AAPT PHYSICS EDUCATION
AMERICAN JOURNAL of PHYSICS

Traditionally taught students learn; actively engaged students remember

Scott V. Franklin, Eleanor C. Sayre, and Jessica W. Clark

Citation: American Journal of Physics **82**, 798 (2014); doi: 10.1119/1.4890508
View online: http://dx.doi.org/10.1119/1.4890508
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/82/8?ver=pdfcov
Published by the American Association of Physics Teachers

The Physics Education Research Section (PERS) publishes articles describing important results from the field of physics education research. Manuscripts should be submitted using the web-based system that can be accessed via the American Journal of Physics home page, http://ajp.dickinson.edu, and will be forwarded to the PERS editor for consideration.

Traditionally taught students learn; actively engaged students remember

Scott V. Franklin[a)]
Department of Physics, Rochester Institute of Technology, Rochester, New York 14623-5603

Eleanor C. Sayre[b)]
Department of Physics, Kansas State University, Manhattan, Kansas 66506

Jessica W. Clark
Department of Physics and Astronomy, University of Maine, Orono, Maine 04469

(Received 4 November 2013; accepted 7 July 2014)

A common narrative in physics education research is that students taught in lecture-based classes learn less than those taught with activity-based reformed methods. We show this narrative is simplistic and misses important dynamics of student learning. In particular, we find students of both methods show equal short-term learning gains on a conceptual question dealing with electric potential. For traditionally taught students, this learning rapidly decays on a time scale of weeks, vanishing by the time of the typical end-of-term post-test. For students in reform-based classes, however, the knowledge is retained and may even be enhanced by subsequent instruction. This difference explains the many previous pre- and post-test studies that have found minimal learning gains in lecture-based courses. Our findings suggest a more nuanced model of student learning, one that is sensitive to time-dependent effects such as forgetting and interference. In addition, the findings suggest that lecture-based courses, by incorporating aspects designed to reinforce student understanding of previously covered topics, might approach the long-term learning found in research-based pedagogies. *© 2014 American Association of Physics Teachers.*

[http://dx.doi.org/10.1119/1.4890508]

I. INTRODUCTION

Students in traditionally taught courses learn less than peers taught with reform-based methods. These results rest upon hundreds of "pre/post" studies that give students an assessment before instruction and then give the same students a second, matched test well after instruction (often at the end of the quarter or semester). Pre/post testing of students is functionally the standard for assessing learning in physics.[1] Thornton and Sokoloff[2] used the method to establish the validity of the FMCE and demonstrate the efficacy of active engagement classrooms, a study reproduced on a much larger scale by Hake.[3] Pre/post testing fails, however, to reveal the full dynamism of student learning.

Research on forgetting and interference[4–7] shows that the dynamics of learning are quite nuanced and difficult to measure. For example, learning curves, such as those predicted by the Rescorla-Wagner model,[6] for repeated training increase quickly at first and then level off, suggesting a simple error-reduction model of learning. "Forgetting curves," such as those studied by Ebbinghaus,[8] in a wide variety of tasks reveal that memory performance decays (approximately) exponentially after training ceases, with timescales varying from seconds to decades. A third phenomenon, interference, occurs when two pieces of related information (or tasks) are learned. Performance on one can significantly decrease when the second is learned either before (proactive) or after (retroactive), and the amount of interference increases with the degree of similarity between the two pieces of information.[4,7] This (and other) research on learning, forgetting, and interference has strong implications for traditional pre/post testing, raising the question of whether seemingly significant learning gains may actually result from learning that is short-lived or context-dependent.

In this study, we use the Response Curve Methodology (RCM),[9] which has been shown to produce a much more detailed picture of student understanding before, during, and after instruction. In RCM, described in more detail below, different groups of students take the same test, but at different times. As long as each group is large enough to be representative of the population, the comparison between groups corresponds to the population's understanding at different points in time, which can be within several days of each other. The statistics require large student populations—as well as the time and effort to generate appropriate student groupings—but RCM has been used to study the dynamics of students' understanding in electrostatics,[9,10] circuits and magnetic field direction,[11] Newton's third law,[12–14] force and motion,[11] and vector subtraction.[15]

In this paper, we draw attention to a comparison between two institutions that use different teaching methods in introductory physics, one traditional and the other reform-based. Viewed through the "pre/post" test lens, the results look no different from the many previously published studies doing similar comparisons. The RCM, however, reveals a striking dynamism in the traditionally taught students' understanding that contradicts the popular claim that students do not learn in these environments.

II. METHODS

A. Populations

Our first institution is a large, four-year, public, research-intensive, doctoral-granting institution (the "University"). Each year approximately 700 students take introductory calculus-based physics in the off-sequence quarters, which are offered in a very traditional format: conventional chalk-and-talk lecture for three one-hour sessions each week, 1 h of recitation, and a 2 h cookbook-style confirmatory laboratory. On average, each lecture section enrolls 170 students and two sections are taught each of two quarters.

Our second institution (the "Institute") is[16] a large, four-year, private university with high undergraduate enrollment, balanced arts and sciences/professions, and some graduate students in non-physics disciplines. Each year about 1000 students take introductory calculus-based physics, which is offered in a workshop format that integrates lecture, experiment, and short group activities. Adapted from the SCALE-UP project,[17] the classes meet for three two-hour sessions each week, with students seated at tables of six and working in small groups. Classrooms accommodate up to forty-two students, with enrollment in each section varying. Multiple sections of calculus-based electromagnetism (E&M) are offered every quarter, with most students taking the course in the fall of their sophomore year. In order to ensure that the same material is covered in all sections, a mid-term and final exam are common across all sections.

For both institutions at the time of this study, the academic year was divided into four 10-week quarters (including the summer term). A comparison of the two schools' E&M course syllabi revealed that the order of topics and depth of coverage was comparable. Though the University offered E&M as the second course in the sequence while the Institute offered it as the third course, this does not seem to have affected students' pre-test scores.

The demographics of the overall student population that enrolls in these two institutions are similar[18] (Table I), as are the subset of students who enroll in the introductory calculus-based physics courses.

Table I. Enrollment profiles for the Institute and the University.

	University	Institute
Percent of applicants admitted	64%	58%
ACT composite (25th/75th percentiles)	26/30	25/30
Total UG enrollment	42,916	14,849
Enrollment by gender	53% (male)	72% (male)
Full-time students	91%	82%

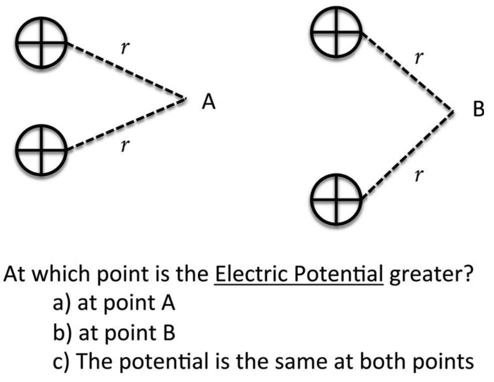
B. Response curve methodology

First used in physics by Sayre and Heckler,[9] the RCM gives short conceptual quizzes to different sub-groups of the population in successive weeks. This avoids test-retest effects that would occur if the same group took the quiz twice. Tests can be administered in a controlled setting (e.g., an education research lab[9–11,15]), as short quizzes,[12] or online.[14] At the University, data are collected in a research lab and binned by week to comprise each group; students are quasi-randomly assigned to groups and offered homework credit for completion.[9] At the Institute, data are collected as short quizzes given at the beginning of the first workshop each week. Different sections are therefore different groups, with Sec. I (for example) completing the relevant quiz in week 3 and Sec. II taking the quiz in week 8. The particular week in which a section took a specific quiz was randomly assigned, and all quizzes were administered at the beginning of class, once per week, in paper format. Students had five to ten min to complete each quiz, which were sometimes appended to an instructor-generated quiz that was not used for research. Previous analysis suggests that different sections are not statistically significantly different at any given time,[14] and thus these two variants on the RCM are comparable.

For this type of analysis to be valid, the group responses to a single question should be approximately normally distributed and have similar variances, which we have confirmed. Because sections are statistically independent, we can compare the performance of different sections across weeks, essentially capturing student understanding on a weekly time scale. A time plot of average performance, termed the *response curve*, is sensitive to the particulars of the week—the current topic of instruction and coincidence with exams or homework. The conventional pre/post test corresponds to the first and last points on such a curve, and can miss much of the dynamic evolution of understanding. Error bars on the response curve are determined using a binomial distribution, and data are collapsed across the fall and winter quarters (by week) to increase sample size. Students in, for example, Week 3 in the Winter see the same activities, labs, and lectures as those in Week 3 in the Fall.

C. Electric potential task

While between-student testing allows us to probe multiple concepts simultaneously, this study focuses on a single question involving electric potential, shown in Fig. 1. This was



At which point is the Electric Potential greater?
a) at point A
b) at point B
c) The potential is the same at both points

Fig. 1. Conceptual question probing student understanding of the scalar nature of electric potential.

but one question in a (slightly) longer quiz involving charged particles, electric forces, fields, and potentials. We chose this topic because there is ample time before instruction to gather baseline data, and ample time afterwards to measure longer-term effects after the evanescent results of instruction.

As seen in Fig. 1, students are asked to compare the electric potential that results from two different configurations of charged particles. Because the potential is a scalar quantity, it depends only on the distance from particle to location and is the same at both points A and B. A similar question on the same quiz asked about the relative strength of the electric field, a vector quantity that has a larger magnitude at point A. Prior study[9] showed that the correct response to this vector question was suppressed during instruction on the scalar concepts of potential and resistance but returned later in the quarter during instruction on the vector-based concept of magnetic field. The conclusion was that scalar concepts *interfered* with students' abilities to recall the vector concept, a cognitive dissonance that disappeared when new vector quantities were discussed. A similar, albeit inverse, relationship was reported for scalar potential in the traditional course. In this study, we look at the effects of different teaching methods on the shape of the curve. In particular, we are interested in whether the shape of the curve during and after instruction is similar or different.

III. RESULTS

Our results for both traditionally taught University and activity-based Institute students are shown in Fig. 2. The data are normalized by the score of the students answering the question in the first week of the quarter, prior to any instruction. Recall that due to the response curve methodology, individual students do not see the question at any subsequent point in the quarter; each point in the graph is a separate group of students.

For both institutions, electric potential is taught during the end of the second and beginning of the third weeks of the quarter, and tested in the fourth week (at the Institute) or fifth week (at the University). (In each, the order of instruction moves from charges, to electric forces, then to electric fields, and finally to potential.) Thus, we see that the scores remain at or near the pre-test level for the first two weeks of the course. The activity-based class begins to cover the topic toward the end of the second week, and so the scores for week 3 (taken at the beginning of the week) show the effect of this instruction. The traditionally taught students do not learn the material until week 3, and so the first indication of learning is in their dramatic improvement in week 4.

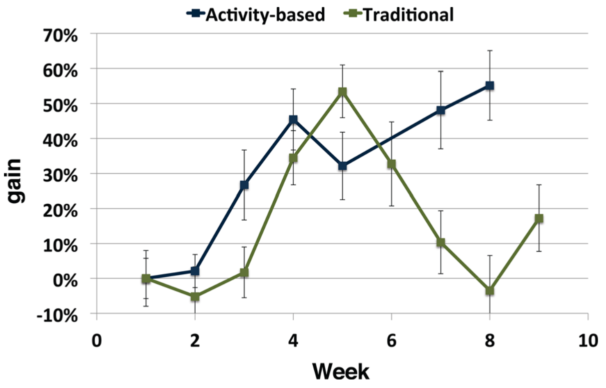


Fig. 2. Student performance on electric potential conceptual question as a function of time, relative to the pre-test score in week 1. Traditionally taught students (dark line, $N \approx 1{,}000$) learn as much as their actively engaged peers (light line, $N \approx 350$) when the material is first taught (weeks 3 and 4) but rapidly forget much of the knowledge by the quarter's end. Error bars on the response curve are determined using a binomial distribution; some error bars are smaller than the points shown.

Figure 2 shows quite clearly that the traditionally taught students learn just as much as their activity-based peers, with both instruction methods leading to approximately the same normalized gains [the ratio of learned to initially unknown material, defined as (post % – pre %)/(100% – pre %)]. Not only are the relative learning gains virtually identical, but so are the short-term fluctuations, with each population showing a small decay in the week after the examination (larger in the traditionally taught students, but this is not statistically significant).

The truly surprising result occurs in the subsequent weeks after the exam. It is here that the true impact of the reform-based classroom is seen to be not in the amount learned, but rather in the amount *remembered* (or, alternately, the long-lasting nature of what has been learned). The performance of the traditionally taught students on the task continues to plunge in weeks 6, 7, and 8, returning to its initial value only 3 weeks after the midterm. The graph shows a small bounce in week 9, as students begin to prepare for the final exam, but it does not approach the mid-term results. One could infer that the learning of the traditionally taught students is short-lived, transient, and (perhaps) superficial.

On the other hand, students in the activity-based classes retain their learned knowledge through the end of the quarter, subsequently showing signs of enhancing the understanding and increasing the learning gains. We emphasize that this result is even more surprising given the vector-based nature of the instruction during the last third of the quarter (true for both institutions). In the introductory sequence, there is no magnetic analog for electric potential, and so students spend several weeks concerned primarily with vector quantities (magnetic field, flux, etc.) On the graph, there is no sign of the interference that scalar quantities show on the recall of electric field. Compared to the learning of the traditionally taught students, this learning seems sustained and deep.

For these students, a traditional pre/post test at the quarter's end would show long-term learning gains of 60% for the activity-based class vs 20% for the traditionally taught students. These results are consistent with the many published studies that report minimal learning gains in traditionally taught classes. Such a study, however, would fail to detect the short-term learning that actually occurred in the traditionally taught students.

IV. CONCLUSIONS

What does it mean that students can demonstrate significant, if short-term, learning even in traditional lectures? At minimum, we should revise our understanding of assessment, accounting for temporal decay or other temporal phenomena. As a method for assessing teaching, pre/post testing does not measure long-term understanding, which must be taken into account by physics education researchers who use pre/post testing as a research tool. It is also important for physics departments that use pre/post testing to assess teaching or curricular revisions. If one can dramatically manipulate students' pre- and post-scores by seemingly small changes in testing date, then the timing of the tests must be considered as we compare the performance of students in different classes.

More broadly, the assumption of students' long-term understanding suffuses physics programmatic design. In the four-year curriculum, many topics are revisited in several classes, with an assumption that material covered previously is remembered. The example in this paper involves electrostatics that recurs in junior-level electrodynamics courses. Our research suggests that students in traditionally taught classes may learn enough to pass their exams but do not remember the topics for subsequent courses. The focus of future reform should be on the retention, not just the acquisition, of knowledge, and account for the amount of material that a student can forget.

A revised focus on retention might increase the effectiveness of even the traditional class environment. Traditionally taught courses, with their large lecture sections and cookie-cutter laboratories, scale well to ever-increasing student enrollments. Efforts to improve such courses could improve the learning at many institutions that cannot afford to use fully reform-based pedagogy the way that the Institute has. Lessons from reform-based classes and other research[19,20] suggest that students must revisit prior knowledge, actively integrating new material with old into a coherent framework.

In conclusion, student learning is complicated, context-dependent, and dynamic, and it merits further study in diverse learning environments and should be taken into account in future curricular and programmatic development.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the many instructors at both institutions who allowed us access to their students and classrooms to administer these tasks. This work was supported by the Rochester Institute of Technology and the National Science Foundation under Grant No. DUE 1240782.

[a] Electronic mail: svfsps@rit.edu
[b] Electronic mail: esayre@gmail.com

[1] Matthew A. Kohlmyer, Marcos D. Caballero, Richard Catrambone, Ruth W. Chabay, Lin Ding, Mark P. Haugan, M. Jackson Marr, Bruce A. Sherwood, and Michael F. Schatz, "A tale of two curricula: The performance of two thousand students in introductory electromagnetism," Phys. Rev. ST Phys. Ed. Res. **5**, 020105–020114 (2009).

[2] Ronald K. Thornton and David R. Sokoloff, "Assessing student learning of newton's laws: The force and motion conceptual evaluation and the evaluation of active learning laboratory and lecture curricula," Am. J. Phys. **66**, 338–353 (1998).

[3] Richard R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," Am. J. Phys. **66**, 64–74 (1998).

[4] Leo Postman and Benton J. Underwood, "Critical issues in interference theory," Memory Cogn. **1**, 19–40 (1973).

[5] Goerge B. Semb, John A. Ellis, and John Araujo, "Long-term memory for knowledge learned in school," J. Educ. Psychol. **85**(2), 305–316 (1993).

[6] Robert A. Rescorla and Allan R. Wagner, "A theory of Pavlovian conditioning," in *Classical Conditioning II: Current Theory and Research*, edited by Abraham H. Black and William F. Prokasy (Appleton Century Crofts, New York, 1972), pp. 64–99.

[7] Mark E. Bouton, "Context, time, and memory retrieval in the interference paradigms of pavlovian learning," Psychol. Bull. **114**(1), 80–99 (1993).

[8] Hermann Ebbinghaus, *Memory* (Teachers College, Columbia University, 1913).

[9] Eleanor C. Sayre and Andrew F. Heckler, "Peaks and decays of student knowledge in an introductory E&M course," Phys. Rev. ST Phys. Ed. Res. **5**, 013101–013106 (2009).

[10] Eleanor C. Sayre and Andrew F. Heckler, "Evolution of student knowledge in a traditional introductory classroom," in *AIP Conference Proceedings*, Vol. 1064, 195–198 (2008).

[11] Andrew F. Heckler and Eleanor C. Sayre, "What happens between pre- and post-tests: Multiple measurements of student understanding during an introductory physics course," Am. J. Phys. **78**(7), 768–777 (2010).

[12] Jessica Clark, Eleanor C. Sayre, and Scott V. Franklin, "Fluctuations in student understanding of Newton's 3rd law," in *AIP Conference Proceedings*, Vol. 1289, 2010, pp. 101–104.

[13] Eleanor C. Sayre and Scott V. Franklin, "Learning, retention, and forgetting in university physics," in *Proceedings of the 2011 American Educational Research Association Annual Meeting*, American Educational Research Association, 2011.

[14] Eleanor C. Sayre, Scott V. Franklin, Stephanie Dymek, Jessica Clark, and Yifei Sun, "Learning, retention, and forgetting of Newton's third law throughout university physics," Phys. Rev. ST – Phys. Ed. Res. **8**(1), 010116–010125 (2012).

[15] Tianren Wang and Eleanor C Sayre, "Maximum Likelihood Estimation (MLE) of students' understanding of vector subtraction," in *AIP Conference Proceedings*, Vol. 1289, 2010, pp. 329–332.

[16] 2010 Carnegie classification, <http://classifications.carnegiefoundation.org/>.

[17] Robert J. Beichner, Jeffery M. Saul, Rhett J. Allain, Duane L. Deardorff, and David S. Abbott, "Introduction to scale-up: Student-centered activities for large enrollment university physics," presented at the Annual Meeting of the American Society for Engineering Education, Seattle, Washington (2000).

[18] U.S. Department of Education, Institute of Education Sciences, technical report, National Center for Education Statistics (2013).

[19] Anton E. Lawson, "Using the learning cycle to teach biology concepts and reasoning patterns," J. Biol. Educ. **35**(4), 165–169 (2001).

[20] David E. Meltzer and Ronald K. Thornton, "Resource Letter ALIP-1: Active-Learning Instruction in Physics," Am. J. Phys. **80**, 478–496 (2012).