

A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms

Chance Hoellwarth, Matthew J. Moelter, and Randall D. Knight
Department of Physics, California Polytechnic State University, San Luis Obispo, California 93407

We present data on student performance on conceptual understanding and on quantitative problem-solving ability in introductory mechanics in both studio and traditional classroom modes. The conceptual measures used were the Force Concept Inventory and the Force and Motion Conceptual Evaluation. Quantitative problem-solving ability was measured with standard questions on the final exam. Our data compare three different quarters over the course of 2 years. In all three quarters, the normalized learning gain in conceptual understanding was significantly larger for students in the studio sections. At the same time, students in the studio sections performed the same or slightly worse on quantitative final exam problems.

I. INTRODUCTION

Recent research into physics education has looked at both what students learn and how they learn it.¹ Although much work has examined student conceptual understanding, the connection between conceptual understanding and problem-solving skills has not been as well studied.^{2,3} Active-learning instructional strategies can promote conceptual understanding, as measured by the Force Concept Inventory⁴ (FCI) and the Force and Motion Conceptual Evaluation⁵ (FMCE). But what happens to problem-solving skills as conceptual understanding is increasingly emphasized? Are we sacrificing problem-solving skill development to make gains in conceptual understanding?

In many cases, active-learning modes are adopted by an entire course so that there is no longer a comparison group using traditional instruction. At California Polytechnic State University (Cal Poly), we teach many simultaneous sections of the same course and are able to employ both active-learning and traditional teaching modes during the same term. Therefore, we are able to compare our active-learning sections directly to traditional sections with regard to conceptual learning and problem-solving ability.

In this paper we will present the results from comparisons done over three separate quarters in the first-quarter course of a yearlong, introductory calculus-based physics sequence. We will begin by describing the student populations and the active-learning course we have developed at Cal Poly. Then we will discuss the tools we use to compare the two instructional modes, present the results, and draw some conclusions.

II. THE STUDENTS AND THE COURSES

A. Student characteristics

The students taking introductory calculus-based physics at Cal Poly are predominately engineering students, but a small

subset are science and mathematics students. Students are given no information about the distinction between studio and traditional modes of instruction at registration. Most students enroll in either a studio or a traditional section based on schedule preferences, section availability, and scheduling priority. The student populations are not preselected in any fashion.

We have examined the composition of the student groups in terms of their academic background. Table I shows high school grade-point averages and SAT scores for the students involved in the present study. It is clear from Table I that there are no statistically significant differences between the studio and traditional student populations. The male/female ratio also was almost identical in all groups.

B. Studio course

The Cal Poly studio classroom opened in the winter quarter of 1998. The primary goals of this environment are to eliminate the boundary between lecture and laboratory and to promote active-learning instruction. The physical layout is based loosely on the Rensselaer model.⁶ The studio classroom can accommodate 48 students, but the class size has usually been held to about 40. There is a computer for each pair of students, although space limitations require students to work in groups of four on some of the experiments. The computers are used for classroom activities which include RealTime Physics experiments,^{7,8} Interactive Lecture Demonstrations,⁹ some VideoPoint¹⁰ analyses of motion movies, and similar activities that we have developed at Cal Poly. These activities are all grounded in physics education research and designed to promote conceptual understanding and the laboratory skills necessary to reach conclusions based on experimental data. The computers are not used for simulations or spreadsheets.

The studio classes have minimal lecture. Classes meet in two-hour blocks three times a week for a total of six hours.

Table I. High school grade point averages, SAT-Math scores, and SAT-Verbal scores for students in each of the three quarters. The numbers in parentheses in the first column are the number of students in each group. The uncertainties are the standard deviation of the mean (standard error). None of the differences are statistically significant.

	H.S. GPA	SAT-Math	SAT-Verbal
<i>Fall 1998</i>			
Traditional (105)	3.73 (± 0.05)	635 (± 7)	580 (± 7)
Studio (120)	3.76 (± 0.04)	634 (± 7)	569 (± 7)
<i>Winter 1999</i>			
Traditional (39)	3.86 (± 0.05)	663 (± 10)	565 (± 12)
Studio (147)	3.77 (± 0.03)	648 (± 6)	581 (± 6)
<i>Spring 2000</i>			
Traditional (103)	3.73 (± 0.03)	628 (± 7)	563 (± 8)
Studio (71)	3.72 (± 0.05)	636 (± 9)	581 (± 10)

Class time is primarily divided between computer-based activities including experiments making use of computer-based data collection and small-group work. The group work consists of guided discussion, pencil-and-paper exercises, and practice problem solving. Problem solving is typically done in groups of two to four students and uses a structured approach (Van Heuvelen-type worksheets),¹¹ but not formal collaborative groups.¹² Individual instructors are free to mix and match the various components as they see fit. Instructor-focused activities are limited to summaries and some example problem solving. One or two undergraduate assistants are available to interact with the students. The assistants receive some training with the RealTime Physics labs, but their ability to recognize and respond to student misconceptions is minimal.

The studio course uses the textbook *Physics: A Contemporary Perspective* by Knight,¹³ a research-based textbook designed to support an active learning environment. Because there is comparatively little lecture, students must acquire basic information through reading the textbook. This reading is strongly encouraged by almost daily collection of exercises from the student workbook that accompanies the text.¹⁴

C. Traditional course

The traditional classes have a clear distinction between lecture and laboratory. Classes of 35 to 40 students also meet six hours per week for three hours of formal lecture and one three-hour recitation/laboratory. The recitation/laboratory instructor is usually not the same as the lecture instructor. The weekly recitation/laboratory session is loosely broken up into two parts. A recitation (about 45 minutes) for problem solving and homework discussion begins the period. The remainder of the three hours is a standard measurement-oriented experiment using a locally written lab manual.

The textbook for the traditional sections was in flux throughout this period, and a different text was used during each of the quarters discussed here. *Physics* by Resnick, Halliday, and Krane¹⁵ was used in Fall 1998; *Fundamentals of Physics* by Halliday, Resnick, and Walker¹⁶ in Winter 1999; and *University Physics* by Benson¹⁷ in Spring 2000.

The prerequisite for both the studio and the traditional course is one quarter of calculus. Both cover the same topics: kinematics, Newton's laws, conservation laws, and an introduction to rotational kinematics and dynamics. The studio

sections spend approximately 10% more time on kinematics and Newton's laws, and somewhat less time on rotational motion.

III. COMPARISON TOOLS

A. Measures of conceptual understanding

Conceptual understanding was measured with either the Force Concept Inventory (Fall 1998) or the Force and Motion Conceptual Evaluation (Winter 1999 and Spring 2000). We will not attempt to evaluate how well these tests measure conceptual knowledge or discuss any differences between the two.¹⁸ We note that the questions require conceptual knowledge in order to answer them correctly and are the type that we hope students would be able to answer correctly with one quarter of instruction; the tests give similar results. We will report both preinstruction and postinstruction results as well as the normalized gain.

The normalized gain $\langle g \rangle$ is the ratio of the actual gain to the maximum possible gain, $\langle g \rangle = (\text{post-pre}) / (100 - \text{pre})$. This measure accounts for differences in the initial starting knowledge of students so that different classes can be compared directly. In his 6000 student survey of the Force Concept Inventory, Hake¹⁹ found that traditional-instruction classes had normalized gains in the range 0.19–0.27 and that interactive engagement classes had normalized gains in the range 0.34–0.62. For the Force and Motion Conceptual Evaluation, traditional-instruction classes had gains in the range 0.11–0.21, and classes using the RealTime Physics materials had gains in the range 0.38–0.66.²⁰

B. Measures of problem-solving ability

Quantitative problem-solving ability was measured with four or five common problems on the final exams of both studio and traditional sections. Most problems required two or more pieces of content knowledge, such as both Newton's laws and kinematics or both energy and momentum conservation. None required the graphical interpretation skills or conceptual knowledge that is emphasized in the studio sections but not in the traditional sections. All would be considered standard problems that we hope students would be able to solve at the end of one quarter of physics. Representative problems from two different quarters are shown in Fig. 1.

In Fall 1998, the exam problems were written by instructors who were not associated with the studio and who were not teaching the course that quarter. In other quarters, the exam problems were written jointly by the studio and traditional instructors. Each final exam paper was identified by a code number, not by student name or section. The exams from all of the sections were shuffled together to randomize them, then each exam problem for all students was graded by a single instructor.

IV. RESULTS

For each of the comparison quarters, Table II shows the preinstruction score, postinstruction score, and normalized gain for the relevant conceptual measure. The last column shows the score on the quantitative final exam problems. The data sets are matched. That is, each student had pre- and postinstruction conceptual assessment and final exam scores. The uncertainty shown with each score is the standard deviation of the mean, σ / \sqrt{N} , also called the standard error. This

Problem A

Frank starts from rest and pushes a 10 kg crate up a ramp that is tilted at a 15° angle. The coefficient of kinetic friction is 0.20. Frank pushes the crate with a horizontal force P , parallel to the floor. If Frank pushes on the crate with a constant 55 N, how long does it take the crate to reach a speed of 3.0 m/s?

Problem B

A 5-gram bullet moving with an unknown initial speed is fired into and passes through a 2-kg wood block. The exit speed of the bullet is 100 m/s. The wood block is initially at rest on the rough (coefficient of friction equal to 0.20) horizontal surface. In addition, the block is connected to a spring of force constant 500 N/m. The block moves a distance of 10 cm to the right after impact (before oscillating back to the left). We are only concerned with examining the system from before impact to the maximum compression after impact.

- Find the increase in the potential energy for the system.
- Calculate the work done by friction on the block.
- Find the speed of the block right after impact (before compression of the spring).
- Find the original speed of the bullet before impact.

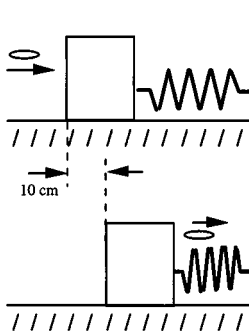


Fig. 1. Sample problems from final exams of two different quarters.

quantity is the appropriate measure for comparing the means of two groups. There are no differences on the preinstruction score between the studio classes and the traditional classes.

A. Fall 1998

In the Fall 1998 quarter, we attempted to minimize any possible influence of the individual instructors by having the same three instructors teach both a studio section and a traditional section. Two of the three instructors had previously taught in the studio. The FCI posttest was given as part of the final exam.

The results are shown in the top section of Table II. There are several noteworthy aspects of this data. Students in the studio sections scored significantly higher on the conceptual test, with nearly twice the normalized gain, and scored an average 5 points higher on the quantitative final exam problems. This latter difference is not statistically significant.

The normalized gain for the studio sections is consistent with other active-learning courses.¹⁹ However, the gain is unusually high for the traditional sections. This high gain may be because the instructors of the traditional sections also taught studio sections, so there could have been some carry-over from the studio courses, such as an increased emphasis on concepts. Or, because the FCI was given for credit as part

of the final exam, perhaps the students took it more seriously and did correspondingly better. However, we have not seen this effect repeated and so we believe it may be an anomaly.

B. Winter 1999

There were five studio sections and three traditional sections. The instructors of the traditional sections were a subset of the instructors in the studio and had all previously taught in the studio. The FMCE posttest was given during the last meeting of the class or lab and did not count toward the students' grades. The five instructors jointly wrote the common final exam problems.

The results for this quarter are shown in Table II. In this case, the traditional sections scored an average 5 points higher on the quantitative final exam problems. This difference of approximately 10% is not statistically significant. The students in the studio sections again scored much higher on the post-FMCE. The difference in normalized gain is approximately 200%. These conceptual learning gains for both the traditional and studio sections are consistent with the results of other studies.¹⁹

Table II. Student performance on conceptual understanding and quantitative problem solving. The Force Concept Inventory (FCI) was used to measure conceptual understanding in Fall 1998. The Force and Motion Conceptual Evaluation (FMCE) was used in Winter 1999 and Spring 2000. The uncertainties are the standard deviation of the mean.

	No. of sections	No. of students	Pre-FCI (%)	Post-FCI (%)	Normalized gain	Problem solving (%)
<i>Fall 1998</i>						
Traditional	3	105	47.2 ± 2.0	68.0 ± 2.1	0.39 ± 0.04	57.6 ± 1.3
Studio	3	120	48.9 ± 1.4	79.0 ± 1.4	0.60 ± 0.03	62.6 ± 1.8
			Pre-FMCE (%)	Post-FMCE (%)		
<i>Winter 1999</i>						
Traditional	3	39	33.5 ± 4.0	49.0 ± 3.8	0.23 ± 0.08	70.9 ± 2.1
Studio	5	147	33.0 ± 2.0	76.9 ± 1.8	0.65 ± 0.04	65.4 ± 1.4
<i>Spring 2000</i>						
Traditional	4	103	25.9 ± 1.9	41.0 ± 2.6	0.20 ± 0.04	72.7 ± 1.9
Studio	3	71	27.8 ± 2.6	75.8 ± 2.6	0.66 ± 0.07	60.7 ± 2.8

C. Spring 2000

The instructors for the studio and traditional sections were different. The instructors in the traditional sections had never taught in the studio and had no previous experience with active-learning techniques. The FMCE posttest was again given during the last meeting of class or lab. The traditional sections covered more material, through rotational dynamics, while the studio sections only introduced rotational kinematics. The common final exam problems (without rotational motion) were written by all the instructors involved.

The results are shown in Table II. This time, students in the traditional sections scored an average 12 points higher on the quantitative problems than students in the studio sections. This difference of approximately 20% is statistically significant. However, the studio sections had a normalized gain on FMCE that was more than three times that of the traditional sections. The normalized gains for the traditional and studio sections were again consistent with previous work.¹⁹

V. CONCLUSIONS

We have reported data for three different quarters in which we have compared studio and traditional instruction. Although there are significant variations in course characteristics from quarter to quarter, we believe we can identify some important trends.

The first is that the studio format promotes large conceptual learning gains. Using the FMCE, the average normalized gain was 0.65 for the studio sections and 0.21 for the traditional sections, a roughly 200% improvement in the studio sections. This result is consistent with other quarters, not reported here, when a common final exam was not given.

Second, any difference in quantitative problem-solving ability is much smaller than the difference in conceptual learning. The difference between studio sections and traditional sections was not statistically significant in two of the three comparison quarters. The 12 point difference in Spring 2000, a loss of roughly 20% in the studio sections, is statistically significant. Although we cannot with certainty pinpoint the reason why one quarter in three revealed a gap in problem-solving ability, we note that Spring 2000 was the least controlled of the three comparison quarters. Taken together, the three comparison quarters imply that students in the studio sections have, at most, a slight decrease in problem-solving ability.

From these results, we conclude that students must be taught both concepts and problem-solving skills explicitly if we want students to be proficient at both. An implicit assumption in many traditional-format classes is that students will automatically acquire conceptual understanding in the process of learning to solve physics problems. Our results, in agreement with other studies, show that this is not the case. More importantly, we have shown that a course in which concepts are explicitly emphasized produces large gains in conceptual understanding without significantly sacrificing students' quantitative problem-solving ability.

ACKNOWLEDGMENTS

We thank Cal Poly colleagues J. Boone, R. Echols, R. Saenz, N. Sungar, L. Wall, W. Wilson, and R. Zammit for their participation, President W. Baker for his enthusiasm, and Dean P. Bailey of the College of Science and Math for support and release time. Professor M. Wittmann of the Uni-

versity of Maine helped analyze the FMCE data. We also received support from Department of Education Grant No. FIPSE #116P980003, administered by Professor D. Sokoloff at the University of Oregon.

¹Lillian C. McDermott and Edward F. Redish, "Resource Letter: PER-1: Physics Education Research," *Am. J. Phys.* **67**, 755–767 (1999).

²R. N. Steinberg and M. S. Sabella, "Performance on multiple-choice diagnostics and complementary exam problems," *Phys. Teach.* **35**, 150–155 (1997).

³B. Thacker, E. Kim, K. Trefz, and S. M. Lea, "Comparing problem solving performance of physics students in inquiry-based and traditional introductory physics courses," *Am. J. Phys.* **62**, 627–633 (1994).

⁴D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach.* **30**, 141–158 (1992); the most recent information is at (<http://modeling.asu.edu/modeling.html>).

⁵R. K. Thornton and D. 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**, 338–352 (1998).

⁶Karen Cummings, Jeffrey Marx, Ronald Thornton, and Dennis Kuhl, "Evaluating innovation in studio physics," *Am. J. Phys.* **67**, S1, S38–S44 (1999).

⁷We use the Universal Laboratory Interface along with Motion Detectors and Force Probes from Vernier Software & Technology, (www.vernier.com); we also used Rotary Motion Sensors (CI-6625) from Pasco Scientific, (www.pasco.com).

⁸David R. Sokoloff, Ronald K. Thornton, and Priscilla W. Laws, *Real Time Physics, Module 1: Mechanics* (J Wiley, New York, 1999). The students did Lab 1: 1, 2, 3-1, 4; Lab 2: 1, 3; Lab 3: 1, 2; Lab 4: 1; Lab 5: 1, 2-1; Lab 8: 1, 2; Lab 11: 1, 2; and Lab 12: 1, 2. An entire lab was typically spread out with one or two investigations per class meeting. We did not assign any of the prelab preparation sheets and for each lab the corresponding homework was assigned and collected. The materials were slightly modified to account for local experimental arrangements.

⁹Tools for Scientific Thinking—Interactive Lecture Demonstrations (Vernier Software, Beaverton, OR). We did ILD #3.

¹⁰*VideoPoint* software is available from Lenox Softworks, (www.lsw.com).

¹¹A. Van Heuvelen, "Overview, Case Study Physics," *Am. J. Phys.* **59**, 898–907 (1991).

¹²The use of structured group work is discussed in P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1. Group versus individual problem solving," *Am. J. Phys.* **60**, 627–636 (1992); P. Heller and M. Hollabaugh, "Teaching problem solving through cooperative grouping. Part 2. Designing problems and structuring groups," *Am. J. Phys.* **60**, 637–644 (1992).

¹³Randall D. Knight, *Physics: A Contemporary Perspective* (Addison-Wesley, New York, 1997), preliminary ed., Vol. I.

¹⁴Randall D. Knight, *Student Workbook to Accompany Physics: A Contemporary Perspective* (Addison-Wesley, New York, 1997), preliminary ed., Vol. I.

¹⁵Robert Resnick, David Halliday, and Kenneth S. Krane, *Physics* 4th ed. (J Wiley, New York, 1992).

¹⁶David Halliday, Robert Resnick, and Jearl Walker, *Fundamentals of Physics* 5th ed. (J Wiley, New York, 1997).

¹⁷Harris Benson, *University Physics* (J Wiley, New York, 1996), revised ed.

¹⁸For discussion about the FCI see Charles Henderson, "Common concerns about the Force Concept Inventory," *Phys. Teach.* **40**, 542–547 (2002); Richard Steinberg and Mel Sabella, "Performance on multiple-choice diagnostics and complementary exam problems," *Phys. Teach.* **35**, 150–155 (1997); Robert C. Hudson and Frank Munley, "Re-score the force concept inventory!," *Phys. Teach.* **34**, 261–267 (1996); Pat Heller and Doug Huffman, "Interpreting the Force Concept Inventory: A reply to Hestenes and Halloun," *Phys. Teach.* **33**, 502 (1995); David Hestenes and Ibrahim Halloun, "Interpreting the Force Concept Inventory: A response to March 1995 critique by Huffman and Heller," *Phys. Teach.* **33**, 503 (1995); Doug Huffman and Pat Heller, "What does the Force Concept Inventory actually measure?" *Phys. Teach.* **33**, 138–143 (1995).

¹⁹Richard 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).

²⁰Michael Wittman, private communication.