section had its own teaching staff, composed of one faculty instructor, two graduate student teaching fellows (TAs), and two undergraduate learning assistants (LAs). LAs are undergraduates who have previously taken an intro-

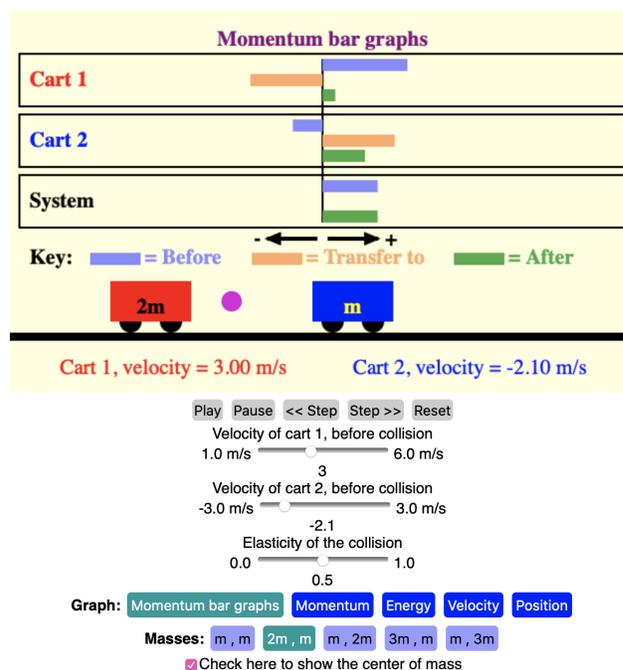


FIG. 1. Simulation portion of the hybrid momentum lab. Students are able to define the relative masses and initial velocities of the carts, and play the two-cart collision in real time. The hybrid lab consisted of three collisions using this simulation and three collisions using two physical carts.

ductory physics class and return to the class to assist in instruction [17]. LAs are available to help students understand course topics, in addition to graduate TAs and the faculty instructor. TAs and LAs circulated the studio classroom.

B. Simulation Development

The simulation used with the hybrid lab was created by one of the course faculty instructors. The simulation was written in HTML5 (Javascript) and featured an interactive animation of two moving carts colliding in one dimension. The simulation depicted momentum bar graphs of three different types of collisions: completely elastic, inelastic, and completely inelastic. It also depicted line graphs of the carts' energy vs. time, momentum vs. time, and position vs. time. Students were able to change settings of the two-cart collision, such as the carts' initial velocities, mass ratio, and collision elasticity. A screenshot of the momentum lab simulation and possible settings is shown in Figure 1. The hybrid lab was designed to be as similar in content to the hands-on lab as possible. All lab materials, including the simulation, are available on our website [18].

The simulation portion of the hybrid labs were reviewed by LAs prior to being used in the classroom. We paired LAs for this course and conducted think-aloud interviews as they worked through the lab. Afterwards, we asked LAs for feedback on the lab's content and usability. We recorded audio and video of the think-aloud interviews and screen-recorded

Week	Assessment
1	EMCS Pre-test
5	Midterm 1
7	Momentum Lab
8	Momentum Homework, Momentum Quiz
10	Midterm 2
13	EMCS Post-Test
15	Final Exam

TABLE I. Timeline of class assessments used in this study. The class ran for 15 weeks during the Fall 2019 semester. The Week 7 assessment in bold was the experimental two-cart collision lab.

the lab simulations. We transcribed the interviews and assigned codes based on feedback and suggestions for revision. We made minor changes to the user interface and clarity of instruction of the lab handouts based on our observations and their suggestions.

C. Experimental Design

Students were instructed in momentum concepts during week 6 of the course using a 105-minute lecture, and in week 7 using a 105-minute lab. For the lab activity, students were assigned to either the experimental or control lab group by section.

While students could choose their class sections, they did not know which sections were part of the experimental or control group. Students in the control group completed an entirely hands-on momentum lab using moving carts on tracks. Students in the experimental group completed a hybrid lab, a combination of the hands-on lab and a simulated lab. In each group, students analyzed two completely elastic, two inelastic, and two completely inelastic two-cart collisions. Students in the experimental group analyzed one type of each collision using the simulation, and the remaining collisions using the hands-on lab set. Simulated and hands-on collisions were divided this way to ensure that students in the experimental and control groups were exposed to the same physical concepts, and the hybrid lab contained conceptually equivalent simulation and hands-on parts. Students were instructed to use both equal and unequal masses, with the same initial velocities, for the three types of collisions. They were also free to explore collisions that were not listed on the lab worksheet. After students completed the lab, they completed a lab assessment with momentum questions, working in groups of three.

D. Assessments

We evaluated student performance using matched-set statistical analysis of the Energy and Momentum Concept Inventory (EMCS), a multiple-choice survey of energy and momentum topics [19], administered as a pre- and post-test; topical assessments on momentum (one quiz, one lab assessment, and one homework assignment); two midterm exams; and a final exam. A course timeline is shown in Table I. The course

specific assessments were written to specifically address the scenarios and concepts covered in the momentum lab.

EMCS Pre- and Post-Tests: We used the Energy and Momentum Conceptual Survey (EMCS), to assess student understanding of physics concepts. We administered the EMCS during the first week of the course and again during the 13th week of the course to all students. Students received credit for completing the assessments, but performance on the EMCS pre- and post-tests did not affect students' course grades. We analyzed the entire EMCS pre- and post-tests, the EMCS questions on all momentum topics (questions 3, 5, 7, 10, 11, 13, 14, 16, 18, 19, 21, 23) and the EMCS questions on collision momentum (questions 3, 5, 14, and 21).

Course Assessments on Momentum: We analyzed student performance on momentum-specific course assessments. These assessments included a homework assignment and quiz, and topical momentum questions on the second midterm. We also included the momentum lab assessment, which consisted of two parts: Part 1 was based on data collection during the lab, and Part 2 was based on analysis of momentum bar graphs. All of these assessments affected students' final course grades.

Exams and Course Performance: We analyzed scores for midterm 1, midterm 2, the final exam, and student performance in the course as a whole using numerical final course grades. Midterm 1 occurred before the experimental lab, and we include it to compare student scores before exposure to the hybrid or hands-on lab. We also used course performance to determine the accuracy of course assessments in measuring student understanding. We employ an index of discrimination for this purpose, outlined in Section IID 1.

1. Index of Discrimination

As a way of validating original course assessments, we measured an index of discrimination for each assessment, which compares the number of correct answers from students with overall course grades in the top 25% (C_h), and the number of correct answers from students with overall course grades in the bottom 25% (C_l), using

$$i = \frac{C_h - C_l}{N/4} \quad (1)$$

where N is the total number of final course grades [20–22]. If $i = 1$, the assessment discriminates perfectly between the highest- and lowest-performing students and is likely an accurate indicator of students' understanding.

We applied this metric to original questions on momentum topics presented in exams, quizzes, homework sets, and labs. We considered free-response questions graded with partial credit to be "correct" if a student achieved at least 80% of the possible points.

We calculated this index for the topical course assessments used in our analysis. The index of discrimination for topical momentum questions on midterm 2 was $i_1 = 0.67$, and for the momentum quiz, the index was $i_2 = 0.53$. These were closed-book assessments completed individually. Therefore,

Assessment	$N_{\text{exp}}, N_{\text{control}}$	Experimental $\bar{x} \pm \sigma$	Control $\bar{x} \pm \sigma$	Pearson's r	Cohen's d
EMCS Pre-Test (out of 25)	162, 130	6.54 ± 2.70	6.68 ± 3.00	$-0.02 (p = 0.7)$	–
EMCS Post-Test (out of 25)	162, 130	10.20 ± 4.36	9.35 ± 4.28	$0.10 (p = 0.1)$	–
EMCS Gain (out of 1)	162, 130	0.18 ± 0.27	0.15 ± 0.23	$0.06 (p = 0.3)$	–
EMCS Pre-test Momentum (out of 12)	162, 130	3.92 ± 1.70	4.00 ± 2.05	$0.02 (p = 0.7)$	–
EMCS Post-test Momentum (out of 12)	162, 130	4.86 ± 2.72	4.85 ± 2.26	$0.10 (p = 0.1)$	–
EMCS Pre-test Collisions (out of 4)	162, 130	1.62 ± 0.95	1.71 ± 0.91	$-0.05 (p = 0.4)$	–
EMCS Post-test Collisions (out of 4)	162, 130	1.48 ± 1.15	1.69 ± 1.05	$-0.97 (p = 0.1)$	–

TABLE II. Statistical analysis of EMCS assessment scores. Entries of "–" for Cohen's d indicate that no statistically significant differences between experimental and control groups were found for any of the assessments listed.

we find that the questions on momentum presented in the midterm exams and the topical quiz are reasonably good assessments of students' understanding of momentum topics with indexes greater than 0.3 [20–22]. We include the momentum questions on midterm 2 and the momentum quiz in our analysis because they provide insight into individual student performance.

The lab assessment and homework assignment were open-book and usually completed in groups with TAs and LAs available for questions, unlike the exams and quizzes which were closed-book and completed individually. Due to their collaborative nature, we do not expect these assessments to be as discriminating as the individual assessments. The index of discrimination for the lab assessment on momentum was $i_3 = 0.14$, and the index for the homework on momentum was $i_4 = 0.23$, both below the accepted lower limit of $i = 0.3$, demonstrating that students who score well on the lab and homework assessments do not necessarily score in the top 25% in overall course grades. These non-discriminating assessments, which are group activities and employ lab materials, notes, and help from teaching staff, demonstrate different aspects of student learning than exams and quizzes and are still valuable in our analysis. We include these assignments for further evidence on the effect of hybrid or hands-on labs on student performance.

E. Data Analysis

We included only matched score sets in our analysis of EMCS pre- and post-tests, course exams, momentum-specific assessments, and overall course performance. However, we acknowledge that excluded scores may not have been missing completely at random (MCAR). From the 374 students who completed the course, we included score sets from students who completed both EMCS pre- and post-tests of $N_{\text{EMCS}} = 292$ ($N_{\text{exp}} = 162$ and $N_{\text{control}} = 130$); students who completed assessments on momentum topics (quiz, homework, and midterm 2) of $N_{\text{momentum}} = 369$ ($N_{\text{exp}} = 215$ and $N_{\text{control}} = 154$); students who completed the momentum lab of $N_{\text{lab}} = 362$ ($N_{\text{exp}} = 218$ and $N_{\text{control}} = 144$); and students who completed all exams and have a reported final grade of $N_{\text{exams}} = 374$ ($N_{\text{exp}} = 218$ and $N_{\text{control}} = 156$). We analyzed pre- and post-test scores for the EMCS along

with normalized gain. For all assessments, we calculated two-tailed Pearson's correlation coefficients (r) and Cohen's d effect sizes, along with associated p -values, between the experimental (hybrid lab) and control (hands-on lab) groups.

III. Findings and Discussion

EMCS Pre- and Post Test: All statistical analysis on the EMCS pre- and post-test data resulted in Pearson's r and Cohen's d values with $p > 0.05$. We found no statistically significant differences in total scores on the EMCS pre-test, EMCS post-test, or the normalized gain between the experimental and control groups. We also found no statistically significant differences in performance on the EMCS pre-test and post-test questions on collision momentum topics or questions on all momentum topics (Table II).

Momentum-Specific Assessments: All statistical analysis on momentum-specific assessments resulted in Pearson's r and Cohen's d values with $p > 0.05$. We found no statistically significant differences in total scores on the homework or quiz on momentum topics, momentum-specific questions on midterm 2, or the momentum lab assessment between the experimental and control groups. Additionally, we investigated part 1 (data collection) and part 2 (analysis of momentum bar graphs) of the lab assessment individually. We found no statistically significant difference between performance on individual parts of the lab assessment between the experimental and control groups (Table III).

Course Exams: All statistical analysis on the course exams resulted in Pearson's r and Cohen's d values with $p > 0.05$, indicating no statistically significant differences in total scores on midterm 1, midterm 2, and the final exam between the experimental and control groups (Table IV).

Overall Course Performance: All statistical analysis on the overall course grades resulted in Pearson's r and Cohen's d values with $p > 0.05$. We found no statistically significant differences in overall course grades between the experimental and control groups (Table IV).

Overall, we find that the use of hands-on two-cart collision labs and hybrid two-cart collision labs had an equal effect on the EMCS, exams, topical momentum assessments, and overall course scores. Although $p > 0.05$ for all assessments, we present Pearson's r and p -values to show the range of p -

Assessment	$N_{\text{exp}}, N_{\text{control}}$	Experimental $\bar{x} \pm \sigma$	Control $\bar{x} \pm \sigma$	Pearson's r	Cohen's d
Midterm 2 (Momentum Questions) (out of 10)	215, 154	7.98 ± 2.06	8.04 ± 1.87	$-0.02 (p = 0.7)$	–
Quiz (out of 5)	215, 154	3.71 ± 1.19	3.80 ± 1.11	$-0.04 (p = 0.5)$	–
Homework (out of 20)	215, 154	17.21 ± 2.13	17.16 ± 1.87	$0.01 (p = 0.8)$	–
Lab (out of 10)	218, 144	9.22 ± 0.83	9.31 ± 0.69	$-0.05 (p = 0.3)$	–
Lab: Part 1 (Data Collection) (out of 6)	218, 144	5.50 ± 0.61	5.61 ± 0.50	$-0.10 (p = 0.07)$	–
Lab: Part 2 (Bar Graph Analysis) (out of 4)	218, 144	3.72 ± 0.45	3.70 ± 0.47	$0.03 (p = 0.6)$	–

TABLE III. Statistical analysis of momentum-specific assessment scores. Entries of "–" for Cohen's d indicate that no statistically significant differences between experimental and control groups were found for any of the assessments listed.

Assessment	$N_{\text{exp}}, N_{\text{control}}$	Experimental $\bar{x} \pm \sigma$	Control $\bar{x} \pm \sigma$	Pearson's r	Cohen's d
Midterm 1 (out of 50)	218, 156	40.43 ± 8.43	39.90 ± 7.89	$0.03 (p = 0.5)$	–
Midterm 2 (out of 50)	218, 156	35.30 ± 9.45	35.15 ± 8.90	$0.01 (p = 0.9)$	–
Final Exam (out of 60)	218, 156	42.80 ± 10.51	42.92 ± 10.22	$0.03 (p = 0.5)$	–
Course Grade (out of 100)	218, 156	82.81 ± 11.13	82.52 ± 10.17	$0.02 (p = 0.7)$	–

TABLE IV. Statistical analysis of exam scores and overall course grades. Entries of "–" for Cohen's d indicate that no statistically significant differences between experimental and control groups were found for any of the assessments listed.

values throughout the assessments we investigated.

IV. Conclusions

We investigated student performance on topical momentum assessments before and after completing a two-cart collision momentum lab. Students were divided into two groups, one who used a traditional hands-on lab setup, the other using a hybrid system with both a simulation and hands-on equipment. We did not find evidence of improved or diminished student performance between groups.

This result suggests that the mode of presentation for momentum labs (hands-on or hybrid) in in-person, studio-style introductory mechanics classes does not affect student performance or knowledge gain on momentum topics. The null result is strengthened by the fact that our assessments ranged from individual, closed-book exams to collaborative activities (Section II D). This adds to the current body of knowledge on the effects of simulation, hands-on, or hybrid labs on student understanding and course performance, based on the mixed results of the research literature [2, 10–15, 23]. Because hybrid and hands-on labs may affect course performance equally, it may be useful to implement hybrid labs in classes with limited equipment or lab space, or classes conducted during semesters with blended online and in-person learning due to external circumstances.

V. Limitations and Future Work

This study may not have minimized instructor effects on student performance. There were four different faculty instructors, ten different TAs, and ten different LAs over 5 class sections, which may have caused differences in student performance due to varying teaching methods and graders. Of

note, students in one of the five class sections in this study had a higher average course grade of 86.0 ± 9.4 ($N = 80$) and higher average grades on most assessments than the other four sections. This section was one of three assigned to the experimental group.

In addition, we acknowledge that students in this physics course are mostly junior and senior life science majors at a large, private university, and our subjects are not necessarily representative of all physics students in the United States in demographics or mathematical preparedness [24].

This study focused only on momentum conservation labs using two-cart collisions, which may limit the generalizability of our results. In the future, we intend to conduct similar studies with labs on other concepts, especially more complex topics that are less readily visualized [7].

Some skills only gained in hands-on labs, such as understanding real data and using instruments and tools, are not best measured with the EMCS or other assessments. These learning gains should not be overlooked, and in the future, it may be useful to include assessments on these topics, such as the Physics Lab Inventory of Critical Thinking [25].

Finally, it would be useful to investigate the effects of hands-on vs. simulated or hybrid labs combined with in-person vs. virtual lectures in a blended online and in-person class. This may provide insight on combinations of virtual and in-person instruction that maximize student learning gain and optimize usage of lab materials, accessibility, and time in lab sessions.

Acknowledgments

We would like to thank the Learning Assistants in this course for their participation and feedback. This work is supported by the NSF under IUSE Grant No. DUE-1712159.

-
- [1] NSTA Position Statement, The integral role of laboratory investigations in science instruction, (2007).
- [2] 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).
- [3] S. Chen, H. C. Lo, J. W. Lin, J. C. Liang, H. Y. Chang, F. K. Hwang, G. L. Chiou, Y. T. Wu, S. W. Y. Lee, H. K. Wu, C. Y. Wang, and C. C. Tsai, Development and implications of technology in reform-based physics laboratories, *Phys. Rev. ST Phys. Educ. Res.* **8**, 020113 (2012).
- [4] W. K. Adams, A. Paulson, and C. E. Wieman, What levels of guidance promote engaged exploration with interactive simulations?, *AIP Conf. Proc.* **1064**, 59 (2008).
- [5] J. R. Brinson, Learning outcome achievement in non-traditional (virtual and remote) versus traditional (hands-on) laboratories: A review of the empirical research, *Comp. Educ.* **87**, 218 (2015).
- [6] J. V. Nickerson, J. E. Corter, S. K. Esche, and C. Chassapis, A model for evaluating the effectiveness of remote engineering laboratories and simulations in education, *Comp. Educ.* **49**, 708 (2007).
- [7] S. R. Singer, M. L. Hilton, and H. A. Schweingruber, *America's lab report investigations in high school science* (National Academies Press, 2006).
- [8] S. Chen, The view of scientific inquiry conveyed by simulation-based virtual laboratories, *Comp. Educ.* **55**, 1123 (2010).
- [9] J. Son, P. Narguizian, D. Beltz, and R. Desharnais, Comparing physical, virtual, and hybrid flipped labs for general education biology, *Online Learning* **20**, 10.24059/olj.v20i3.687 (2016).
- [10] A. Jimoyiannis and V. Komis, Computer simulations in physics teaching and learning: a case study on students' understanding of trajectory motion, *Comp. Educ.* **36**, 183 (2001).
- [11] K. Y. Yang and J. S. Heh, The impact of internet virtual physics laboratory instruction on the achievement in physics, science process skills and computer attitudes of 10th-grade students, *J. Sci. Educ. Technol.* **16**, 451 (2007).
- [12] K. E. Chang, Y. L. Chen, H. Y. Lin, and Y. T. Sung, Effects of learning support in simulation-based physics learning, *Comp. Educ.* **51**, 1486 (2008).
- [13] Z. Zacharia and C. 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).
- [14] Z. C. Zacharia and T. de Jong, The effects on students' conceptual understanding of electric circuits of introducing virtual manipulatives within a physical manipulatives-oriented curriculum, *Cognition and Instruction* **32**, 101 (2014).
- [15] M. Jariwala, E. Allen, and A. Duffy, Investigating simulation use on student learning outcomes in introductory physics, in *Physics Education Research Conference 2019*, PER Conference (Provo, UT, 2019).
- [16] R. Beichner, The student-centered activities for large enrollment undergraduate programs (scale-up) project, in *Research-Based Reform of University Physics*, Vol. 1 (2007).
- [17] V. Otero, S. Pollock, and N. Finkelstein, A physics department's role in preparing physics teachers: The Colorado learning assistant model, *Am. J. Phys.* **78**, 1218 (2010), <https://doi.org/10.1119/1.3471291>.
- [18] *Visualizations in Introductory Physics: Education Research (VIPER)* (<http://physics.bu.edu/~duffy/VIPER/>) (Retrieved 05/27/2020).
- [19] C. Singh and D. Rosengrant, Multiple-choice test of energy and momentum concepts, *Am. J. Phys.* **71**, 607 (2003).
- [20] G. J. Aubrecht and J. D. Aubrecht, Constructing objective tests, *Am. J. Phys.* **51**, 613 (1983), <https://doi.org/10.1119/1.13186>.
- [21] L. Ding and R. Beichner, Approaches to data analysis of multiple-choice questions, *Phys. Rev. ST Phys. Educ. Res.* **5**, 020103 (2009).
- [22] J. Day and D. Bonn, Development of the concise data processing assessment, *Phys. Rev. ST Phys. Educ. Res.* **7**, 010114 (2011).
- [23] L. D. Conlin, E. Kuo, and N. R. Hallinen, How null results can be significant for physics education research, *Phys. Rev. Phys. Educ. Res.* **15**, 020104 (2019).
- [24] S. Kanim and X. C. Cid, Demographics of physics education research, *Phys. Rev. Phys. Educ. Res.* **16**, 020106 (2020).
- [25] C. Walsh, K. N. Quinn, C. Wieman, and N. G. Holmes, Quantifying critical thinking: Development and validation of the physics lab inventory of critical thinking, *Phys. Rev. Phys. Educ. Res.* **15**, 010135 (2019).