AMERICAN JOURNAL of PHYSICS

Classroom demonstrations: Learning tools or entertainment?

Catherine Crouch, Adam P. Fagen, J. Paul Callan, and Eric Mazur

Citation: Am. J. Phys. **72**, 835 (2004); doi: 10.1119/1.1707018
View online: http://dx.doi.org/10.1119/1.1707018
View Table of Contents: http://ajp.aapt.org/resource/1/AJPIAS/v72/i6
Published by the American Association of Physics Teachers

Additional information on Am. J. Phys.

Journal Homepage: http://ajp.aapt.org/
Journal Information: http://ajp.aapt.org/about/about_the_journal
Top downloads: http://ajp.aapt.org/most_downloaded
Information for Authors: http://ajp.dickinson.edu/Contributors/contGenInfo.html

ADVERTISEMENT

Downloaded 18 May 2013 to 142.150.190.39. Redistribution subject to AAPT license or copyright; see http://ajp.aapt.org/authors/copyright_permission

All submissions to PERS should be sent (preferably electronically) to the Editorial Office of AJP, and then they will be forwarded to the PERS editor for consideration.

Classroom demonstrations: Learning tools or entertainment?

Catherine H. Crouch[a)]
Department of Physics and Division of Engineering and Applied Sciences, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138

Adam P. Fagen[b)]
Program in Molecular Biology and Education, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138

J. Paul Callan[c)] and Eric Mazur[d)]
Department of Physics and Division of Engineering and Applied Sciences, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138

(Received 20 September 2002; accepted 20 February 2004)

We compared student learning from different modes of presenting classroom demonstrations to determine how much students learn from traditionally presented demonstrations, and whether learning can be enhanced by simply changing the mode of presentation to increase student engagement. We find that students who passively observe demonstrations understand the underlying concepts no better than students who do not see the demonstration at all, in agreement with previous studies. Learning is enhanced, however, by increasing student engagement; students who predict the demonstration outcome before seeing it, however, display significantly greater understanding.
© *2004 American Association of Physics Teachers.*
[DOI: 10.1119/1.1707018]

I. INTRODUCTION

Classroom demonstrations, a standard component of science courses in schools and universities, are commonly believed to help students learn science and to stimulate student interest. There is little doubt that well-performed demonstrations achieve the latter objective; one study found that demonstrations are among students' favorite elements of introductory undergraduate physics courses.[1] However, research on student learning from demonstrations suggests that traditional demonstrations may not effectively help students grasp the underlying scientific concepts or recognize and correct scientific misconceptions they may have.[2–4]

Science education research shows that most students learn more from instruction that actively engages them rather than from traditional methods in which they are passive spectators.[5] A number of approaches to instruction that are designed to engage students more actively have therefore been developed. Many of the most successful approaches consist of a set of carefully refined student activities designed to address research-identified student difficulties with the material. These approaches specify both the instructional methods and the content to be covered.[6] For example, Sokoloff and Thornton's Interactive Lecture Demonstrations (ILD)[7] replace 1 h of lecture per week with a sequence of five to seven highly interactive, demonstration-based activities.[7]

In our study, we examined whether student learning from demonstrations, which were originally developed for traditional use, could be enhanced simply by varying the mode of presentation. We find that students who passively observe demonstrations understand the underlying concepts no better than students who do not see the demonstration at all, in agreement with previous studies. Learning in enhanced, however, by increasing student engagement; students who predict the demonstration outcome before seeing it, however, display significantly greater understanding.

II. DEMONSTRATION PEDAGOGIES

We examined three different modes of presentation: (1) *observe*, the traditional approach to demonstrations, in which students watch the demonstration and hear the instructor's explanation, (2) *predict*, in which students record their predictions of the demonstration outcome, observe the demonstration, and hear the instructor's explanation; and (3) *discuss*, in which students record predictions, observe the demonstration, discuss it with fellow students, and finally hear the instructor's explanation. We compared results from these three modes with those from a no-demonstration (control) group who did not see the demonstration at all.[8]

Predictions were elicited by asking the entire class a question and giving students a few minutes to think and record their predictions, without discussion. In the predict mode, the question was posed on a viewgraph together with a multiple-choice list of possible answers, in a manner similar to a ConcepTest.[9] Student predictions were recorded with an electronic polling system.[10] In the discuss mode, the question was posed in open-ended form on a worksheet, on which students recorded their predictions. After the students made their predictions, they were shown the viewgraph used for predict mode and they reported the answer choice closest to

Downloaded 18 May 2013 to 142.150.190.39. Redistribution subject to AAPT license or copyright; see http://ajp.aapt.org/authors/copyright_permission

Table I. Rates of correct responses by outcome (R_{outcome}) and explanation (R_{expln}) for each mode (combined data from all seven demonstrations), and average time required for each mode. The p values indicate the statistical significance of a particular rate of correct response compared to the *no demonstration* group; $p<0.05$ is considered statistically significant (Ref. 15). Effect size indices $h_{outcome}$ and h_{expln} are a measure of the differences between a particular treatment group and the *no demonstration* group, normalized in a manner suitable for proportion data; h is calculated as specified by Cohen (Ref. 16). Values of h between 0.2 and 0.5 are considered small, and h values between 0.5 and 0.8 are considered medium.

		Outcomes			Explanations			
Mode	N	R_{outcome}	p_{outcome}	h_{outcome}	R_{expln}	p_{expln}	h_{expln}	Time (min)
no demo	297	61%	⋯	⋯	22%	⋯	⋯	0
observe	220	70%	0.03	0.19	24%	0.64	0.05	11
predict	179	77%	<0.001	0.35	30%	0.04	0.18	13
discuss	158	82%	<0.0001	0.47	32%	0.02	0.23	21

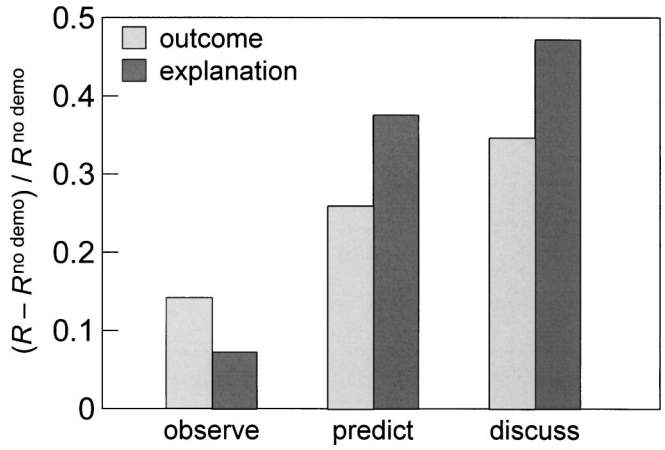

Fig. 1. Improvement in rate R of correct outcomes (light gray) and explanations (dark gray) for each mode over the *no demonstration* (control) group, normalized to the rate for the control group.

their original prediction through the electronic polling system. The worksheets also prompted students to record their observation of the demonstration, explicitly compare their predictions to their observations, and discuss the demonstration with two or three other students.[11]

We performed this study in a 133-student introductory physics course for premedical students. In addition to 2.5 h of lecture per week, the class met weekly for collaborative study in groups of three or four, in sections of 15–20 students per group, directed by teaching assistants. During the section meetings, we presented a series of seven demonstrations.[12] Every week, each section saw that week's demonstration in one of the three modes, or did not see the demonstration at all. Assignments of modes to sections rotated from week to week so that all students participated in each mode or in the control group roughly the same number of times. Students who were absent from a section in a given week were also assigned to the control group.[13]

III. ASSESSMENT

At the end of the semester, we administered a free-response test[11] to assess student learning from the demonstrations. The test presented physical situations identical to the demonstrations, without mentioning that they had been shown in class. We asked students to predict the outcome of the situation and explain the reason for their prediction. Several follow-up questions were designed to help determine whether students understood the underlying physics.

Of the 133 students enrolled in the course, 122 completed the test and responded to all questions, giving answers that displayed genuine effort. We analyzed those 122 tests and classified the responses to each question separately by outcome (correct or incorrect) and by explanation (correct or incorrect). Outcome correctness was based on the questions that were identical to the demonstrations, and did not consider the correctness of answers to follow-up questions. Two different graders were involved in classifying explanations and some cross-checking of classifications was done to ensure consistency. Only complete, fully correct outcomes and explanations were classified as correct.

IV. RESULTS AND DISCUSSION

Table I shows the aggregate results of the end-of-semester test for all seven demonstrations:[14] (1) the rates R of correct outcomes and correct explanations for each mode and for the no-demonstration (control) group, (2) p values (statistical significance testing[15]) and effect size indices h (normalized measure of differences in proportions[16]) for $R^{\text{mode}}-R^{\text{no demo}}$ (the differences in R between treatment groups and the control group), and (3) the average class time required for each mode. Figure 1 displays the improvement in the rates of correct outcomes and explanations for each mode over those rates for the control group, normalized to the rates for the control group. Especially noteworthy is that students in the observe group—those who saw the demonstration in traditional fashion—score only marginally better on explanations than students in the control group, and the difference is not statistically significant. The students in the observe group display no greater understanding of the underlying concepts than those who did not see the demonstration at all.

Increasing student engagement by asking students to predict the outcome of the demonstration doubles both the normalized improvement and the effect size index for outcomes, and nearly quadruples them for explanations, without substantially increasing the time spent. On average, the prediction adds only 2 min to the time required to show and explain the demonstration. The additional improvement in the discuss mode over the predict mode is not large, especially in explanations, despite the additional 8 min per demonstration.

The overall rate of fully correct explanations is modest even for predict and discuss (30% and 32%, respectively), perhaps indicating that there are limits to what students can learn from individual demonstrations. Many successful research-based strategies for teaching physics involve a sequence of activities designed to address particular student difficulties with the material.[6] Finally, student learning from certain demonstrations, even when performed interactively, may be limited because the demonstrations themselves are not designed to address particular student difficulties. Kraus[3] has found that simply discussing certain traditional demonstrations has a limited impact on understanding. If so, demonstrations that are designed to address difficulties should lead to a greater improvement in student learning. We conducted a follow-up study using the same protocol, in which one-third of the demonstrations were drawn from the ILD materials,[7] and half were standard demonstrations. With the

Downloaded 18 May 2013 to 142.150.190.39. Redistribution subject to AAPT license or copyright; see http://ajp.aapt.org/authors/copyright_permission

ILD, the predict and discuss groups indeed display greater improvement relative to the observe and control groups, than with standard demonstrations.[17]

We also compared the rate of correct predictions made by students in class to the end-of-semester test results. The predict and discuss groups differed greatly in the accuracy of students' predictions: for all seven demonstrations combined there were 26% correct predictions for the predict mode in comparison to 41% correct for the discuss mode. While making predictions in class, students saw a list of possible outcomes in the predict mode, but did not see this list in the discuss mode until after making predictions. The more open-ended prediction process in the discuss mode may have pushed students to think more carefully about the demonstration. It is also possible that the presence of plausible distracters in predict mode increased the rate of errors in the predictions. Previous research by Steinberg and Sabella[18] and Rebello and Zollman[19] gives mixed results in comparing the difficulty of open-ended and multiple-choice question formats, when the multiple-choice distracters are based on common student misconceptions.

It is surprising that the discuss group performed only marginally better than the predict group on our end-of-semester test, in spite of the higher rate of correct in-class predictions, the more open-ended process of prediction, the post-demonstration discussion between students, and the additional time spent. According to McDermott,[5] for students to discover and correct their own misunderstandings, instruction should elicit students' ideas, then confront students with errors in those ideas, and finally offer students the opportunity to resolve the errors. In this study, both the predict and discuss modes elicit students' ideas through the prediction stage and confront those who make incorrect predictions with their errors through the demonstration itself. However, only the discuss mode offers students the opportunity to explicitly resolve their misunderstandings through discussion and writing. Perhaps for typical mechanics demonstrations, in contrast to activities specifically designed to address common student difficulties, enhancement of learning comes primarily from securing students' attention and exposing erroneous ideas through the prediction process. We indeed found[17] that the discuss mode confers more benefit with research-based ILD than with standard demonstrations.

V. CONCLUSIONS

Our results lead to two clear conclusions: First, despite popular beliefs to the contrary, students learn little, if anything, from traditionally presented classroom demonstrations. Second, giving students a couple of minutes to predict the outcome and record their predictions costs very little time and yields better understanding. Involving students by having them predict the outcome of demonstrations is a simple step toward increasing student engagement and improving learning from demonstrations. We are presently investigating the benefits of this prediction process in more depth with research-based demonstrations.

ACKNOWLEDGMENTS

This work was funded by the Division of Undergraduate Education of the National Science Foundation under Grant No. DUE-9980802. We thank Wolfgang Rueckner and Nils Sorenson for preparing apparatus for demonstrations; the staff and students of Physics 1a at Harvard University; Howard Georgi, Edw. S. Ginsberg, Dudley Herschbach, and Pamela Kraus for valuable feedback on the manuscript; and Philip Sadler, David Sokoloff, Stamatis Vokos, and Michael Wittmann for helpful conversations.

[a]Present address: Department of Physics and Astronomy, Swarthmore College, Swarthmore, Pennsylvania 19081.
[b]Present address: National Research Council, 500 Fifth St. NW, Washington, District of Columbia 20001.
[c]Present address: McKinsey & Co., 600 Fourteenth St. NW, Washington, District of Columbia 20005.
[d]Author to whom correspondence should be addressed. Electronic mail: mazur@physics.harvard.edu

[1]R. Di Stefano, "Preliminary IUPP results: Student reactions to in-class demonstrations and to the presentation of coherent themes," Am. J. Phys. **64**, 58–68 (1996).
[2]I. A. Halloun and D. Hestenes, "Common sense concepts about motion," Am. J. Phys. **53**, 1056–1065 (1985).
[3]P. A. Kraus, "Promoting active learning in lecture-based courses: demonstrations, tutorials, and interactive tutorial lectures," Ph.D. thesis, University of Washington (1997), available on University Microfilms, Inc., UMI No. 9736313, and references therein. Kraus includes an excellent, extensive bibliography and review of education research on demonstrations.
[4]W.-M. Roth, C. J. McRobbie, K. B. Lucas, and S. Boutonné, "Why may students fail to learn from demonstrations? A social practice perspective on learning in physics," J. Res. Sci. Teach. **34**, 509–533 (1997).
[5]See, for example, L. C. McDermott, "Oersted Medal lecture 2001: 'Physics education research—the key to student learning,'" Am. J. Phys. **69**, 1127–1137 (2001).
[6]Several such approaches are listed in Sec. VII of L. C. McDermott and E. F. Redish, "Resource letter: PER-1: Physics education research," Am. J. Phys. **67**, 755–767 (1999).
[7]D. R. Sokoloff and R. K. Thornton, "Using interactive lecture demonstrations to create an active learning environment," Phys. Teach. **35**, 340–347 (1997).
[8]A preliminary version of this study was first reported at an American Association of Physics Teacher's Meeting: J. P. Callan, C. Crouch, and E. Mazur, "Classroom demonstrations: education or mere entertainment?" AAPT Announcer **29**, 89 (1999).
[9]E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, Upper Saddle River, NJ, 1997).
[10]We recorded predictions both for later analysis and to serve as an attendance record; in the *observe* mode, the instructor recorded attendance separately.
[11]See EPAPS Document No. E-AJPIAS-72-007406 for an example of a *predict* mode viewgraph and a *discuss* mode worksheet in an online appendix. Also provided is the end-of-semester test used for assessment. In *discuss* mode, students worked in the same groups in which they did other collaborative exercises throughout the section meeting. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps. See the EPAPS homepage for more information.
[12]The seven demonstrations were, in the order presented during the semester: (1) Driving a radio-controlled model car on a lightweight roadbed with little friction beneath the roadbed and the floor, so that the roadbed moves when the car starts; (2) colliding a rubber ball and a putty ball with a bowling pin to see which ball knocks the pin over; (3) comparing tension in a string when fixed to a wall at one end versus when attached to weights at both ends; (4) comparing the time of travel for balls on two tracks which have the same starting and ending points, but one track dips lower than the other in the middle; (5) demonstrating the minimum starting height for a model car to travel around a vertical loop without falling; (6) for a puck revolving on tabletop at the end of a string, showing the effect of string length on angular speed; (7) for a beam supported at each end by platform scales, showing effect of position of load on scale readings. The end-of-semester test used to assess understanding (Sec. IV) describes each demonstration in detail and can be found online (Ref. 11).
[13]We confirmed that the populations in each mode were equivalent by calculating the average student final grade for each mode. In calculating the averages, each student was included as many times as he or she had participated in that mode. The average final grades for the *observe*, *predict*, and *discuss* modes were 71.58%, 71.59%, and 71.56%, respectively; the average final grade for the *no demo* mode was 70.15%. The difference

Downloaded 18 May 2013 to 142.150.190.39. Redistribution subject to AAPT license or copyright; see http://ajp.aapt.org/authors/copyright_permission

between the *no demo* group and the other groups is not statistically significant at the $p<0.05$ level.

[14]The rates of correct outcomes and explanations were computed by aggregating (from all seven demonstrations) all student responses that were associated with a given demonstration mode, and then calculating the fraction of correct responses. In other words, the rate is not an average over rates for individual demonstrations.

[15]The p values given are for a two-tailed test of the hypothesis $R^{\text{mode}}=R^{\text{no demo}}$; such values are commonly interpreted as the probability that the treatment group differs from the control group. A difference is statistically significant if $p<0.05$. Our calculation of p values follows the approach outlined in Sec. 8.2 of D. S. Moore and G. P. McCabe's, *Introduction to the Practice of Statistics* 3rd ed. (Freeman, New York, 1998).

[16]Cohen's effect size index h is calculated from the difference in the arcsine transformation of the proportions; see J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* 2nd ed. (Erlbaum Associates, Hillsdale, NJ, 1988), Chap. 6. Data in which the individual values are binary and the average is a number between zero and one are called proportions.

[17]A. P. Fagen, ''Assessing and enhancing the introductory science course in physics and biology: Peer instruction, classroom demonstrations, and genetics vocabulary,'' Ph.D. thesis, Harvard University (2003).

[18]R. N. Steinberg and M. S. Sabella, ''Performance on multiple-choice diagnostics and complementary exam problems,'' Phys. Teach. **35**, 150–155 (1997).

[19]N. S. Rebello and D. A. Zollman, ''The effect of distracters on student performance on the Force Concept Inventory,'' Am. J. Phys. **72**, 116–125 (2004).

AMERICAN JOURNAL OF PHYSICS ON THE INTERNET

For access to the online version of AJP, go to AIP's Online Journal Publishing Service: http://aapt.org/ajp.

Browsing abstracts and tables of contents of online issues (beginning with January 1999) and searching of titles, abstracts, etc., back to 1975 is unrestricted.

Institutional and library (''nonmember'') subscribers have access via IP addresses to the full text of articles that are online; to activate access, these subscribers should contact AIP, Circulation & Fulfillment Division, 800-344-6902; outside North America 516-576-2270 or subs@aip.org.

Individual (''member'') subscribers to the paper version who wish (for an additional fee) to add access to the online version should similarly contact AAPT or go to the AAPT website: http://www.aapt.org/.

AJP's home page at the editorial office (http://www.kzoo.edu/ajp/) contains the Table of Contents of the next month's issue several weeks before publication, the Tables of Contents for the last several years, the ''Statement of Editorial Policy,'' ''Information for Contributors,'' membership rates and a membership application, library (''nonmember'') subscription rates, etc.

Downloaded 18 May 2013 to 142.150.190.39. Redistribution subject to AAPT license or copyright; see http://ajp.aapt.org/authors/copyright_permission