CHAPTER 9: Conceptual Understanding

I. OVERVIEW

As we discussed in chapter 2, conceptual understanding plays an important role in the way experts solve complex problems. Unfortunately, the examples of interviews and qualitative exam problems in chapter 2 (Mazur[1] and Hammer[2]) as well as the surveys of concept tests discussed in chapter 4 (Hestenes *et al.*,[3] Hake,[4] Thornton and Sololoff[5]) indicate that many students have little improvement in their understanding of physics concepts after traditional lecture instruction than when they began. There are two main causes for this situation. One is the persistence of students' common sense beliefs based on the students' previous real world experiences. The second is the heavy emphasis in many traditional lecture courses on typical end-of-chapter problems that students can often solve without understanding or applying the relevant concepts (see the discussion in chapters 2 & 6). In classes like this, conceptual understanding becomes part of the hidden curriculum, a course goal that is neither explicitly stated nor adequately reinforced through grading.

However, as we saw in chapters 2 and 4, PER indicates that instruction that takes the students' common sense beliefs into account and provides a mechanism for conceptual change[6] can be effective for improving students' conceptual understanding of physics. One of the main goals of this dissertation is to see if students taught with the three researched-based curricula discussed in the previous chapter, Tutorials (TUT), Group Problem Solving (GPS), and Workshop Physics (WP), show significant improvement in students' conceptual understanding of physics. In this chapter, we will look at two aspects of students' conceptual understanding: how well students learn the basic concepts as measured by concept tests and how well they use concepts and representations in solving specially-designed qualitative problems.

Research-based multiple-choice concept tests, like the Force Concept Inventory (FCI)[7] and the Force Motion Concept Evaluation (FMCE)[8], have been developed with questions that can trigger and identify students' common sense beliefs. These concept tests are a good indication of how well a student understands basic concepts in introductory physics. However, this type of understanding is only one part of the understanding needed for students to achieve a good functional understanding of physics.

Students should also be able to use their conceptual understanding in solving complex problems. In addition, they should be able to express their conceptual understanding in the multiple representations often used by expert problem solvers. The studies on physics problem solving discussed in chapter 2 indicate that expert problem solvers make heavy use of concepts and conceptual understanding in their solutions. In particular, they often use a detailed qualitative representation of the problem before applying the relevant mathematical model. This qualitative representation is then used both as a guide to solving the problem and as a means to evaluate the solution. This ability to make use of conceptual knowledge and representations in problem solving is one reason experts can respond more flexibly to new, complex problems than novices. Many students in traditional introductory physics courses lack this skill.

In section II, I present results on students' understanding of Newton's laws of motion and force for each of the three research-based curricula as well as for traditional instruction for comparison. To determine students' understanding of these basic concepts, I have gathered, processed, and analyzed concept test data for each of the ten schools. Unless otherwise specified, the data presented in section II is "matched." Only responses from students who took the tests both at the beginning and at the end of the semester or quarter of the sequence are presented in the tables and graphs. Overall results and results from specific concept clusters from the two concept tests are presented for each of the three research-based teaching methods in turn. Particular attention will be paid to questions on students' understanding of velocity-time graphs (University of Maryland only) and Newton's third law; two concepts that PER has shown to be difficult for many students to learn in traditional instruction.[9]

Section III looks at how well students use concepts and representations in solving long exam problems (as opposed to short answer or multiple choice problems). Because of the logistical difficulties discussed in the chapter 8, exam data was only available from University of Maryland. The results from four exam problems are presented.

Results from problem interviews with students are presented in section IV. The two-rock problem protocol[10] discussed in chapter 7 was used in interviews with student volunteers from Maryland, Dickinson, and Ohio State. The two-rock problem is very difficult for students. However as we discussed in chapter 7, it also provides an opportunity to see how well the student understands the concepts of velocity vectors, kinematics, and energy as well as probing their approach to physics.

II. STUDENTS' UNDERSTANDING OF BASIC CONCEPTS

A. Overall concept of Force and Motion

1. University of Maryland Tutorials

The results of pre- and post-course FCI's from tutorial and non-tutorial classes at the University of Maryland are given in Table 9-1. Sixteen first semester classes in the introductory physics sequence for engineering majors gave the FCI as pre- and post-tests. Nine were taught with tutorials, seven with recitations. Two of the instructors, C and D, taught classes with both formats. A comparison of the averages of the pre-test scores for all students taking the pre-test with the matched subset show that the matched data is consistent with the unmatched data. Therefore, the matched samples are a reasonable representation of the classes in question.

The results are displayed as a figure of merit (*h*) histogram in Fig. 9-1. Recall that *h* is the fraction of the possible gain achieved from the pre to post FCI results (*h* is described in detail in chapter 4 in the section on Hake's 6000 student study). The tutorial classes systematically produced better overall FCI gains than the non-tutorial classes. The average fractional gains of the classes are ($< h > \pm$ Std. Error)

$< h > = 0.19 \pm 0.03$ (7 classes, with recitations)
$< h > = 0.35 \pm 0.01$ (9 classes, with tutorials)

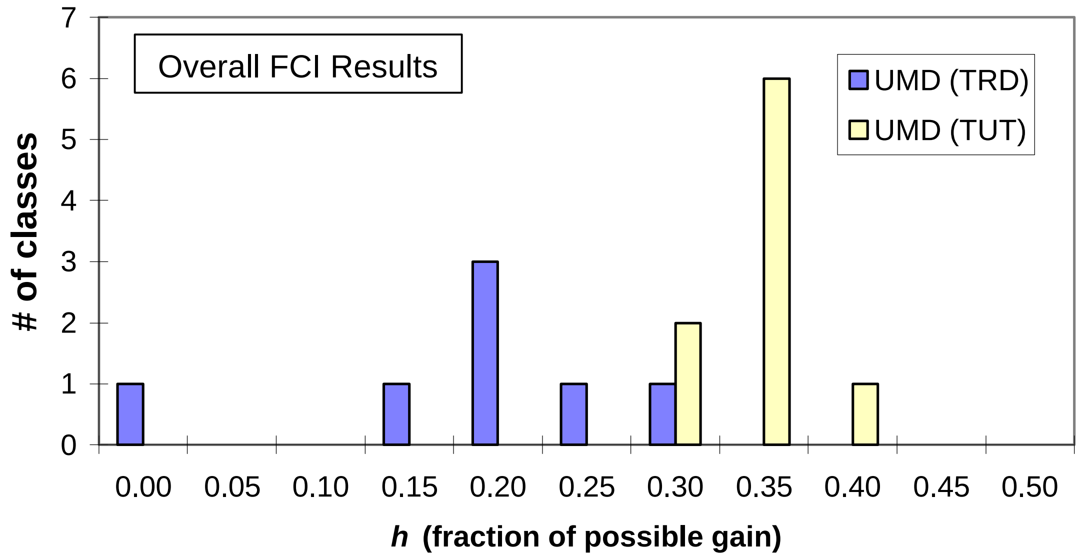
Note that this average is taken as equally weighted over lecture classes, not by students. Every tutorial class had a larger *h* than all but one of the non-tutorial classes and that was a small class taught by an award-winning lecturer.[11] However, even the tutorial results are somewhat disappointing, achieving only about 1/3 of the possible

Table 9-1: Overall FCI results for University of Maryland Traditional and Tutorial classes. Instructors are coded by letter. Numbers indicate that instructor has taught more than one class and are sequenced chronologically. *h* is defined as the fraction of the possible gain.

Traditional Lecture w/ Recitations	N	FCI Pre	FCI Post	*h*
A	42	55.4	55.9	0.01
B1	27	51.2	65.5	0.29
B2	18	60.2	70.1	0.25
C1	35	41.8	54.2	0.21
C2	39	44.5	52.9	0.15
D2	21	53.0	52.6	0.20
E	76	51.4	61.8	0.21
Average	36.9	51.1	60.4	0.19
Std. Dev.	19.5	6.3	6.3	0.09
Std. Error	7.4	2.4	2.4	0.03

Traditional Lecture w/ Tutorials	N	FCI Pre	FCI Post	*h*
C3	46	51.5	69.6	0.37
D1	55	53.9	67.8	0.30
F1	38	47.8	66.7	0.36
F2	102	54.5	72.8	0.40
G1	69	50.3	67.5	0.35
G2	65	45.3	61.1	0.29
H	59	47.6	67.2	0.37
I	24	52.7	69.4	0.35
J	88	50.3	68.9	0.37
Average	60.7	50.4	67.9	0.35
Std. Dev.	24.1	3.1	3.1	0.04
Std. Error	8.0	1.0	1.0	0.01

Figure 9-1. Overall FCI figure of merit histogram for classes at the University of Maryland. Figure of merit *h* = fraction of possible gain on the full FCI, for tutorial classes (TUT) and traditional (TRD) lecture classes.



gain. Both results, however, are consistent with Hake's findings from his study of 6000 introductory physics' students discussed in chapter 4. He found that 14 classes taught in the traditional lecture style and 48 classes taught with active engagement (research-based) curricula had the following average fractional gains:[12]

Traditional Classes	$\langle h \rangle = 0.23 \pm 0.04$ (std. dev.)
Active Engagement Classes	$\langle h \rangle = 0.48 \pm 0.14$ (std. dev.)

where *h* is averaged over classes, not students. Both the non-tutorial and the tutorial results from the Maryland classes are within one standard deviation of the respective average h-values measured by Hake. Note that the active-engagement activity in the Maryland Tutorial classes was limited to only one hour of four per week while many of Hake's active engagement classes had a much higher fraction of and total time spent on active engagement activities.

Assuming that all 16 University of Maryland classes are drawn from the same population, the probability that the difference of the means is random is less than 2% using a 2-tailed t-test with pooled variance.[13] If class A is excluded as an outlier, the probability that the difference in the means is random is less than 1%.

The same amount of instruction was offered students in both environments (3 hours of lecture and 1 hour of small class section). The primary difference between the tutorial and traditional classes is that the tutorial classes spend one hour per week on explicit concept building in a small-class group-learning-style environment, while the traditional classes spend one problem-solving hour per week in a small-class lecture-style environment.

2. FCI results at other schools

The FCI was also used as a pre/post evaluation for the first term in the introductory sequence at 7 of the 9 other schools participating in this study. The overall FCI results are shown in Table 9-2. Overall FMCE results are given for the other two schools as well as more recent WP classes at Dickinson College are shown in Table 9-3. The values for each school in the table except Ohio State are averaged over the number of classes. The Ohio State data is averaged over students. The reader is reminded that the Workshop Physics classes at Drury, Nebraska Wesleyan, & Skidmore and the Group Problem Solving classes from Minnesota & Ohio State were in the first two years of the implementation of their respective curricula.[14]

The pre-course FCI scores for the eight schools appear to represent two distributions. The pre-course averages at Drury and the three large state universities cluster around 50%. This is significantly larger (std. error ≈ 0.03) than the 40% pre-course average at the other schools. A similar difference can be seen in the pre-course FMCE data between Moorhead State and the other two schools. The pre-course average for classes at Moorhead and at Maryland is about 40% while the classes at Carroll and Dickinson start with an average around 26%. These differences do not correlate with the selectivity of the college.

The average fractional gain (*h*) for all of the research-based classes is significantly greater than the average for the traditional lecture classes at Maryland and the community college. The average *h* for all classes using research-based instruction each fall within one standard deviation of Hake's average *h* value for active-engagement

Table 9-2: Overall FCI Scores for all curricula (scores ± std. error)

University of Maryland (F93-S97)

	N (# of classes)	FCI Pre	FCI Post	*h*
Recitations	258 (7)	51.1 ± 2.4	60.4 ± 2.4	0.19 ± 0.03
Tutorials	546 (9)	50.4 ± 1.0	67.9 ± 1.0	0.35 ± 0.01

Other Traditional Lecture courses

	N (# of classes)	FCI Pre	FCI Post	*h*
PGCC (F94&F95)	40 (4)	37.2 ± 2.2	43.0 ± 4.4	0.09 ± 0.04

Group Problem Solving

	N (# of classes)	FCI Pre	FCI Post	*h*
MIN (F94)	524 (5)	49.0 ± 1.0	65.2 ± 0.7	0.32 ± 0.004
MIN (F95)	653 (5)	49.7 ± 1.1	72.9 ± 2.0	0.46 ± 0.03
OSU (F95)	258 (2)	50.4	69.4	0.38

Workshop Physics

	N (# of classes)	FCI Pre	FCI Post	*h*
DC (F92)	62 (3)	42.2	66.7	0.42
DRY (F96)	8 (1)	47.8	77.6	0.57
NWU (F95&F96)	68 (6)	38.0 ± 1.7	62.3 ± 1.9	0.39 ± 0.03
SKD (F96)	33 (2)	41.2	63.9	0.39

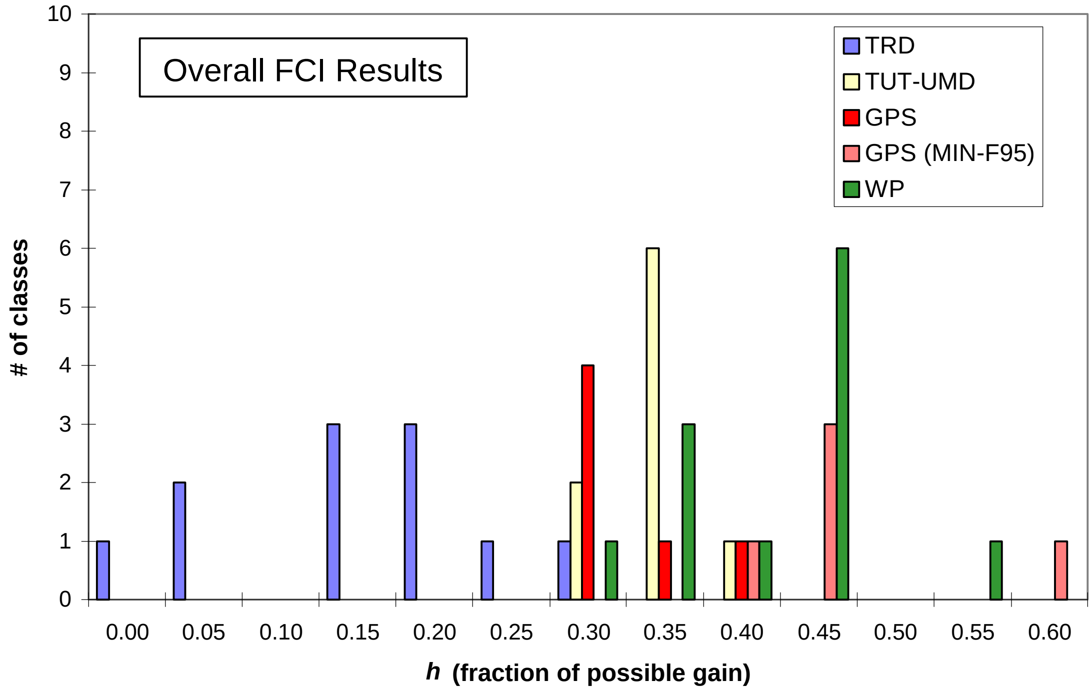
Average *h* for classes using research-based curricula: *h* = 0.39 (N = 2150)

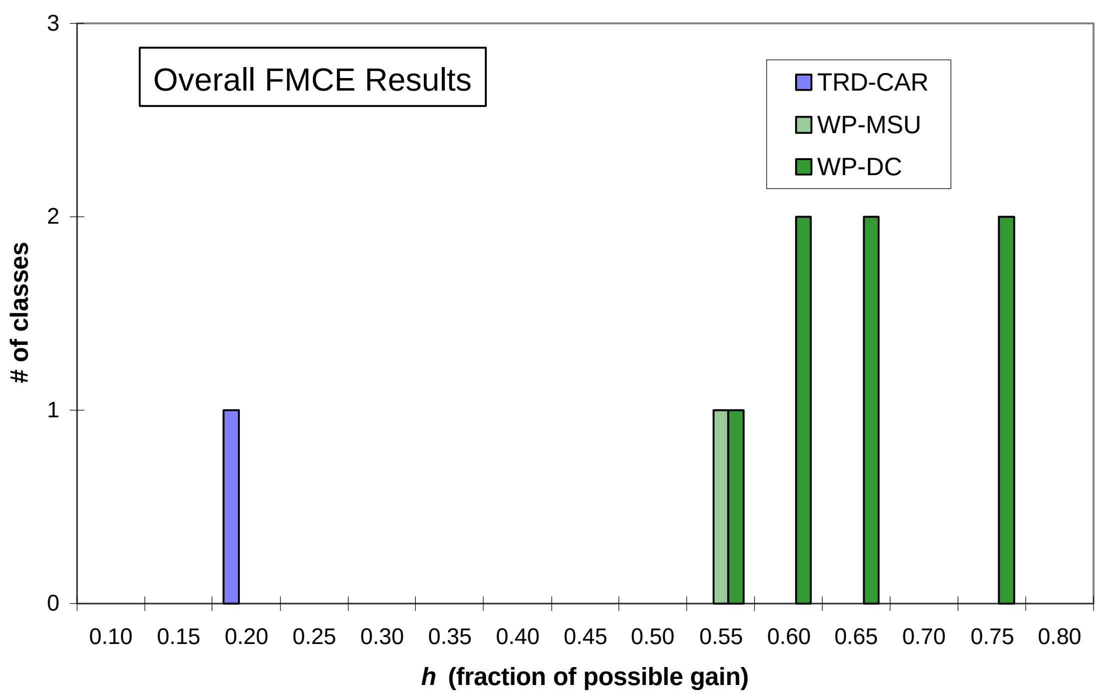
Average *h* for classes using traditional lecture instruction: *h* = 0.15 (N = 298)

Table 9-3: Overall FMCE results for Workshop Physics schools

	N (# of classes)	FMCE Pre	FMCE Post	*h*
CAR (F95)	24 (1)	25.9	39.4	0.18
MSU (F95)	16 (1)	37.5	72.2	0.56
DC (F94-F96)	154 (7)	24.1 ± 1.2	73.4 ± 2.2	0.65 ± 0.03

Figure 9-2. FCI (above) and FMCE (below) figure of merit histogram for classes from all ten schools participating in the study. The figure of merit *h* = the fraction of the possible gain on the full FCI, for traditional lecture classes (TRD), Tutorial classes (TUT), Group Problem Solving classes (GPS) and Workshop Physics classes (WP).





classes. The average *h* for all the traditional lecture classes at Maryland and the community college is smaller, but not significantly different than Hake's value.[15]

The average *h*-value for each class in this study using research-based curricula is better than all but one of the traditional lecture classes at Maryland and all but one of Hake's traditional undergraduate classes.[16] The worst *h*-values for the research-based curriculum classes are as good as these two best traditional classes. The best average *h*-values in this study were achieved by the Workshop Physics classes at Dickinson and Drury and by the fall 96 Group Problem Solving classes at University of Minnesota. Not surprisingly, these programs were considered the most successful implementations of their respective curricula by the instructors and outside evaluators. Also, Dickinson and Minnesota are the development sites for their respective curricula. The class at Drury College was in the second year of their implementation and had the second highest *h*-factor for any class participating in this study. This class is unusually small even for Drury and should not necessarily be considered typical.

The highest *h*-factors at Dickinson, NWU, Minnesota, and Maryland were achieved by classes that were taught by instructors directly involved in the development and/or implementation of the research-based curricula. I suggest that this may be due to better integration of the active-engagement elements and major themes in these courses. A senior graduate student in the physics education group at Ohio State made a similar observation in their implementation of the Group Problem Solving curriculum.[17] The highest FCI *h*-factor achieved by any class participating in this study was 0.59 for the GPS class taught by Ken Heller at Minnesota, who was involved with the development of the GPS curriculum.

The increase in the average *h*-factors at Minnesota between the fall 94 and fall 95 quarters is very large even discarding Heller's class as an outlier. When asked about this, the physics education researchers responsible for implementing the GPS curriculum at Minnesota commented on two differences between the two quarters.[18] The implementation of the problem solving labs was better in 1995 and the post-course FCI was given as part of the final exam in 1995 instead of in the laboratory in the last week of classes. The PER people at Minnesota suggest that perhaps the increase in scores is due to the students taking "the FCI much more seriously as part of the final exam."

There are three possible explanations for this remarkable improvement:

1. The students actually understood the concepts much better because of improvements in the curriculum or because they finally understood the concepts better after studying for the final exam.
2. The students don't take the FCI as seriously when they are not graded on their score.
3. In the context of the exam, the students are responding with what they were taught rather than what they believe.

There is not enough data here to distinguish between these three possibilities. That is left for future research. However, it is interesting to note that in Hake's study the highest *h*-factors were observed in classes where the post FCI was incorporated into the final exam.

Last, since Carroll College, Moorhead State University and the more recent Workshop Physics classes at Dickinson use the FMCE instead of the FCI for their evaluation of students conceptual understanding, FMCE results from both schools are included in Table 9-3 for comparison with the FCI results. Laws feels these FMCE results for Dickinson are more representative of their Workshop Physics classes due to improvements in the curriculum and some unusual difficulties in the 1992 fall semester.

By the end of the first term, the Workshop Physics classes at Moorhead State and Dickinson show significantly greater improvement in understanding of basic concepts compared to the traditional class at Carroll College both in terms of actual score and fractional gain *h*.

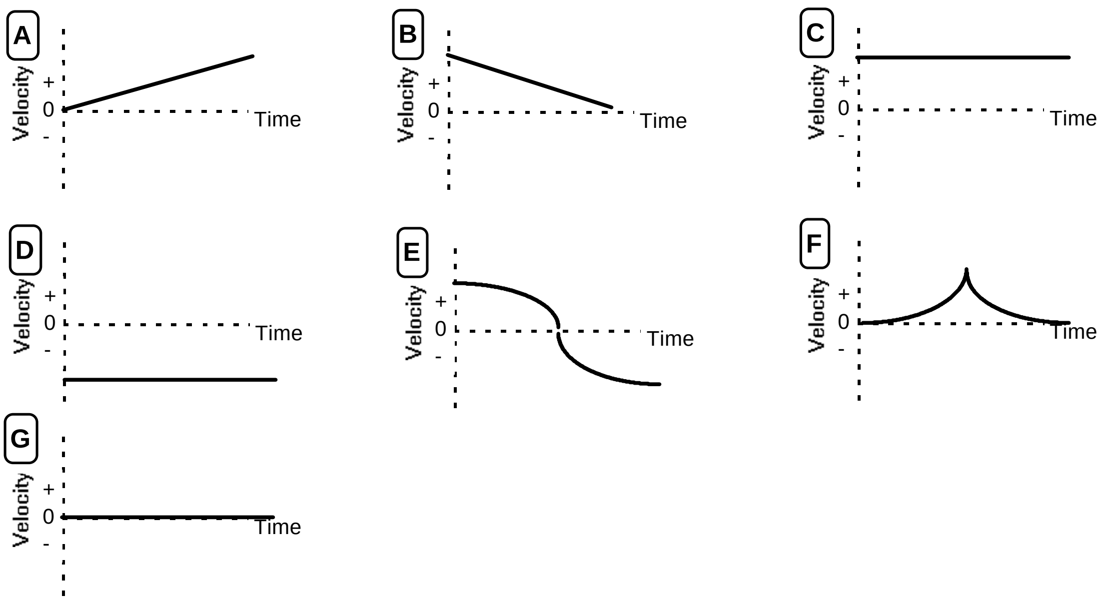
B. Velocity Graphs

Velocity graphs are essential to an understanding of mechanics and address the general issue of the relationship between a quantity and its rate of change. Velocity graphs are also known to be difficult for many introductory physics students (see the brief discussion on representations in chapter 2).[19] Thornton and Sokoloff found that student understanding of velocity graphs could be significantly improved using an MBL curriculum they developed.[20] They evaluated the effect of their curriculum using a set of multiple-choice velocity graph questions (VQ) in which students were required to match a description of a motion to a velocity graph. The VQ questions are shown in Figure 9-3. Thornton and Sokoloff demonstrated that students who were given four hours of their group-learning guided-discovery active-engagement MBL curriculum were significantly more successful in choosing the correct graphs than those who only received traditional lecture instruction.

The results are dramatic, with a large fraction of the students missing all but the simplest of the five velocity graph questions after traditional instruction.[21] After the MBL activities, the error rate drops to below 10% on all questions. This result is very robust and has been confirmed at dozens of colleges and universities. The results of

FIGURE 9-3. THORNTON-SOKOLOFF VELOCITY GRAPH QUESTION (VQ)

An object's motion is restricted to one dimension along the + distance axis. Answer each of the questions below by selecting the velocity graph below that is the best choice to describe the answer. You may use a graph more than once or not at all.



a. Which velocity graph shows an object going away from the origin at a steady velocity?

b. Which velocity graph shows an object that is standing still?

c. Which velocity graph shows an object moving toward the origin at a steady velocity?

d. Which velocity graph shows an object changing direction?

e. Which velocity graph shows an object that is steadily increasing its speed?

Thornton and Sokoloff are cited as an indication that interactive-engagement MBL activities are highly effective. However, some members of the physics education community question whether the improvement is due to the MBL activity or to the extra time on the topic? The following study addresses this question.

The VQ were given in two of the Maryland classes taught by Redish, my advisor. In the first class[22] (G0), Professor Redish did his best to teach the material explicitly in lecture, devoting nearly three full lecture hours to the topic of instantaneous velocity. Lecture demonstrations using the same MBL apparatus as in the tutorial were conducted with much student interaction and discussion. The professor had the students watch and plot the professor's motion as he walked a variety of paths, and a number of problems relating to students' personal experience were presented, but no worksheets were distributed. While the students were prompted for predictions for many of the situations, they were not given time to explain their ideas and predictions with one another. In the recitation sections, graduate-teaching assistants spent one hour going over textbook problems on the same material.

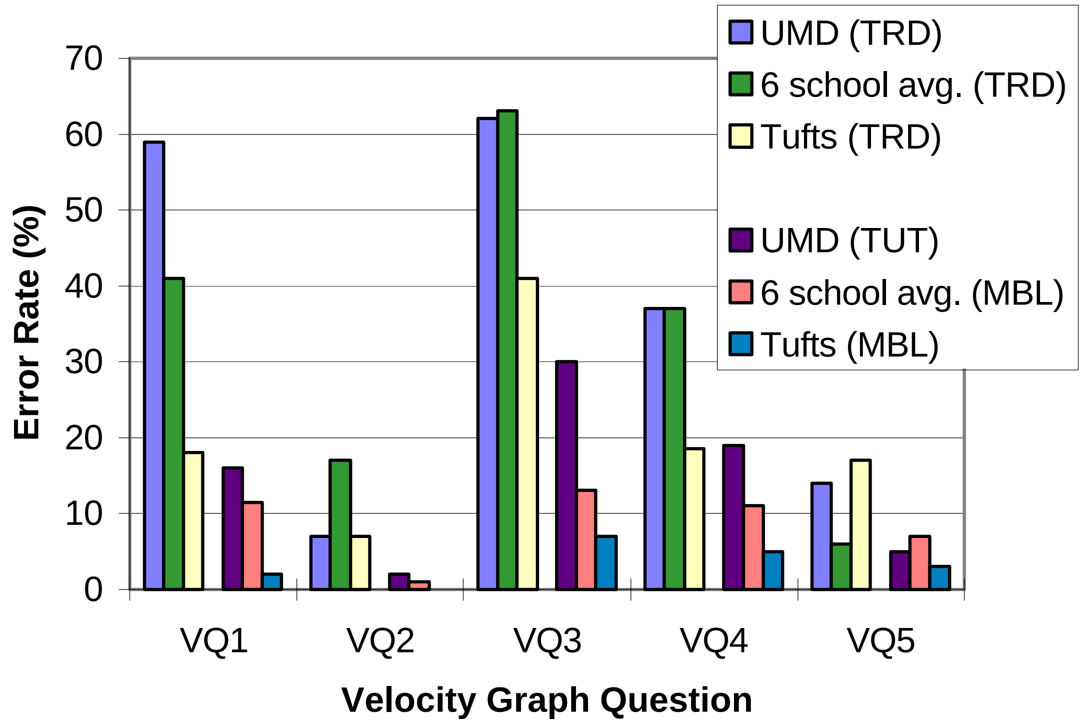
For class G2, the tutorial system was in place, and the one-hour MBL velocity tutorial written by Redish and the author (this tutorial is described in more detail in chapter 8) was given. The professor reduced the lecture time on the topic to a single hour, which was more typical of a traditional lecture and had little student interaction. In both classes, the questions were given as part of an examination and were not previously given to the students as homework. The results for the error rates are given in Table 9-4 and shown in Fig. 9-4. Note that since only post instruction data was taken, this data is not matched.

Table 9-4: Percentage error on the VQ with and without MBL. TRD indicates traditional lecture instruction tutorial; TUT indicates class was taught with tutorials.

Instruction without MBL	VQ1	VQ2	VQ3	VQ4	VQ5
University of Maryland (TRD) N = 100	59	7	62	37	14
Tufts (Thornton & Sokoloff)[23] N = 177	18	7	41	18.5	17
Six School Average (Thornton & Sokoloff)[24] N = 505	41	17	63	37	6

Instruction with MBL	VQ1	VQ2	VQ3	VQ4	VQ5
University of Maryland (TRD) N = 100	16	2	30	19	5
Tufts (Thornton & Sokoloff)[25] N = 177	2	…	7	5	3
Six School Average (Thornton & Sokoloff)[26] N = 505	11.5	2	13	11	7

Figure 9-4. Error rate on velocity questions (VQ). TRD indicates classes taught without tutorials or MBL, TUT indicates Maryland class was taught with tutorials, MBL indicates classes taught with 4 hours of MBL activities.



The Maryland results with four hours of traditional instruction on velocity and velocity graphs but no tutorial (class F0) resembled the 6-school average of traditional lecture classes reported in Thornton's lecture at the Raleigh conference.[27] Not surprisingly, the Maryland result from class F2 with one hour of MBL tutorial and one hour of lecture was substantially improved from four hours of traditional instruction, but not as good as the improvement shown with four hours of Thornton and Sokoloff's MBL activities.

These results are consistent with those given by Thornton and Sokoloff. The fact that these results have been obtained with both the lecturer and the time of instruction controlled strongly supports the finding by Thornton and Sokoloff that cooperative group activities with MBL can be more effective for helping students learn to understand graphical representations than traditional instruction. In this case, the improvement was achieved with two hours of instruction compared to four hours of traditional instruction so the improvement was not due to additional time on task. These results demonstrate that MBL group activities can play a significant role in improving student understanding of the concept of velocity. It is not simply the extra time that is responsible. It also suggests that simply enhancing lectures is not necessarily effective in producing an improvement in the learning of the velocity concept for a significant number of students.

C. Newton's Third Law

1. FCI Newton 3 Cluster at University of Maryland

Student understanding of Newton's third law was evaluated using the four FCI questions 2, 11, 13, and 14 (N3 FCI) shown in Figure 9-5. The results from the 16 traditional and tutorial classes at Maryland discussed previously are given in Table 9-5 and shown as a histogram in Figure 9-6. The table gives the fraction of students answering each of the N3 FCI questions correctly at the beginning (pre) and end (post) of the semester. A figure of merit, h = (class post-test average- class pre-test average)/(100 - class pre-test average), is calculated for each question in analogy with the Hake figure of merit for the full FCI. The four h-values are then averaged in the last column to give a figure of merit for the Newton 3 FCI cluster, h_{N3}.

The fractional gain results are systematically better for the tutorial classes. Indeed, every tutorial class has a higher value of h_{N3} than every non-tutorial class (though a similar statement is not true for the h-values for every individual question). The average values of h_{N3} for each group of classes are ($< h > \pm$ Std. Error):

$< h_{N3} > = 0.28 \pm 0.04$ (7 classes, with recitations)
$< h_{N3} > = 0.60 \pm 0.03$ (9 classes, with tutorials)
$< h_{N3} > = 0.41$ (1 class with no MBL tutorials)
$< h_{N3} > = 0.64 \pm 0.05$ (4 classes, with Newton 3 MBL tutorial)
$< h_{N3} > = 0.60 \pm 0.02$ (4 classes, with MBL tutorials but no Newton 3 MBL tutorial)

where the last four classes had other MBL tutorials including the velocity tutorial and the Newton's 2^{nd} law tutorial discussed in chapter 8. Using pooled variance, the

2. Imagine a head-on collision between a large truck and a small compact car. During the collision:

(A) the truck exerts a greater amount of force on the car than the car exerts on the truck.

(B) the car exerts a greater amount of force on the truck than the truck exerts on the car.

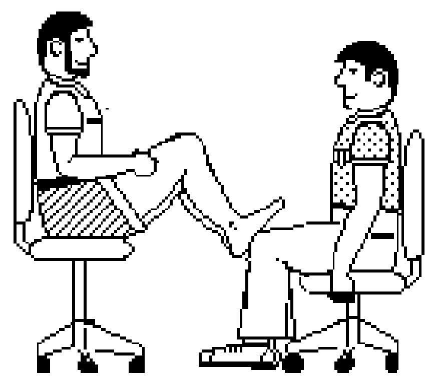
(C) neither exerts a force on the other, the car gets smashed simply because it gets in the way of the truck.

(D) the truck exerts a force on the car but the car does not exert a force on the truck.

(E) the truck exerts the same amount of force on the car as the car exerts on the truck.

11. In the figure at right, student "a" has a mass of 95 kg and student "b" has a mass of 77 kg. They sit in identical office chairs facing each other.

Student "a" places his bare feet on the knees of student "b", as shown. Student "a" then suddenly pushes outward with his feet, causing both chairs to move.



In this situation:

(A) neither student exerts a force on the other.

(B) student "a" exerts a force on student "b", but "b" does not exert any force on "a".

(C) each student exerts a force on the other, but "b" exerts the larger force.

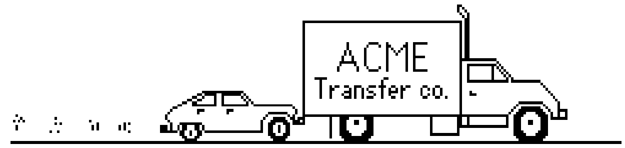
(D) each student exerts a force on the other, but "a" exerts the larger force.

(E) each student exerts the same amount of force on the other.

Figure 9-5. Newton’s third law FCI questions (N3 FCI) continued

Refer to the following statement and diagram while answering the next two questions.

A large truck breaks down out on the road and receives a push back into town by a small compact car.



13. While the car, still pushing the truck, is <u>speeding up</u> to get up to cruising speed:

(A) The amount of force with which the car pushes on the truck is equal to that of the truck pushing back on the car.

(B) The amount of force of the car pushing on the truck is smaller than that of the truck pushing back on the car.

(C) The amount of force of the car pushing against the truck is greater than that of the truck pushing back on the car.

(D) The car's engine is running so it applies a force as it pushes against the truck, but the truck's engine is not running so the truck cannot push back against the car. The truck is pushed forward simply because it is in the way of the car.

(E) Neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

14. After the car reaches the constant cruising speed at which its driver wishes to push the truck:

(A) The amount of force of the car pushing on the truck is equal to that of the truck pushing back on the car.

(B) The amount of force of the car pushing on the truck is smaller than that of the truck pushing back on the car.

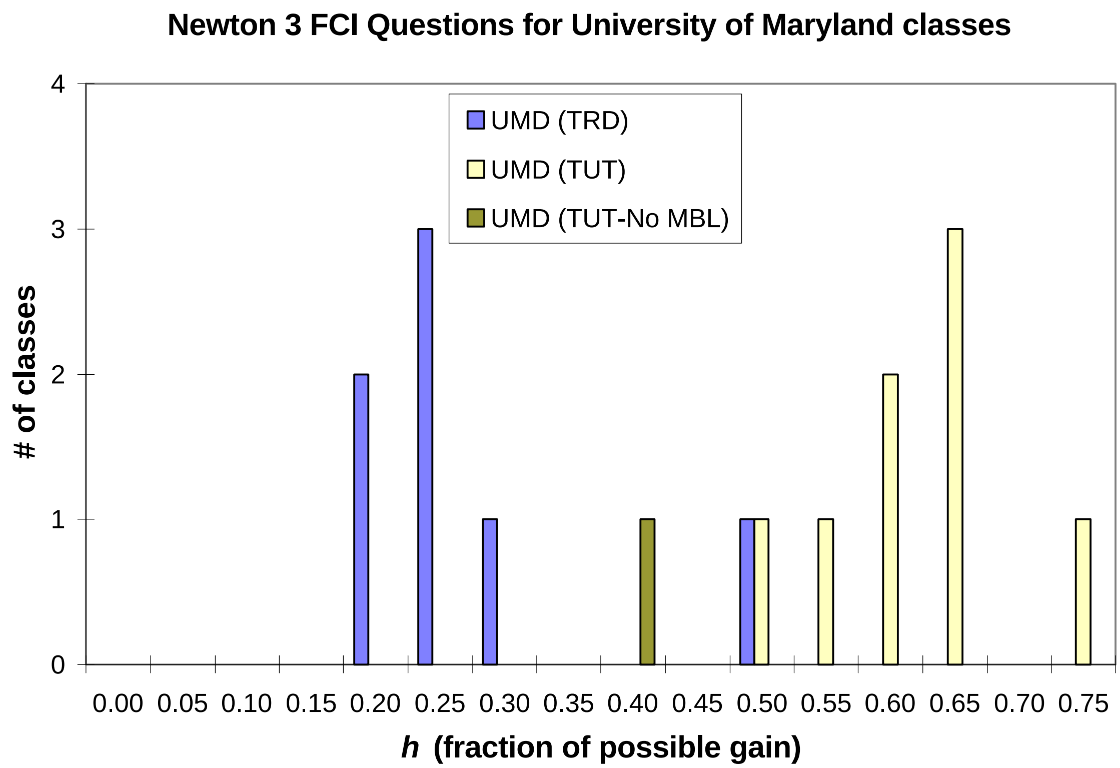
(C) The amount of force of the car pushing on the truck is greater than that of the truck pushing back on the car.

(D) The car's engine is running so it applies a force as it pushes against the truck, but the truck's engine is not running so it can't push back against the car. The truck is pushed forward simply because it is in the way of the car.

(E) Neither the car nor the truck exert any force on the other. The truck is pushed forward simply because it is in the way of the car.

Table 9-5. N3 FCI results for Maryland classes.

Figure 9-6. Histogram of average figures of merit (h_{N3}) for the Newton 3 FCI cluster for University of Maryland traditional (blue bars) and tutorial classes (yellow bars). All classes which used MBL tutorials with or without the Newton 3 MBL tutorial are shown in solid yellow. The crosshatched bar represents the one tutorial class taught without any MBL tutorials.



standard error of the last two $\langle h_{N3} \rangle$'s is $\sigma = 0.05$. Thus, if the other MBL tutorials are used, there is no significant difference in the student scores on the Newton 3 FCI cluster when the Newton 3 MBL tutorial is not used.

In the first semester in which tutorials were tested, there were no MBL tutorials and there was no tutorial specifically oriented towards Newton 3. The first MBL tutorials were implemented the following year. As a result, the first Maryland tutorial class, G1, used tutorials but no MBL tutorials. The same instructor taught the class with all three MBL mechanics tutorials in a later semester. This gives us a control for individual lecturer as well as for the presence of tutorials. (No special effort was devoted to Newton 3 in lecture in either case.) The result was:

$\langle h_{N3} \rangle = 0.41$ (F1: with no MBL tutorials)
$\langle h_{N3} \rangle = 0.65$ (F2: with velocity, Newton2, and Newton 3 MBL tutorials)

However, in later semesters a non-MBL Newton 3 tutorial was substituted for the MBL version while keeping the velocity and the Newton 2 MBL tutorials. As can be seen in the results reported on page 9-19, no significant difference in either the overall FCI or the FCI Newton 3 cluster results was observed.

2. FCI Newton 3 cluster for other research-based and traditional lecture curricula

A similar analysis was performed on the FCI data for other schools where the itemized FCI data was available (some schools only submitted the overall FCI score for each student). FCI Newton 3 cluster results from Ohio State, Minnesota, Dickinson, Skidmore, and Nebraska Wesleyan are summarized in Table 9-6. The distribution of fractional gains for classes taught with the four curricula is shown in a histogram in

Table 9-6. N3 FCI results for all four curricula

Figure 9-7. Histogram of average figures of merit (h_{N3}) for the Newton 3 FCI cluster for all four curricula: Traditional lecture classes at University of Maryland (blue bars), Tutorial classes at University of Maryland (yellow bars), Group Problem solving at University of Minnesota (F94) and Ohio State (red bars), and Workshop Physics classes at Dickinson College, Skidmore College, and Nebraska Wesleyan University.

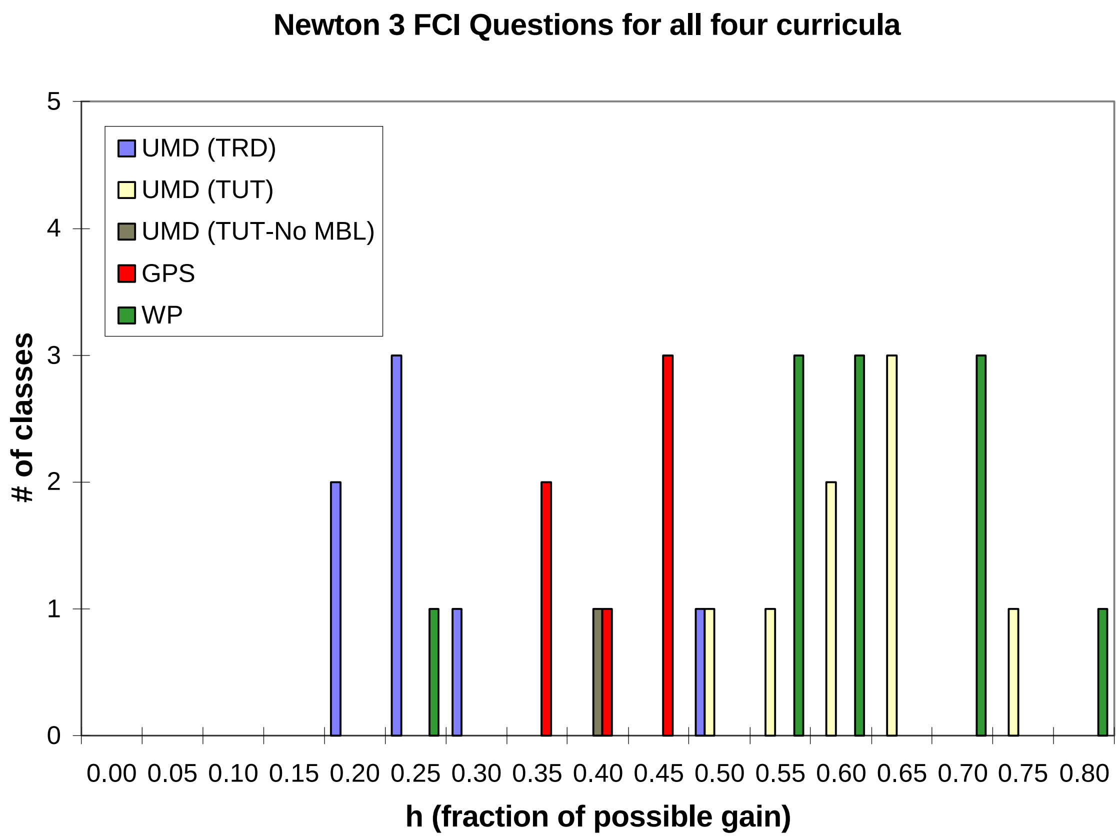


Figure 9-7. The Ohio State and the Skidmore data only represent the results of two classes each. Accordingly, the results from Ohio State are averaged over students instead of classes.

The average pre-course Newton 3 scores vary from 26% at Skidmore to 41% at Maryland with most of the scores clustering at 31%. Surprisingly, the pre-course averages for Minnesota and Ohio State are significantly smaller than Maryland’s.

One of the Maryland Tutorial classes and four of the Workshop Physics classes achieved a fraction gain $h > 0.70$ (two of the classes at Dickinson, one of the classes at Skidmore, and one of the classes at Nebraska Wesleyan). The result for the Skidmore class is especially notable since the pre-course score for the Newton 3 cluster was the smallest of any class participating in this study. In fact, the students at both Skidmore and Nebraska Wesleyan had significantly lower average pre-course Newton 3 FCI scores than the other four schools whose results are shown in Table 9-6.

The school average fractional gains for classes using one of the three research-based curricula were significantly higher than those using traditional lecture instruction at Maryland. The scores for all the research-based classes are significantly larger than all the traditional lecture classes except for classes B1 and one Workshop Physics class at Dickinson. Class B1 had the best overall and Newton 3 FCI results of the traditional lecture classes. As mentioned earlier, the Dickinson class had many difficulties including poor attendance.

The best fractional gains for the research-based classes were achieved by classes that incorporated at least some MBL group learning activities to address student difficulties with Newton’s third law. The Group Problem Solving classes from Ohio as well as all but two of the Tutorial and Workshop Physics classes. This result is not

solving, not conceptual understanding explicitly like Tutorials and Workshop Physics.

Also, no itemized FCI dat

implementation[28]

taught with the GPS curriculum in both the primary and secondary implementations

show greater improvement in

Maryland students taught with traditional instruction.

S AND APPLICATION OF CONCEPTS IN COMPLEX PROBLEMS

-choice questions tell us whether students "have" the desired information, it gives no information on whether they can access it in an appropriate complex problem. Accordingly, the main two questions in this section are the following:

Q1. on multiple choice mechanics questions imply that the students can use these

Q2. Mechanics is a large part of the first term course, but it is only one of many topics covered in the typical calculus based introductory physics sequence. Are students -based curricula better able to apply their improved conceptual

To answer these questions, an analysis of student responses for four pro specially designed for this study, are presented in this section. These problems look at students' use and/or understanding of velocity graphs & Newton's third law, position

interference.

A. Mechanics

In order to address question Q1, I developed an examination problem that required students to display both an understanding of a velocity graph and to use Newton's third law in a complex physical situation. The problem is shown in Figure 9-8 (this problem is a variation of the problem in Figure 6-2a.). The problem in Figure 9-8 was given on the final exam in one tutorial class and one non-tutorial class, classes G2 and C2 respectively (FCI data from these two classes is shown in Tables 9-1 and 9-5). Overall, performance on the problem was better for the tutorial than for the non-tutorial students. However, here we will only discuss issues related to the velocity graph and Newton's third law.

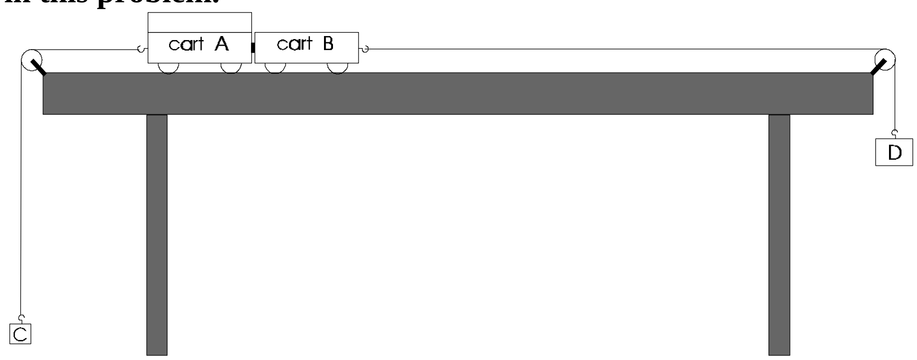
1. Velocity

Part of the examination question asked the student to generate a velocity vs. time graph for a complicated situation. The critical elements of a fully correct solution show the velocity starting at 0, increasing linearly until t = 3 seconds, and then decreasing linearly to some negative value.[29]

Students from both classes struggled with this question. Table 9-7 shows a breakdown of student responses. Only a small fraction of the students in either class were able to draw a graph that reflected the critical features, but the tutorial students did better than the students in the recitations. After traditional instruction, 12% of the

Figure 9-8. Long qualitative exam problem requiring both the construction of a velocity graph and an application of Newton's third law.

Two carts, A and B ($Mass_A > Mass_B$), are placed on a table then stuck together with Velcro. Using pulleys, two small blocks, C and D ($mass_C < mass_D$), are connected by light strings to the carts as shown below. Initially, the carts are held in place. **Ignore all friction in this problem.**



At t = 0, the carts are released. At t = 3 seconds, the Velcro pulls apart and the two carts separate. At some later time, cart A returns to its starting point.

a. Draw and label two separate free-body diagrams, **one for each cart**, for a time after the carts start moving but before the Velcro pulls apart.

b. Rank all the horizontal forces from both your diagrams by magnitude, from largest to smallest. **Explain the reasoning that you used to rank the forces**.

c. Briefly describe the motion of cart A from t = 0 until it returns to its starting point. On the graph below, qualitatively sketch the velocity vs. time for this time period.

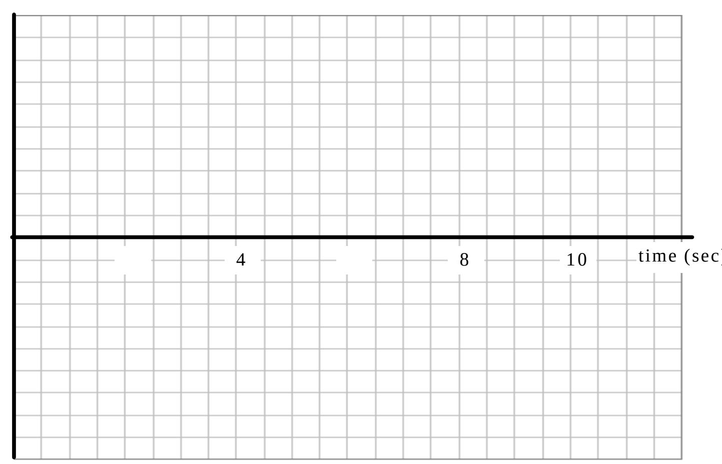


Table 9-7: Results on student constructions of the velocity graph in the long qualitative exam problem from two classes at the University of Maryland. The Tutorial class is in bold.

	% correct	% apparently correct, but ending at v = 0	% other incorrect response
Traditional (N = 50)	12	10	78
Tutorial (N = 82)	**22**	**21**	**57**

Table 9-8: Results on student use of Newton's third law in the long exam problem from two classes at the University of Maryland. The Tutorial class is in bold.

	% correct	% stated third law force pair have different magnitudes	% used the same symbol but did not compare forces	% with no identification of contact forces	% other incorrect response
Traditional	42	22	6	14	16
Tutorial	**55**	**40**	**0**	**1**	**4**

Table 9-9: Results on Newton 3 FCI questions for University of Maryland classes C2 and G2. See Tables 9-1 and 9-5 for comparison with other Maryland classes. The Tutorial class is in bold

	% correct Pre	% correct Post	h_{N3}
Traditional	35.9	50.0	0.23
Tutorial	**39.6**	**69.2**	**0.51**

students drew a correct graph. After MBL tutorials, 22% of the students drew a correct graph.

Analysis of the incorrect graphs along with the accompanying explanations revealed some of the students' difficulties. Many students showed in a variety of ways that they had the well-documented confusion between position and velocity (see the discussion on representations in chapter 2). Some drew graphs that at first glance appear correct: the graph increased linearly for the first 3 seconds and then decreased linearly after. However the graph ended at v=0, and some of these students indicated that this coincided with the cart returning to its starting location (an example of this type of solution is shown in Figure 6-2a). Many students drew graphs that had incorrect combinations of linear segments, including discontinuities in velocity. Others drew dramatically curved features in their velocity-time graphs. Most of these graphs indicated severe conceptual difficulties even if interpreted as a position vs. time graph. It is worth noting that it is clear from their explanations that most of these students intended to draw a velocity vs. time graph.

Both the percentage of correctly drawn graphs and the nature of the incorrect graphs confirm that while student difficulties understanding kinematics is pervasive even after instruction, the modified instruction described earlier in this paper appears to be helping address these difficulties somewhat. Although the VQ were not given in these classes, approximately 70% of the students in the comparable tutorial class F2 answered all of the multiple choice questions correctly, while only about 40% of those in the recitation class A1 answered them all correctly. The relative results on the long-problem are qualitatively consistent with the results of the VQ, but the absolute number of students getting correct answers on the long-problem was substantially lower (22% of the tutorial students correct vs. 12% of recitation students correct). Since no classes were evaluated with both the VQ and the long problem, we cannot completely answer Question Q1, but our indications are that the VQ does not suffice. Our results suggest that answering multiple-choice questions correctly is not sufficient to guarantee a robust and fully functional understanding of the relevant concepts for a significant number of students.

2. Newton 3

Another part of the same examination question tested student facility with dynamical concepts, specifically Newton's 2nd and 3rd laws of motion. The students were asked to draw a free body diagram of each cart shown in Figure 9-8 and to rank the magnitudes of the horizontal forces. Note in particular that by Newton's third law, the magnitude of the force of cart A on cart B is equal to that of cart B on cart A.

The breakdown of student responses to this part of the question is shown in Table 9-8. In the tutorial classes, 55% of the students correctly identified and compared the third law force pair. In the non-tutorial class 42% identified and correctly compared these forces.[30] Many students identified that the two carts were exerting forces on one another, but stated explicitly that the two forces were not of equal magnitude. In addition, there were also many students who did not even recognize that the two carts exert forces on each other. This was particularly common in the non-tutorial class.

These results should be compared with the results on the Newton's third law FCI questions for the same two classes shown in Table 9-9. The two classes' pre-course Newton 3 FCI scores are not significantly different ($\Delta < \sigma$), but the post-test results of

69% and 50% respectively are very similar to the ratio of correct responses to the exam problem for the two classes. The discrepancy between the multiple-choice and long-answer problems (in this case both questions were done by both groups) also suggests that the answer to question Q1 might be: the short answer results provide an indication but overestimates the students' knowledge.

B. Beyond Mechanics

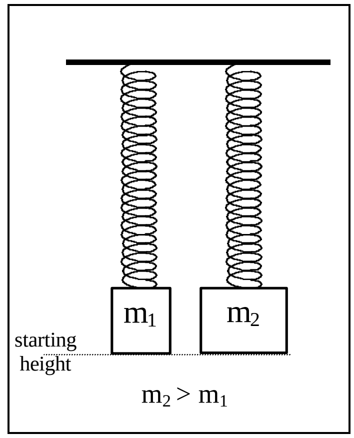
1. Harmonic Oscillator

As we have seen in chapter 2 and in the section above, even with enhanced instruction many students seem to have difficulty using velocity graphs in complex mechanics' situations; however, students seem to have much less trouble with position graphs of mechanics situations. This has been demonstrated by Beichner (see chapter 2)[31] as well as Thornton and Sokoloff.[32] Our own pretests at the beginning of the sequence indicate that most students have little trouble with simple position graphs in mechanics. But what about position graphs for non-linear motion?

In a study of students' understanding of oscillations and waves, we noticed that many students had trouble connecting sinusoidal graphs to physical quantities such as velocity, acceleration, and force. Redish and I designed and implemented an MBL tutorial for harmonic oscillators to help the students overcome this difficulty using both force and motion sensors. This tutorial went through several iterations. After the second iteration, the problem shown in Figure 9-10 was given on the final exam. Keep in mind the harmonic oscillator tutorial is the first tutorial at the beginning of the

Figure 9-9. Vertical harmonic oscillator problem for first tutorial class

The figure at the right shows two identical, massless, frictionless, springs, each with spring const
from a bar. Attached to one spring is a mass m_1
to the other spring is a mass m_2 $_2 > m$. At t = 0, the two masses are connected to the springs and released



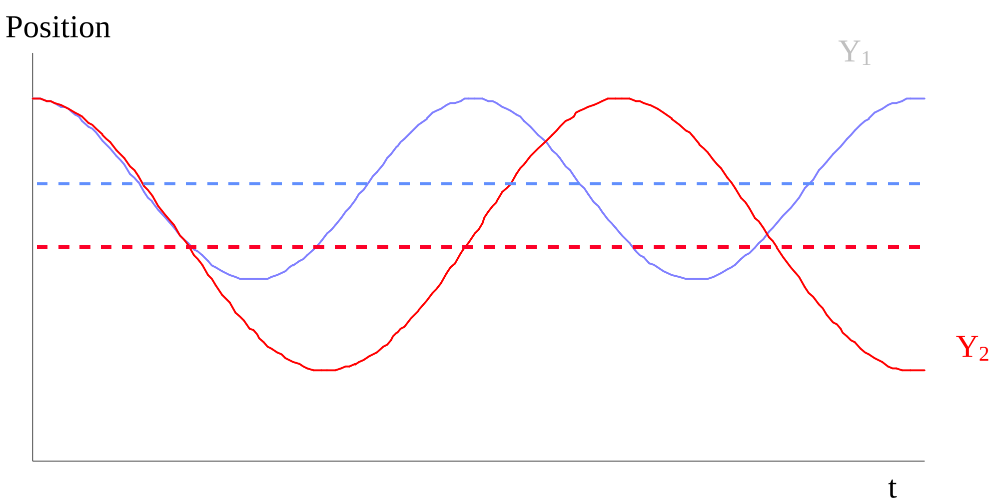
(***Note***: *When the masses are at their starting height, springs are at their unstretched lengths.)*

single *y* *t* *y* *t* is time), sketch the motion of each of the masses. Label

to which mass.

B. i) Determine the *y* *t* *m* .

y *t* *m* .



1 1 1 ω t)

Y = D + A cos(2

2 1 2 1 ω > 2

-10:

to the two harmonic oscillators problem shown in Figure 9-

oscillator tutorial was modified (N = 72 students). If the solution is not clear, the student is giv

The sketch of the motion of each of the masses on a single axes.

Starting points of the drawn curves:

(starts at maximum displacement)

Sine Curves

Other 5%

Distance from initial positions to equilibrium positions of the two masses:

d_1 $_2$ 28% Correct

Same equilibrium distance

Can't tell

A < A_2

A = A 8%

A_1 $_2$ 6%

Comparison of the periods of oscillation of the two mass:

T_1 $_2$ 49% Correct

T_1 $_2$ 28%

T_1 $_2$ 11%

Other (either inconsistent or indeterminate) 12%

Only 19 % of the sketches drawn were basically corre sinusoidal starting from the same position at maximum displacement and the equilibrium position $_1 < d$.

Equations for the two oscillating masses:

The equations are correct 6%

ith the sketch

The equations are incorrect & inconsistent with the sketch 46%

10%

Equations are incomplete or no equations given 8%

semester. Even so, the results from my analysis of the student responses to this problem were very surprising. The results are summarized in Table 9-10.

A graph was considered basically correct if the two curves were sinusoidal, they started from the same point at maximum displacement, and the equilibrium positions of the two curves were different. Although 85% of the students clearly showed that the amplitude for m_2 was greater than that of m_1, only 20% of them drew a graph that was basically correct. The main difficulties are listed below:

- Almost two-thirds of the student drew sinusoidal curves that started at the equilibrium positions of the two masses;
- Fewer than half the graphs showed that the period of mass m_2 was greater than the period of mass m_1; and
- Only 30% of the students showed in their graphs that the two masses would oscillate around different points.

In addition, only 45% of the students wrote an equation that was consistent with their graph. Another 45% of the students wrote an equation that was both incorrect and inconsistent with the graph including 10% of the class who explicitly used terms from equations that describe traveling or standing sinusoidal waves.

As one can tell from the above results, very few students were able to present a consistent and correct solution. However, many students had several parts of the correct answer. This is consistent with the view expressed in earlier chapters that students' conceptual knowledge is fragmented and not well organized. The fact that almost two-thirds of the graphs started the masses in the equilibrium position and that fewer than half the equations were consistent with the graphs suggest that, at least on this exam, many of the students saw only weak connections between the graph, the equation, and the physical situation, even for a position-time graph.

After I discussed the results with the PER group at Maryland, two members of

[33]) revised the harmonic oscillator tutorial to help

sical

quantities. The revised tutorial is included in Appendix B. The exam problem in Figure

-9 was used as a homework problem for the new tutorial. I designed a new harmonic

context. In this

problem, the two masses are the same but one spring constant is four times greater than

Figure 9- em was placed on the final

exam after the new tutorial was taught at the beginning of the second semester. My

-11.

This time a graph was considered basically correct if the period of cart one wa

longer, the two curves had equal amplitudes, and both curves started from the same

maximum displacement. Almost two thirds of the students drew the curves correctly.

Roughly 80% of the students recognized that the curves should be sinusoidal, start at

maximum displacement, and have equal amplitudes. Although only 36% of the students

indicated that the period of cart 1 was twice that of cart 2, 78% of the students indicated

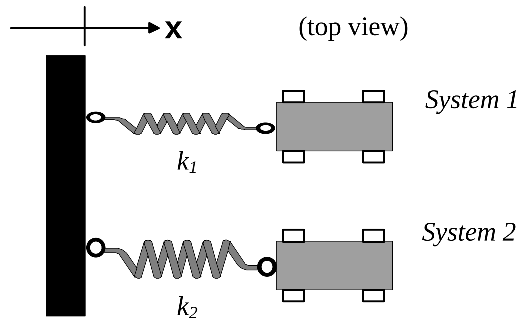
own an

equation that was roughly consistent with the graphs they had drawn. One third of the

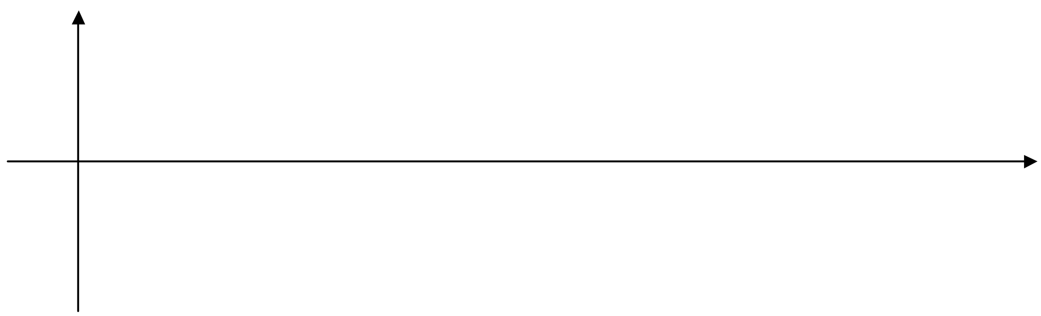
used terms from equations that describe traveling waves and 6% who used one or more

Figure 9-10. Horizontal harmonic oscillator problem for second tutorial class

Two identical **frictionless**
mass *m*
(see figure at right). Initially, both springs
unstretched and both carts are at rest at
$x = x_0$ ***x* is the distance from the**
. The
massless springs in systems 1 and 2 have
k_1 k_2
where $k \; = 4k_1$



Both c *d* and released at seconds. Sketch the motion of each cart after time $t = 0$ sec on the axes below. Clearly identify which curve corresponds to which cart. Label axes clearly.



b) Determine the equation which gives *x* as a function of *t* for cart 2. Explain how you arrived at your answer.

Correct Response: $x = x_0 + d \cos w_2 t$ where $w_2 = \sqrt{k_2/m}$. Note $w_2 = 2w_1$.

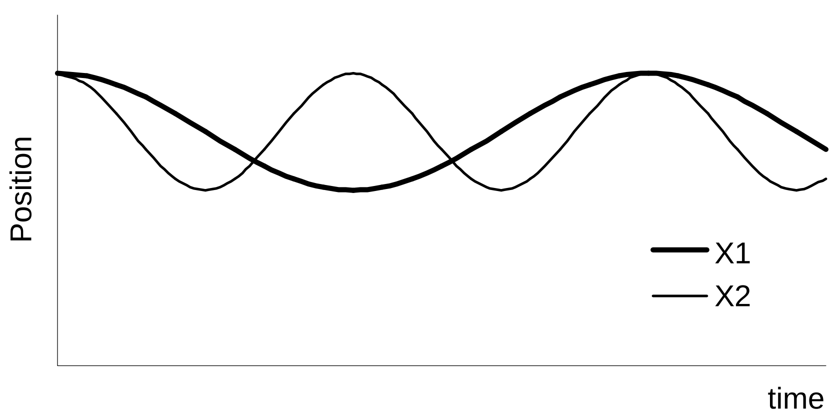


Table 9-11: Summary of student responses from a University of Maryland Tutorial class to the two harmonic oscillators problem in Figure 9-10 after the harmonic oscillator tutorial was modified (N=67 students). If the solution is not clear, the student is given the benefit of the doubt.

The sketch of the motion of each of the carts on a single axes.

Starting points of the drawn curves:		
Cosine Curves (starts at maximum displacement)	79%	Correct
Sine Curves (starts at equilibrium position)	16%	
Other	4%	
Comparison of the amplitudes of oscillation for the two masses:		
$A_1 = A_2$	84%	Correct
$A_1 > A_2$	4%	
$A_1 < A_2$	7%	
Drew only 1 curve	4%	
Comparison of the periods of oscillation of the two mass:		
$T_1 = 2T_2$	36%	Correct
$T_1 > T_2$ *but the ratio is not clear*	30%	Correct
$T_1 = 4T_2$	12%	
$T_1 = T_2$	10%	
$T_1 < T_2$	10%	
Can't tell	1%	

After modifications to the tutorial, 66% of the sketches drawn were basically correct, i.e. the period of cart one was longer, the two curves had equal amplitudes, and both curves started from the same maximum displacement.

Equations for the two oscillating masses:	
The equations are correct	65%
Roughly consistent	57%
No equation was given	10%
The equations have some elements from wave motion	13%
Student used a kinematic equations in their solution	6%

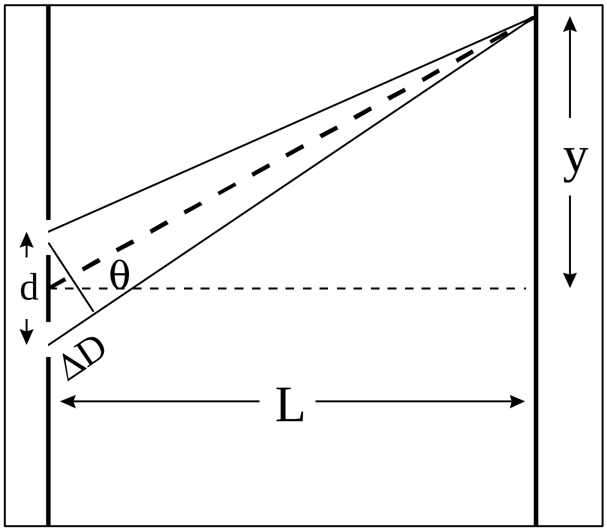
The results of the two exam problems strongly suggest that the revised MBL harmonic oscillator tutorial helped to improve student understanding of the connection between the position graphs, the equations, and the physical situation. Even though the first harmonic oscillator problem is marginally harder because of the vertical oscillations around different equilibrium, less than a third (29%) of the students from the first tutorial class drew graphs that were basically correct even when the equilibrium positions are removed from consideration compared with two thirds of the students in the second tutorial class. An evaluation of the student understanding of the graphs of other physical quantities pertaining to harmonic oscillators such as velocity and/or force is left for future studies.

2. Two-Slit Interference

One area covered by the tutorials in the third semester of the introductory physics sequence at University of Maryland is physical optics. The problem shown in Figure 9-11 was part A of an exam problem written by Richard Steinberg of the Maryland PER group to see how students taught with physical optics tutorials would do on a semi-conventional interference problem. This problem was given on exams in one tutorial class and two traditional lecture classes. The student responses to part A were analyzed by Sabella and Steinberg.[34] The results of their analysis are summarized in Figure 9-12. The tutorial class did significantly better on this problem, 60% vs. 16%. It is interesting to note that 40% of the students taught with traditional instruction applied the conditions for the first maximum next to the central maximum (instead of the first compared to only 9% of the tutorial students. This indicates that a much larger

Figure 9-11: 2-Slit Interference Problem

Light with λ = 500 nm is incident on two narrow slits separated by d = 30 μm. An interference pattern is observed on a screen a distance L away from the slits. The first dark fringe is found to be 1.5 cm from the central maximum. Calculate the distance, L, to the screen. Show all work.



Model Solution:

$$\Delta D = d \sin q$$

$$\Delta D = (m + \tfrac{1}{2}) l = \tfrac{1}{2} l$$

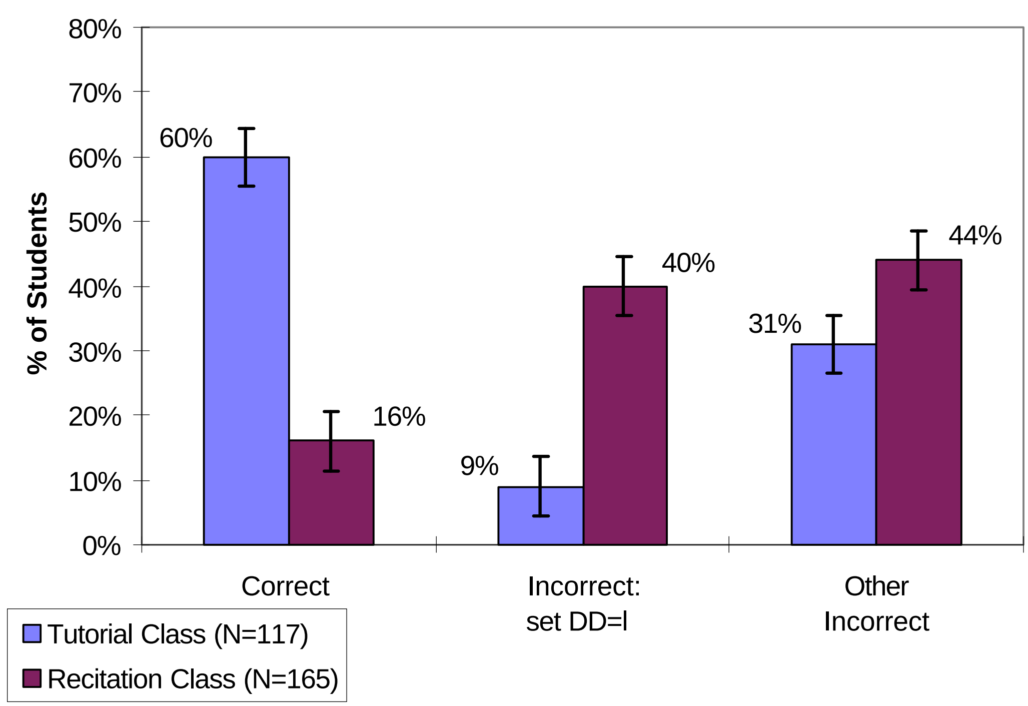
Assume $L >> y$, then

$\sin q$ is almost equal to $\tan q$

since q is small so $\tan q = y/L$

therefore $L = 2dy/l = 1.8\, m$

Figure 9-12: Graph of Student Responses to Double Slit Interference Problem.
Note: DD = ΔD



fraction of the traditional class was using an equation without considering the conditions of the physical situation. Other incorrect responses included algebraic mistakes and using higher order minima.

IV. PROBLEM SOLVING INTERVIEW

The two-rock problem protocol discussed in chapter 7 was given as a think-aloud interview exercise after expectation interviews with some of the students during site visits at Maryland, Dickinson, and Ohio State. This problem was used in interviews by Hammer in his dissertation study of student expectations (see chapter 2). Since all our students are volunteers, most of the students interviewed were in the top half of their introductory physics class. Even so, only one or two students at each of the three schools were able to work through the problem with only a few hints. The students who were able to work through the problem demonstrated a good conceptual understanding of kinematics but only one of these students thought about using conservation of mechanical energy to compare the speed of the rocks when they hit the ground.

The student volunteers were first asked what they thought the answers would be and why. Almost all of the students were able to answer just one of the two questions correctly. Many of them recognized that the rock thrown downward would hit the ground first but then assumed that meant the rock thrown downward would hit at the higher speed. The other students thought the two rocks would have the same speed when they hit the ground. This implied to them that the two rocks would hit the ground at the same time. Because acceleration was the same for both rocks, the students connected equal time to equal speed. Very few of the students seemed to recognize the vector nature of velocity or look at this problem in terms of energy. The latter is an indication that word cues in the problem trigger the students to think of this problem as a kinematics problem.

Many of the students were reluctant to attempt a mathematical solution. This was particularly true of the Dickinson Workshop Physics students. After discussing what they qualitatively predicted the answer would be, they wanted to test their prediction with an experiment. This is understandable since prediction followed by experiment is a major component of the Workshop Physics curriculum. But it is disappointing since students are also expected to be able to apply analytical tools as well

Most of the student volunteers were able to recall the kinematic equations; however, very few of them were able to derive them or state that the kinematic equations require constant acceleration when asked if the kinematic equations are always valid. Many of the student volunteers were surprised when we went over the derivation of the kinematic equations using integration and claimed they had never seen that derivation before. And as observed previously in Hammer's dissertation study,[35] few of the students had a good conceptual understanding of the equation $v(t) = v_0 + at$.

One thing that makes this problem difficult is that while it is possible to derive a symbolic expression for the time it takes each rock to fall, the symbolic expressions are difficult to compare. Most of the students were able to derive the two equations below:

$$t_1 = \frac{-v_0 + \sqrt{v_o^2 + 2ah}}{a}$$

$$t_2 = \sqrt{\frac{2h}{a}}$$

where t_2 is the time it takes the rock thrown horizontally with speed v_0 to fall a distance h while t_1 is time it would take the rock thrown straight down with the same speed to cover the same distance. The students were encouraged to find a symbolic solution to the problem. Most of the students plugged numbers into the two equations above. Only a few students were able to come up with additional ways of approaching the problem on their own including using average velocity or conservation of mechanical energy. This implies that if the students' knowledge is flexible enough they are able to come up with another approach to solve the problem.

Now, recognizing that these students were put on the spot in what was essentially a videotaped oral exam, they did not do that badly. However, the inability of most of the students to derive the kinematic equations using integration or conceptually understand the velocity equation is more disturbing, particularly since all the students had at least one term of physics instruction that emphasized learning with understanding.

IV. SUMMARY

In this chapter, results have been presented from two types of measurements of students' conceptual understanding of physics in the calculus-based introductory course:

1. How well students know basic physics concepts as measured by multiple choice tests like the FCI and FMCE when taught with Tutorials, Group Problem Solving, Workshop Physics, or traditional lecture instruction at various institutions.
2. How well students use physics concepts in solving complex problems when taught with tutorials or traditional lecture instruction.

Pre- and Post-course concept-test data were collected from classes at all ten schools participating in this study. The overall FCI and FMCE results from the 10 schools clearly show that classes taught with one of the research-based curricula improve significantly more than traditional lecture classes. This is even true for classes where the research-based curriculum is in the first year of adoption at the institution. Instructors at University of Maryland who taught with both tutorials and traditional recitations had better fractional gains from the classes with Tutorials.

Similar results were observed for the Newton's third law cluster of FCI questions. All the Tutorial classes at Maryland taught with MBL tutorials had significantly larger fractional gains than the traditional lecture classes at Maryland, even though both classes had similar pre-course scores on these questions. Although some the Workshop Physics classes started with much lower scores initially, the fractional gains of the classes taught with Workshop Physics and Group Problem Solving were also significantly better than the fractional gains for the traditional classes at Maryland. The Workshop Physics classes and the Tutorial classes, which used at least some MBL activities, appeared to improve more than the classes that did not. In addition, a Tutorial class which used the MBL velocity tutorial did better on velocity graph questions than a traditional lecture class taught by Redish who spent twice as much time (4 hours vs. 2 hours) going over velocity in the traditional class.

Four specially designed exam problems were used to see how well University of Maryland students were able to apply concepts when solving complex problems. One of these problems, which looked at students' understanding of velocity graphs and Newton's third law, was placed on the final exams of a Traditional class and a Tutorial class. The Tutorial students showed a significantly better understanding of the two concepts. The difference on Newton's third law between the types of classes was similar to the differences in the class scores on the Newton three FCI cluster.

The other three problems were used to look at student understanding of concepts beyond mechanics. Two similar harmonic oscillator problems were given on final exams before and after the harmonic oscillator tutorial was revised. Each problem required students to sketch a graph the motion of two harmonic oscillators on the same axes. The first problem was marginally more difficult because the two oscillators were oscillating vertically around two different equilibrium positions. If this difference is removed from consideration, the fraction of students who were able to draw a qualitatively correct graph of the two oscillators increased from 29% to 66% after the tutorial was revised. The fraction of students who wrote an equation to describe the motion consistent with their graph increased from 45% to 57%. These results indicate that the classes taught with Tutorials not only had higher scores on concept test but also did better on exam problems using similar concepts. Moreover, the students in the tutorial classes showed a better understanding of concepts on exam problems beyond mechanics. In addition, Sabella *et al.* showed that Maryland Tutorial students performed better on a two-slit interference exam problem than Maryland students taught with traditional lecture.[36]

However, only a few students from classes taught with research-based curricula were able to solve a difficult mechanics problem in interviews. Because of the nature of the problem, this was not surprising. What was surprising was that many of the students were unable to derive the kinematic equations and became stuck when the kinematic equations produced an expression that was hard to evaluate. This suggests that the research-based curricula may need further improvement to help students develop a more flexible and functional understanding of physics concepts and connect the concepts better to the equations.

[1] E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, New Jersey, 1997).

[2] D. Hammer, "Two approaches to learning physics," *Phys. Teach*. **27** (9), 664-670 (1989).

[3] D. Hestenes, M. Wells, and G. Swackhamer, "Force Concept Inventory," *Phys. Teach*. **30** (3), 141-153 (1992).

[4] R.R. Hake, "Interactive-engagement vs. traditional methods: A six thousand student study of mechanics test data for introductory physics courses," *Am. J. Phys*. **66** (1), 64-74 (1998).

[5] R.K. Thornton and D.R. Sokoloff, "Assessing student learning of Newton's laws: The Force and Motion Concept Evaluation," *Am. J. Phys*. **66** (4), 228-351 (1998).

[6] G.J. Posner, K.A. Strike, P.W. Hewson, and W.A. Gertzog, "Accommodation of a scientific conception: Towards a theory of conceptual change," *Sci. Educ*. **66** (2), 211-227 (1982).

[7] See Ref. 3.

[8] See Ref. 5.

[9] R.J. Beichner, "Testing students' understanding of kinematics graphs," *Am. J. Phys*. **62** (8), 750-762 (1994); R.K. Thornton and D.R. Sokoloff, "Learning motion concepts using real-time microcomputer-based laboratory tools," A*m. J. Phys*. **58** (9), 858-867 (1990); and D.P. Mahoney, "Research of problem solving: Physics," in the *Handbook of Research on Science Teaching and Learning*, edited by D.L Gabel (MacMillan Publishing Company, New York, 1994), 327-354.

[10] D.M. Hammer, *Defying Common Sense: Epistemological Beliefs in an Introductory Physics Course*, Ph.D. Dissertation, University of California at Berkeley (1991, unpublished).

[11] In terms of absolute final scores, one tutorial class with a low pre-test score finished below some of the non-tutorial classes, and one non-tutorial class with a high pre-test score finished above some of the tutorial classes.

[12] See Ref. 4.

[13] D.C. Howell, *Statistical Methods for Psychology*, 3rd Ed. (Duxbury Press, Belmont, CA 1992), 181-185.

[14] The fall 1995 class at University of Minnesota was the third year in the implementation of the Group Problem Solving curriculum in the calculus-based introductory physics sequence.

[15] See Ref. 4.

[16] See Ref. 4.

[17] Private communication with Chris Cooksey (August 1997).

[18] Private communication from Laura McCullough and Tom Foster (November 1997).

[19] See Ref. 9.

[20] Thornton and Sokoloff, *Tools for Scientific Thinking* (Vernier Software, Portland OR, 1992 and 1993).

[21] R.K. Thornton & D.R. Sololoff, "Learning motion concepts using real-time microcomputer-based laboratory tools," *Am. J. Phys*. **58**, 858-867 (1990).

[22] This class was taught prior to our use of the FCI at University of Maryland and is not shown in table 9-1.

[23] See Ref. 21.

[24] See Ref. 21.

[25] See Ref. 21.

[26] See Ref. 21.

[27] R.K. Thornton, "Tools for Scientific Thinking: Learning physical concepts with real-time laboratory tools," in *The Conference on Computers in Physics Education , Proceedings,* edited by E.F. Redish and J.S. Risley (Addison-Wesley, New York, 1990), 177-189.

[28] Private communication with Tom Foster and Laura McCullough, Fall 1997.

[29] A graph reversed with respect to the horizontal axis would also be considered correct.

[30] A few students in the non-tutorial class used the same symbol for these two forces, but did not state whether the forces were equal, so it was impossible to determine if they were identifying these two forces as having equal magnitudes. Note that many students used the same symbol to represent forces that clearly had different magnitudes.

[31] R.J. Beichner, "Testing student interpretation of kinematics graphs," *Am. J. Phys*. **62** (8), 750-762 (1994).

[32] R. Thornton, "Using large-scale classroom research to study student conceptual learning in mechanics and to develop new approaches to learning," in *Microcomputer-*

Based Labs: Educational Research and Standards, edited by R.F. Tinker, *Series F, Computer and Systems Sciences* **156** (Springer-Verlag, Berlin, Heidelberg, 1996), 89-114.

[33] Richard N. Steinberg is a Post Doc with the University of Maryland Physics Education Research Group; Michael C. Wittmann is a senior graduate student in the group.

[34] M.S. Sabella, R.N. Steinberg, and E.F. Redish, "Student performance on traditional exam questions in two instructional settings," *AAPT Announcer* **27** (2),122 (1997, abstract only).

[35] See Ref. 10.

[36] See Ref. 31.