Eurasia Journal of Mathematics, Science & Technology Education, 2007, **3**(3), 231-237

Evaluation Novelty in Modeling-Based and Interactive Engagement Instruction

Funda Örnek
Balıkesir Üniversitesi, Balıkesir, TURKEY

Received 10 July 2006; accepted 07 June 2007

A calculus-based introductory physics course, which is based on the Matter and Interactions curriculum of Chabay and Sherwood (2002), has been taught at Purdue University. Characteristic of this course is its emphasis on modeling. Therefore, I would like to investigate the effects of modeling-based instruction and interactive engagement on students' physics understanding. For this reason, The Force Concept Inventory (FCI) (Hake, 1992) as pre-and post-test was used to evaluate students learning and understanding following a newly developed approach to teaching mechanics in an introductory physics course. The results lead that it can be concluded that the modeling-based interactive teaching method helps students to improve their understanding and learning physics.

Keywords: Conceptual Understanding, FCI, Learning, Teaching, Physics Models.

INTRODUCTION

This paper provides an explanation of development of modeling-based interactive engagement teaching approach to teaching mechanics in calculus-based introductory physics course. Also, it demonstrates how the Force Concept Inventory (FCI) was used not only to evaluate this new teaching approach but also to analyze the student learning (Savinalnen, A. & Scott, P., 2002). The teaching was done by an instructor in department physics at Purdue University.

In USA and elsewhere, research results indicate that active-based teaching can lead to improvement of gains in student learning and understanding. I believe that this study can contribute to literature in physics education by "*demonstrating how insight into research on teaching and learning in physics have been drawn upon to revise instructional approaches and to thereby improve student learning*" (Savinalnen & Scott, 2002).

Correspondence to: Funda Örnek, PhD., Teaching Assistant, Balıkesir Üniversitesi, Necatibey Eğitim Fakültesi, Fizik Eğitimi Bölümü, Balıkesir, 10100 TURKEY.
E-mail: fundaornek@gmail.com

E-ISSN: 1305-8223

Structure of the Course

The Purdue physics course is a two-semester introductory physics sequence for physics majors. The course, PHYS 162, which treats Particles, Kinematics, and Conservation Laws, is taught in the Fall semester. PHYS 163, which treats Mechanics, Heat, and Kinetic Theory, is taught in the Spring semester.

Physics modeling takes place in each section of the course. Students used a few physics principles and approximations to construct their models to solve problems.

The structure of the course is different than many other physics courses. During the Fall semester, PHYS 162 consists of two lecture sessions, either small-group work or computer-laboratory sections, and workshops in a computer laboratory. Whether the small-group work or computer laboratory will be held is decided by the instructor.

Lectures meet on Mondays, and Wednesdays. During lectures, students are actively involved in their learning. Students interact with each other and with the instructor instead of sitting, listening, watching the instructor, and taking notes. In addition, the instructor performs hands-on experiments.

Small-group work, which is called "recitation" in all traditional physics courses, meets on Tuesdays, and

Thursdays. It has three sections which meet on the same day. Each section has about 24 students and is divided into 8 small groups. A traditional recitation is run by a teaching assistant solving problems in front of the class, whereas the small-group work sections in PHYS 162 are run by the instructor, a teaching assistant, and a student helper who has already taken these courses. Each small group has a small white board on which to solve physics problems. After they solve the physics problems, they share their solutions with the class by presenting their solutions. The purpose is to have students be actively involved. Teaching assistants, instructor, and student helpers are the facilitators.

The computer-laboratory session has three sections as does the small-group work session. All computer sections are scheduled at the same time that the small-group work sections meet. The instructor decides when they will have the computer laboratory or the small-group work. Students always stay in their section of the small-group section. Each student has a computer which he/she can use and write his/her own simulation program. They use a computer program which is called VPython. Again, the instructor, a teaching assistant, and a student helper are present in each computer-laboratory section.

Workshops are held in the same computer laboratory on Fridays to help students with their difficulties understanding the content covered during classes. These workshops are problem-solving and help sessions. Also, they are for students to catch up. There are three sections in a day as well. In each workshop section, the instructor, and a teaching assistant are present. Moreover, not only the instructor, but also the teaching assistants hold office hours for students.

During the Spring semester, everything is the same except for an additional lecture per week and student helpers (they are not available during the Spring semester). Lectures meet on Mondays, Wednesdays, and Fridays at the same time as in the Fall semester.

There are three 1-hour exams and a 2-hour final exam for each course. In addition, students are supposed to do homework, computer problems and daily quizzes. Daily quizzes, which happen all semesters, are given in lecture to identify whether students understand the concepts, and also for attendance, for which credit is given.

Since the course, PHYS 162, which treats Particles, Kinematics, and Conservation Laws, is taught in the Fall semester, covers force concepts, I only used the FCI results in the Fall semester.

Teaching method: Modeling-Based Interactive Engagement

Beginning the Fall Semester of 2001, the Physics Department at Purdue University started to teach a calculus-based introductory physics course by using a modeling-based interactive engagement method. To be precise, "modeling" as used here has a different meaning from "modeling" used in the notation of science education. In brief, modeling in physics is defined as "making a simplified, idealized physics model of a messy real-world situation by means of approximations" (Chabay & Sherwood, 1999). It is also called "physics modeling" in physics education community. In this course, physics modeling and computer simulations are used to promote conceptual understanding along with interactive engagement method. Hake (1998) defines "Interactive Engagement (IE) methods as those designed at least in part to promote conceptual understanding through engagement of students in heads-on (always) and hands-on (usually) activities which yield immediate feedback through discussion with peers and/or instructors..." (p.65) In other words, it is a method that improves students' conceptual understanding by their interactions with each other using their thoughts and some hands-on activities. Then, they can have immediate feedback from their discussions with their peers, their teaching assistants, or their instructors.

Modeling-based interactive engagement instruction entails some features which focus on the development of conceptual understanding:

Physics Modeling

Modeling means something different to physicists. A physics model in the physics-education community is considered as a simplified and idealized physical system, phenomenon, or idealization. According to Greca & Moreira (2001), the physics models determine, for instance, the simplifications, the connections, and the necessary constraints. As an example one can think of a point particle model of a system in classical mechanics. A simple pendulum is another example of a physics model because it is idealized and consists of a mass particle on a massless string of invariant length moving in the homogenous gravitational field of the Earth in the absence of drag due to air (Czudkova & Musilova, 2000).

In Purdue's calculus-based introductory physics courses, students do not use models which are already created in this course. They apply the fundamental principles and create their own models. Modeling involves making a simplified, idealized physics model of a messy real-world situation by means of approximations. Then, the results or predictions of the model are compared with the actual system. The final stage is to refine the model to obtain better agreement, if needed. Sometimes it may not be needed to modify the model to get more exact agreement with the real-world phenomena. Even though the agreement may be excellent, it will never be exact since there are always some influences in the environment that we cannot consider while we are building the models. For instance, while a rock is falling, the gravitational pull of the earth and air resistance are the main influences. However, there are also other effects such as humidity, wind and weather, Earth's rotation, even other planets (Chabay & Sherwood, 1999).

In physics modeling (Chabay and Sherwood, 1999), the following process is followed:

- Start from fundamental principles which are the linear momentum principle, the energy principle, and the angular momentum principle
- Estimate quantities
- Make assumptions and approximations
- Decide how to model the system
- Explain / predict a real physical phenomenon in the system
- Evaluate the explanation or prediction

In summary, physics modeling is analysis of complex physical systems by means of making conscious approximations, simplifications, and idealizations. When students make approximations or simplifications, they should be able to explain why they make them. For instance, in modeling a falling ball, in general, air resistance is neglected. So, there is no force contribution from air resistance. While students do neglect it, they should be able to have reasons for this. As an example of modeling, consider the calculation of the acceleration of a block is pulled to the right with a force F as shown in Figure 1.

To analyze this system, we should start with the momentum principle,

$$\frac{d\vec{p}}{dt} = \vec{F}_{net}.$$

Because of friction between the table and the block, there is frictional force, *f* in addition to the force, F; pulling the block. So, the total force is

$$F_{net} = F - f.$$

From the momentum principle,

$$\frac{d\vec{p}}{dt} = \vec{F}_{net} = F - f.$$

So,

$$\frac{d\vec{p}}{dt} = \frac{d(mv)}{dt} = m\frac{dv}{dt} = ma = F - f$$

from this, it can be concluded that the block moves with a constant acceleration which is given as

$$a = \frac{F - f}{m}.$$

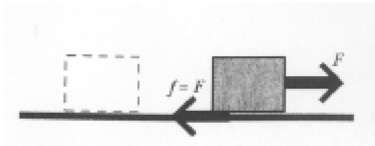

Figure 1. Pulling a Block (Chabay & Sherwood, 2002).

Computer Simulations

In this course, students write programs to simulate physical systems using VPython (Scherer et al., 2000). VPython makes students focus on the physics computations to obtain 3-D visualizations. Students can do true vector computations, which improves their understanding of the utility of vectors and vector notation. For example, students can study the motion of the earth in orbit around the sun by means of writing a program.

Creating simulations by writing computer program using VPython helps students understand physics because they can see how physics principles work.

The modeling-based interactive engagement method defined by Chabay & Sherwood (2002) can offer the potential to promote enhanced learning in conceptual understanding of physics.

RESEARCH METHOD

Research Context and Subjects

This study took place in the Physics Department at Purdue University throughout the Fall, 2004, semester. For this study, I focused on a calculus-based introductory-level physics course which includes lecture, small-group work, which is called "recitation" in a traditional physics course, and computer simulations. Only 16 students completed the pre-test and post-test administration of the FCI.

Theoretical Framework

There are several kinds of research designs which guide quantitative research in education. For the purpose of this study to answer my research questions I used a pre-experimental research design, the one-group pretest-posttest design. A single group is often studied, but no comparison between an equivalent non-treatment group is made. In this design, as shown in Figure 2, a single group of subjects is given a pretest, then the treatment, and then a posttest. The pretest and posttest are the same. The result that is examined is the change from pretest to posttest (McMillan & Schumacher, 2001). In my study, O (observation) meant that pretests and posttest were administered before and after treatment of the modeling-based interactive engagement instruction. X (treatment) was the modeling-based interactive engagement instruction.

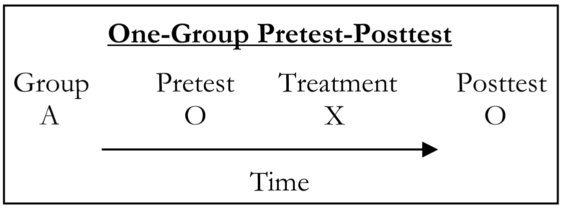

Figure 2. One-Group Pretest and Posttest (McMillan & Schumacher, 2001).

To use this design in my study, I compared pretest results prior to the modeling-based interactive engagement teaching method to the results after completing a semester of experience in this teaching method. I could at least state whether a change in the test results had taken place. What I cannot say is if this change would have occurred even without the application of the treatment. It is possible that any instructional method or mere maturation could have caused the change in grades and not the modeling-based and interactive engagement teaching method itself.

Data Collection

I used the Force Concept Inventory (FCI) which was administered as pre- and post-tests in the fall semester in Phys 162, to determine if the modeling-based interactive engagement teaching would have a significant effect on conceptual learning and understanding of physics. The Force Concept Inventory (Hestenes et al., 1992) is a 29-question multiple-choice test for measuring students' understanding of Newton's Laws. It probes the belief systems students hold concerning force, the primary concept of Newtonian mechanics (Churukian, 2002).

RESULTS

I used the Force Concept Inventory (FCI) which was administered as pre- and post-tests in the fall semester in Phys 162, to determine if the modeling-based interactive engagement teaching would have a significant effect on conceptual learning and understanding of physics. Only 16 students completed the pre-test and post-test administration of the FCI.

To compare students' pre-test scores and post-test scores, I used the Binomial sampling distribution (Blalock, 1960). Before going further, I would like to give some information about the null hypothesis because it is based on median difference instead of mean difference due to skewness (Blalock, 1960). In general:

H_0 (Null hypothesis): Population median difference is zero.

H_a (Alternative hypothesis): Population median difference is not zero.

The null and alternative hypotheses in my study are:

H_0: The modeling-based interactive engagement teaching method has no effect on students' pre-test and post-test FCI scores.

H_a: The modeling-based interactive engagement teaching method has an effect on students' pre-test and post-test FCI scores.

In order to test H_0 hypothesis, I utilized the Sign test which is the application of the Binomial sampling distribution and for a small number of cases (Blalock, 1960). For success, a + sign indicates cases where students' FCI score is increased, for failure, a - sign indicates cases where students' FCI score is decreased. If there are any students who show no change, these students are excluded from the analysis. Assuming that there are equal chances of pluses and minuses, the probability of getting a + sign in any given draw is 0.5 under the null hypothesis (Blalock, 1960).

As $\alpha = 0.05$ level of significance using a two-tailed test; I chose the critical region for which I can reject the null hypothesis. The calculated p-value which determines whether I can reject the null hypothesis or not is 0.022 using a two-tailed test. The probability value (p-value) of a statistical hypothesis test is the probability of getting a value of the test statistic by chance. The p-value is compared with the significance level and, if it is smaller, the result is significant. For example, if null hypothesis can be rejected at $\alpha = 0.05$, this can be reported as $p < 0.05$ (Lomax, 2001).

According to the result of the Binomial distribution obtained from the SPSS, the small p-value, which is 0.022, demonstrates the null hypothesis of "there is no effect on pre-test and post-test FCI scores" is not supported. Why is the null hypothesis rejected? If p-value is less than significance level (α) then H_0 can be rejected in favor of H_a. H_0 is rejected because the p-value = $0.022 < \alpha = 0.5$. Therefore, the result shows a significant change in students' FCI scores. In other words, the modeling-based interactive engagement teaching appears to have an affect on students' pre-test and post-test FCI scores.

Using the Sign test, I determined whether the effect on students' scores was positive or not. I looked at whether each student exhibited any increase or decrease in his/her FCI scores. The results as shown Table 4 are:

- 11 students exhibited increased their score on the post-test relative to the pre-test.
- 2 students showed a decrease in their scores.
- 3 students showed no change in scores.

I used the Wilcoxon Matched-Pairs Signed-Ranks Test (the T test) to compare pre-test and post-test scores and find out if there is significant difference or not. The Wilcoxon test includes the same null hypothesis as used in the t test for paired samples. Blalock (1960) explains how to use a T test:

> … The null hypothesis states that there are no differences between the scores of the two populations. After obtaining difference scores for each pair, these differences are ranked regardless of the sign. The sums of the ranks of both the positive and negative differences are obtained. If the null hypothesis is correct, then the sum of the ranks of the positive differences will be approximately the same as the sum of the ranks of the negative differences. If these sums are quite different in magnitude, the null hypothesis may be rejected. We form the statistics T which is the smaller of these two sums is used. T is obtained by adding the negative signs (negative differences). A T table for N ≤ 25 gives the critical values of T. T should be equal to or less than the values in the T table at the 0.05 level to reject the null hypothesis (p.266-267).

The procedure to calculate rank of difference in Table 1:

- Calculate the differences between pairs.
- Do not count zeros (if the difference is "zero").
- Rank the differences in increasing order according to magnitude of difference (without regard to sign); if multiple observed differences have the same value, "split" the ranks according to the number of observations (for example, the 9th and 11th observations have the same absolute value, 1, then each observation is ranked as 1.5).

Table 2 summarizes the differences between pre-test and post-test scores.

The sum of negative ranks, T, is 10.50. Blalock's T test table, the critical value of T for two-tailed test for N=16 at α = 0.05 is 30. Since T=10.50 is smaller than $T_{crit} = 30$, the null hypothesis – there are no differences between the scores of the two populations – can be rejected. The other way to show any significant difference is to use p-value, 0.014 from the Wilcoxon Signed Ranks test. Since p = 0.014 < α = 0.05, there are statistically significant differences between the scores of pre-test and post-test with the post-test scores being higher. In other words, the modeling-based interactive teaching method appears to help students to improve their scores from pre-test to post-test.

In addition to determining whether the modeling-based interactive engagement teaching method has an affect on understanding and learning of physics as indicted by a difference between the scores of pre-test

Table 1.Computations for Wilcoxon Matched-Pairs Test and the Sign Test.

Pair Number	Pre-Test	Post-Test	Difference Between Post- and Pre-Tests	Rank of Difference	Positive Ranks	Negative Ranks
1	27	27	0	0		
2	29	27	-2	4		4
3	13	13	0	0		
4	13	23	+10	13	13	
5	27	27	0	0		
6	23	29	+6	9	9	
7	19	28	+9	11.5	11.5	
8	23	20	-3	6.5		
9	26	27	+1	1.5	1.5	6.5
10	21	29	+8	10	10	
11	24	25	+1	1.5	1.5	
12	18	27	+9	11.5	11.5	
13	21	25	+4	8	8	
14	27	29	+2	4	4	
15	17	20	+3	6.5	6.5	
16	18	20	+2	4	**4**	
Total					**80.50**	**10.50**

Table 2. Wilcoxon Signed Ranks.

		N	Mean Rank	Sum of Ranks
Posttest score-pretest score	Negative Ranks	2(a)	5.25	10.50
	Positive Ranks	11(b)	7.32	80.50
	Ties	3(c)		
	Total	16		

Note: a posttest < pretest
b posttest > pretest
c posttest = pretest

and post-test, I wanted to see if there was any correlation between the pre-test scores and the first exam grades, and the post-test scores and the final exam scores (there were three exams and one final exam). I used Kendall's Tau correlation coefficient r_b which measures the strength of the relationships between two variables (Blalock, 1960). Kendall's Tau r_b is calculated using the following equation:

$$r_b = \frac{S}{\sqrt{1/2N(N-1)-T}\ \sqrt{1/2N(N-1)-U}}$$

where S is equal to C-D. C is the number of concordant pairs (a given pair is ordered the same way).

D is the number of discordant pairs (a given pair is ordered oppositely)

$T = 1/2\sum t_i(t_i - 1)$, t_i

is the number of ties in each set of ties in the first group.

$U = 1/2\sum u_i(u_i - 1)$, u_i

is the number of ties in each set of ties in the second group.

In this study; let's say the first group is post-test scores and the second group is the final exam scores. Here is an example which explains that ties can be any number which is repetitive such as 27 repeats on pre-test scores or post-test scores. N is the number of cases (N=16).

In Table 3, if every time a given pair such as 1 and 2 is ordered the same way, then +1 is used for C. In other words, each pair increases or decreases. For example, pairs 2 and 3 both decrease. Post-test scores decreased 27 to 13 and the final exam scores decrease 73.53 to 56.69 in pairs 1 and 2. Therefore, we use +1 for this pair. Whenever pairs are ordered oppositely such as pairs 9 and 10, -1 is used for D. For instance, post-test scores increased from 27 to 29, but the final exam scores decreased 89.63 to 85.08. So, we use -1 for this pair. If pairs have ties (same scores), then there is no contribution from these pairs. For example, pairs 1 and 2 have ties in post-test grades which are 27. Therefore, there is no contribution. So, when we make calculations, the contribution from ties will be zero.

To calculate S = C-D, I will just show how to calculate it. I will not calculate completely because it takes too much space. Instead, I calculated using SPSS. Let me give an example from the above table. The contribution of pairs (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)… (1, 10) ….is 0, +1, +1, 0, +1....-- -1…

C= +1+1+1…. ; D=-1….. ; S=C-D=1+1+1-1…..

I calculated the Kendall's Tau r_b using SPSS and tabulated in Table 4.

From Table 4, it is easy to see that there is low correlation between the pre-test scores and the first exam grades because $0.3 < r_b = 0.31 < 0.5$. There is high correlation between post-test scores and the final exam grades since $0.7 < r_b = 0.75 < 0.9$.

Table 3. An Example from the Study to Show How to Calculate S, and to Show Ties.

Pair Number	Post-test Scores (the First Group)	The Final Exam Scores (the Second Group)
1	27	89.73
2	27	73.53.
3	13	56.69
4	23	81.58
5	27	89.73
6	29	97.63
7	28	83.56
8	20	73.65
9	27	89.63
10	29	85.08
11	25	59.99
12	27	96.02
13	25	87.01
14	29	96.33
15	20	59.95
16	20	74.15

Table 4. Kendall's Tau correlations for the Force Concept Inventory (FCI).

	Pre-test vs. the First Exam Score	Post-test vs. the Final Exam Score
Kendall's Tau r_b	0.31	0.75

Some may suggest that the improvement in scores is due to using the same test as a pre-and post-test. The pre-test scores as an advanced organizer that focuses students' attention are ideas that follow during treatment. If there is carryover using the FCI at the beginning and end of semester it should be small. Other studies using the same pre-and post-test structure have shown only small gains. These studies show that any carryover would be small.

CONCLUSION

I wanted to find out whether an introductory physics course that uses modeling-based instruction and interactive engagement lead to better physics understanding.

The gains that students have made in increasing their learning and understanding of physics were determined by the Force Concept Inventory (FCI). The results obtained from the Binomial distribution test, the Sign test, and the Wilcoxon test indicated that there was a statistically significant improvement between students' FCI pre- and post- test scores. In other words, the modeling-based interactive engagement teaching method appears to have a positive effect on students' learning and understanding of physics. This Wilcoxon T-test also indicates that students in this course have made significant improvement compared to a traditional course. Since nonparametric statistics was used for this study due to lack of sample size, it is not possible to calculate and use the same notation for gain which are used in other studies. Therefore, I cannot compare students' gains in this course with students' gains in other universities.

In addition, the result obtained by using Kendall's Tau correlation indicated that conceptual understanding the students have when they begin the course is not related to their final exam grade. Instead, there is high correlation between post-test scores and course grades. That suggests the course grades are based primarily on conceptual understanding rather than other aspects of the course such as algorithmic problem solving skills. This indicates that the amount of conceptual gains students have made have bearing on their final grades. Given the improvement in conceptual understanding and the high correlation between their conceptual understanding and their success as measured by their course grades, one can conclude that the modeling-based interactive teaching method helps students to improve their understanding and learning physics.

Although the participant selection for this research limits the transferability of the findings to the broader population of all undergraduate physics students enrolled an introductory- physics course because it is more likely that selecting participants from an introductory course for non-physics majors or those whose physics and mathematics background are not strong could yield different results and small sample size, overall this study implies that the modeling-based interactive engagement teaching format used for calculus-based introductory physics instruction at Purdue University is successful at teaching physics.

A future research will explore that while including these populations mentioned above would be informative and add to the richness of the findings.

Acknowledgements

I would like to thank to my advisor, Prof. William R. Robinson, Prof. Mark. P. Haugan who is the instructor of PHYS 162 and 163, and Dr. Donna Enersen for their tremendous help. Also, I am thankful to the head of physics department, Prof. Andrew Hirsh, and PHYS 162 and 163 students.

REFERENCES

Blalock, H. M., Jr. (1960). *Social statistics* (2nd ed). New York: McGraw-Hill.

Chabay, R. W. & Sherwood, B. A. (2002). *Vol I: Matter and interactions: Modern mechanics*. New York: John Wiley & Sons, Inc.

Chabay, R. W. & Sherwood, B. A. (1999). Bringing atoms into first-year physics [Electronic version]. *American Journal of Physics*, 67, 1045-1050.

Churukian, A. D. (2002). *Interactive engagement in an introductory university physics course: Learning gains and perceptions* (Unpublished Doctoral dissertation, Kansas State University, 2002). Dissertation Abstracts International, 62, 1685.

Czudkovà, L. & Musilovà, J. (2000). The pendulum: A stumbling block of secondary school mechanics. *Physics Education*, 35, 428-435.

Greca, I. M. & Moreira, M. A. (2002). Mental, physical, and mathematical models in the teaching and learning of physics [Electronic version]. *Science Education*, 1, 106-121.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six thousand-student survey of mechanics test data for introductory physics courses [Electronic version]. *American Journal of Physics*, 66, 64-74.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *The Physics Teacher*, 30, 141-151.

Lomax, R. G. (2001). *Statistical concepts: A second course for education and the behavioral sciences* (2nd ed.). New Jersey: Lawrance Erlbaum.

McMillan, J. H. & Schumacher, S. (2001). *Research in education: A conceptual introduction* (5th ed.). New York: Addison Wesley Longman.

Savinalnen, A. & Scott, P. (2002). Using the force concept inventory to monitor studentlearning and to plan teaching [Electronic version]. *Physics Education*, 37, 53-58.

Scherer, D., Dubois, P., & Sherwood, B. (2000). VPython: 3D interactive scientific graphics for students. *Computing in Science and Engineering*, 82-88.

◈◇◈