A modeling method for high school physics instruction

Malcolm Wells and David Hestenes
Department of Physics and Astronomy, Box 871504, Arizona State University, Tempe, Arizona 85287-1504

Gregg Swackhamer
Glenbrook North High School, 2300 Shermar Road, Northbrook, Illinois 60062

(Received 21 September 1994; accepted 31 December 1994)

The design and development of a new method for high school physics instruction is described. Students are actively engaged in understanding the physical world by *constructing and using scientific models* to describe, explain, predict, and to control physical phenomena. Course content is organized around a small set of basic models. Instruction is organized into *modeling cycles* which move students systematically through all phases of model development, evaluation, and application in concrete situations—thus developing skill and insight in the procedural aspects of scientific knowledge. Objective evidence shows that the *modeling method* can produce much larger gains in student understanding than alternative methods of instruction. This reveals limitations of the popular "cooperative inquiry" and "learning cycle" methods. It is concluded that the effectiveness of physics instruction depends heavily on the pedagogical expertise of the teacher. The problem of cultivating such expertise among high school teachers is discussed at length, with specific recommendations for action within the physics community. *© 1995 American Association of Physics Teachers.*

I. INTRODUCTION

Malcolm Wells is the primary author of this paper, because it is about his contribution to physics teaching. Malcolm has intended to publish an account of his work since his doctoral dissertation was completed in 1987. But the writing was delayed, first because he gave himself to conducting workshops for the benefit of other teachers, and then, in the last few years, because Lou Gehrig's disease has consumed his energy in implacable decline. So it has fallen on his coworkers, D.H. and G.S., to speak for and about Malcolm Wells. We do this gladly to celebrate the life of a truly great teacher, but more—because Malcolm has elevated the craft of teaching, and we believe that his unique contributions can help others surpass themselves and perhaps even Malcolm.

II. MALCOLM'S EDUCATIONAL RESEARCH

The story of Malcolm's research is told by D.H., who directed Malcolm's doctoral work and continued to collaborate with him thereafter. The story has an *unambiguous moral*: to upgrade high school physics, *partnerships* are needed between experienced teachers and physicists involved in educational research.

By any conventional measure, Malcolm was a superior teacher before his partnership with me. Yet his doctoral thesis documents a *large* improvement in the outcomes of his teaching, and it clearly identifies the contribution of educational research to the change. I have been active in theoretical physics research for the duration of our partnership. Though my physics research has deeply influenced my educational research, only the latter has been of direct benefit to Malcolm. Here is the story.

When Malcolm approached me about doctoral research, he was nearly 50, with a long career in high school physics and chemistry teaching behind him. His career began with a powerful boost from PSSC and Harvard Project Physics teacher workshops in the heyday of Sputnik space-race fever. The influence was indelible. He has been a "hands-on" teacher ever since, always eager to build his own apparatus, and always looking for simple demonstrations of deep physics. He also retained a "spirit of adventure" in the physics classroom and a "spirit of kinship" with other physics teachers. This "spiritual imprint" of the PSSC workshops seems to have marked many of Malcolm's generation and sustained them through long careers as physics teachers. The lack of such spirit may contribute to the disturbing dropout rate among the younger teachers in recent years. Malcolm has always been sensitive to this problem. When he got the chance to conduct his own workshops later on, he spared no effort to nourish *camaraderie* among the teachers—even to the extent of rising early every day to purchase fresh donuts, out of his own pocket. He had the teachers bubbling at the coffee breaks in animated discussions about the details of their craft. He had them grappling with practical problems in the workshop sessions. *True camaraderie* comes from collaborative efforts on common problems; it is the strongest kind of professional glue—a source of professional pride and satisfaction. Physics teachers need it to cope with the professional *isolation* most of them face in their schools. They need it as a stimulus to improve; they need it for a sense of belonging. The sporadic successes of teacher workshops in meeting this need demonstrates the importance of permanent institutional mechanisms to support teacher interaction and professional development. Malcolm's work will lead us back to this issue later on.

Since Malcolm's high school is close to Arizona State University, over the years he was able to take every university course in science and education that was relevant to his teaching. When he excelled in physics and chemistry courses, his professors presumed that he would "leave teaching for a more challenging career"—a sad testament to the vision of professors. To understand the depth and richness of the teaching challenge, college professors should spend some time in classrooms or workshops run by superior teachers like Malcolm. With Malcolm's extensive academic background, he could have dashed off a thesis and obtained his doctorate from the college of education in a few months.

Instead, he came to me looking for a doctoral research project that would count as a genuine contribution to physics education. We discussed a variety of possibilities over several years before settling on one that satisfied us both. During this period he became familiar with the details of my educational research program, and I learned about his ceaseless efforts to improve his teaching.

Malcolm was among the first to use computers in high school physics. He did not wait for someone else to tell him how to do it. As soon as the Apple computer was available, he was writing his own programs and designing activities for his students to use on it. He had enough of this for a complete high school physics course when he came to me, so it was a natural subject for his dissertation. The main issue in our discussions was how to prove the pedagogical value of his activities and, more generally, how to establish sound principles for using computers in the physics classroom. Malcolm was hard pressed to come up with a suitable plan for his research until he was shocked by a sudden revelation about his own teaching in 1983.

At that time Ibrahim Halloun was compiling the statistics from our *Mechanics Diagnostic* test as part of his doctoral research. This test measures the difference between Newtonian concepts and the students' personal beliefs about the physical world. The published results[1] show that this difference is large, and conventional introductory physics courses are not effective at reducing the gap. Further, the results are independent of the instructor's qualifications and teaching style. These conclusions have been supported by many other studies since. When examining the *Mechanics Diagnostic* for the first time, most physics teachers think that the questions are too obvious to be informative; then they are shocked by the post-instruction scores of their own students. Malcolm was no exception. In fact, he was the first high school teacher to be confronted by such evidence.

Like many physics teachers, Malcolm is strict about maintaining high academic standards, and he is hard-nosed about requiring students to assume responsibility for their own knowledge. When confronted by an irate parent who demanded to know why his son had received an "*F*," Malcolm replied, "Because there is no lower grade!" Even so, Malcolm is realistic about student capabilities, and he assumes full responsibility for his own role in what they learn. When confronted by the dismal scores of his students on the *Diagnostic*, he soon concluded that the fault was in his teaching and set about doing better. Thus, he was finally launched on his doctoral research.

In his own teaching, Malcolm had already abandoned the traditional lecture-demonstration method in favor of a student-centered inquiry approach based on the *learning cycle* popularized by Robert Karplus.[2] He was thoroughly schooled in all aspects of the learning cycle from a course in "methods of science teaching" by Anton Lawson, who employed it extensively in his research and teaching.[3] Despite all this, the performance of Malcolm's students on the Mechanics Diagnostic was poor. In fact, later data show that it was no better than the typical result from traditional instruction. Malcolm did not try to rationalize this failure by pointing out that his method has many other advantages which are obvious to anyone observing his classes—that the students are captivated by the classroom activities and their capacity for independent investigation improves markedly over the course. Instead, Malcolm confirmed the results of the Diagnostic by interviewing the students himself. He concluded that his instructional method was missing something essential.

Malcolm soon saw how to improve his instruction by following the modeling approach under development at ASU. At that time in 1983, I had just drafted a long paper proposing a theory of physics instruction with modeling as the central theme.[4] Physics professors have told me that the paper is difficult to read, but in my extensive discussions with Malcolm I found that he had mastered every detail relevant to his teaching. His real genius, though, appeared when he implemented the theory. That will be discussed in a later section. Here we review the underlying ideas.

There are several reasons for adopting a modeling approach to physics instruction: First, because it brings instruction closer to emulating scientific practice. Second, because it addresses serious weaknesses in traditional instruction. Finally, as documented below, Malcolm's research gives it strong empirical support. The first two reasons have been discussed at length elsewhere,[4,5] but a brief review is in order here to explain Malcolm's motivation.

The crucial role of *mathematical models* in physics research and applications is common knowledge to practicing physicists. It should be surprising, therefore, that the general concept of a scientific *model* is scarcely recognized in physics textbooks, though their pages are chock-full of specific examples. Change is in the winds, however. In recent blue-ribbon proposals for wholesale reform of the K–12 science and math curriculum, *modeling* has been explicitly identified as a major theme.[6,7] It will be no easy task to implement this theme, but Malcolm Wells has taken the lead.

From the pedagogical perspective, a major reason for adopting the modeling approach is to help students develop a more coherent, flexible, and systematic understanding of physics. The knowledge that students acquire from traditional instruction tends to be fragmented and diffuse. To most students the physics course appears to be "one damn thing after another," so they are forced into rote methods to learn it. Soon they are overwhelmed by the accumulation of rote fragments, with disaffection as an inevitable consequence.

The modeling approach organizes the course content around a small number of *basic models*, such as the "harmonic oscillator" and the "particle subject to a constant force." These models describe *basic patterns* which appear ubiquitously in physical phenomena. Students become familiar with the structure and versatility of the models by employing them in a variety of situations. This includes applications to explain or predict physical phenomena as well as to design and interpret experiments. It also includes the construction of more complex models by modification of the basic models. Explicit emphasis on basic models focuses student attention on the *structure* of scientific knowledge as the basis for scientific understanding. Reduction of the essential course content to a small number of models greatly reduces the apparent complexity of the subject.

Besides a general plan for organizing course content, modeling theory supplied Malcolm with many other ingredients for instructional design. Without going into details given elsewhere,[4] three ingredients are worth mentioning here.

First, an analysis and explicit definition of *model*. The models in physics are conceptual representations of physical systems and processes. Specifications for defining a complete model are outlined in Box 1

Box 1: Model specification

Object / system	Model	
I. Organization	**System Schema**	
• composition • environment • connectivity	• (internal) constituents • (external) agents • connections	
II. Basic Properties	**Descriptors**	*(Examples)*
• intrinsic	• object variables	m, q, I
	• state variables	$\mathbf{x}, \mathbf{v}$
• interactive	• interaction variables	$\mathbf{F}, \mathbf{T}$
III. Structure (internal/external)	**Laws of interaction**	$F = GmM/r^2$
IV. Behavior (temporal structure)	**Laws of change**	$m\dot{\mathbf{v}} = \mathbf{F}$, $\dot{\mathbf{L}} = \mathbf{T}$

(details in Ref. 4). This includes an *interpretation* of the model, specifying how the structure of the model relates to the structure of the object or system it represents. Accordingly, a mathematical model is not fully specified until it has been supplied with an interpretation.

Second, an extensive discussion of qualitative reasoning and *representational tools*, especially *force diagrams and motion maps*. The main point being that such tools are essential for competent modeling and problem solving. The failure of students to learn this from conventional instruction has been established.

Third, a detailed analysis of the procedural knowledge involved in constructing and deploying scientific models, including a characterization of specific *modeling* stages. This provided Malcolm with the key to his chief instructional innovation, the *modeling cycle*. It enabled him to identify clearly what the learning cycle was missing, namely, detailed specification of the modeling processes and techniques involved. The modeling cycle is discussed in Sec. IV.

While Malcolm was getting started on his dissertation in 1983, Halloun and I were conducting a pedagogical experiment in the University Physics course at ASU. Although a detailed account of the experiment has been published,[8] reiteration of the main idea will put Malcolm's work in a broader context.

A primary objective of University Physics is to develop student problem-solving skills. The bane of traditional instruction is that most students cling to a "plug-and-chug" problem-solving strategy that severely limits their skill development. "Well-grounded" teachers are keenly aware that the key to effective problem solving is in the initial *qualitative analysis* of the problem, including the construction and use of suitable diagrams. Employing the traditional didactic approach, they demonstrate good technique in solving many problems, and they can explain their reasoning clearly when necessary. In my experience, such teachers are often nonplused or even angered by evidence suggesting that their approach is ineffective for the vast majority of students—insisting that their presentations are clear and thorough, so any failure reflects on the intelligence, attitude, or preparation of the students.

A different conclusion comes from considering the student viewpoint. The student sees that the "answer" to a problem invariably comes from plugging numbers into equations and chugging a little arithmetic. All that fluff about diagrams and "physical intuition" can be ignored. The key to problem solving is finding the "right equation" in which to plug the "given numbers." If the teacher is "fair" and the course is "well-organized," the right equation is easily extracted from a short list of equations for the "current topic." Exam preparation is reduced to memorizing the list for each topic to be covered. The effectiveness of this strategy is abundantly confirmed by good grades on homework and exams. It fails only when the teacher gets tricky. Tricky teachers are a pain!

Tricky teachers try to tell students that there is a better way than plug-and-chug. But what is it, exactly? They do not even have a name for it!

Modeling theory enables us to do better. My pedagogical experiment with Halloun instructed students in a sharp alternative to plug-and-chug called the *modeling method*. We take the position that the *complete solution* to every physics problem is actually a *model*, not, as often supposed, a mere number, the *answer* to some question posed in the problem. The model supplies the context which makes the answer meaningful. Without the model the significance of the answer (its numerical value, for instance) cannot be evaluated—which explains why plug-and-chuggers seldom question their unreasonable answers. We maintain that expert physicists always presume some model in their answer to a physics problem, though they may be unaware of that fact and seldom explicate the model fully. This suggests that problem-solving performance can be improved by instruction which insists on making the model in every problem explicit.

With the modeling method, every physics problem is solved by creating a model or, more often, adapting a known model to the specifications of the problem. Most problems in introductory physics are solved by deploying a small number of *basic models*. For example, all the standard projectile problems are solved by deploying a single *kinematic model*: the *particle with constant acceleration*. Students are thrilled when they realize this and thrilled again when they understand how all the models in mechanics can be generated by a single theory.

Our modeling method for problem solving is accompanied by a modeling method for teaching it. Implementation of the method in our pedagogical experiment[8] was constrained by the large course, lecture-recitation format at the university. My lectures deviated considerably from standard practice by expounding the modeling perspective exclusively, concentrating on thorough analysis of a small number of exemplary models and illustrating their deployment to solve problems. More subtle aspects of the method were implemented by Halloun in an experimental recitation section. He engaged students in group problem solving with the instructor as mediator. The critical role of the instructor in this process need not be described here, because it is so similar to Malcolm's approach. Results of our experiment will be compared with Malcolm's in Sec. II.

We think that the emphasis on solving textbook problems in physics courses is often excessive and misguided. It may even promote a distorted view of physics, because textbook problems are so artificial. In the modeling approach to instruction, *problem solving is secondary to modeling*. The modeling of physical systems raises all sorts of problems—problems which are more meaningful in the context of modeling than when they have been extracted and presented as textbook exercises—and problems which do not appear in textbooks at all. The modeling method may facilitate the solution of textbook problems by providing deeper physical insight. But it also supports a de-emphasis on textbook problems.

Malcolm developed a quite different or, rather, a complementary version of the modeling method—one which is laboratory based and adapted to scientific inquiry. It emphasizes the use of models to describe and explain physical phenomena rather than solve problems. It aims to teach modeling skills as the essential foundation for scientific inquiry. To accomplish this in a systematic fashion, Malcolm developed the *modeling cycle*, to be described in Sec. III.

In the implementations by both Halloun and Wells, the modeling method has a *student-centered* instructional design. This is believed to be critical to its success, because students must be *actively engaged* in the right kinds of activities to develop modeling skills. In both problem-solving and laboratory activities, students are required to articulate their plans and assumptions, explain their procedures, and justify their conclusions. The modeling method is unique in requiring the students to present and defend an *explicit model* as justification for their conclusions in every case. The instructor must be well prepared to consistently guide this process to a timely and satisfying closure. Specifically, the instructor must be (1) fully conversant with all aspects of the relevant models and (2) acutely aware of likely student misconceptions or knowledge deficiencies.

At last we are prepared to understand how Malcolm corrected the deficiency in his instructional method which was exposed by the Mechanics Diagnostic. As students are led to articulate their reasoning in the course of solving a problem or analyzing an experiment, their naive beliefs about the physical world surface naturally. Rather than dismiss these beliefs as incorrect, Malcolm learned to encourage students to elaborate them and evaluate their relevance to the issue at hand in collaborative discourse with other students. *In the context of modeling activities students have a framework for testing and correcting their own ideas*, especially in regard to relevance and coherence with other ideas.

To sharpen his skills for dealing with student misconceptions, Malcolm mastered the *taxonomy* developed by Halloun and Hestenes,[9] a systematic classification of naive beliefs about mechanics. He used the taxonomy for planning, to ensure that class activities would provide repeated opportunities for confronting all the serious misconceptions. He prepared an agenda of misconceptions to be addressed in connection with each activity. This preparation sensitized him to opportunities for addressing misconceptions in the course of student presentations and discussions.

Halloun made a similar use of the taxonomy in the limited domain of problem solving, but Malcolm had much more freedom to extend the modeling method in his high school course. He concentrated on developing techniques for improving the *quality of student discourse* about scientific subjects. Modeling theory supplied a clear goal: *scientific discourse* featuring the formulation, elaboration, evaluation, and application of well-defined models; discourse exhibiting a suitable mixture of qualitative and quantitative elements. In pursuit of this goal, Malcolm expanded the class time allotted to oral presentations by students. The time for *student postmortems* of laboratory activities was increased to a third of the total activity. The postmortem is devoted to analyzing and consolidating what the students have learned from the experiment. It seems likely that the most significant learning occurs in this period—at least, when the activity is guided with the skill of a teacher like Malcolm Wells.

To facilitate postmortems and other student presentations, Malcolm experimented with a variety of techniques. For example, he tried having students outline their presentations on "butcher paper" to be hung up for other students to see, but that proved to be awkward. Finally, he hit on a brilliant idea. He equipped student groups with "whiteboards." A *whiteboard* is a 24 in.×32 in. section of "kitchen and bath" paneling. It is easy to write and draw on it with colored dry markers, and it is easily erased. The whiteboard soon became an integral part of Malcolm's method.

Teaching students how to use the whiteboard effectively became an important subgoal. For Malcolm the whiteboard is an instrument for improving the quality of student discourse. In preparation for a presentation, student groups are encouraged to outline their model and supporting argument on the whiteboard. Evaluation of the presentation then includes an evaluation of the whiteboard display.

Besides the design and implementation of the instructional innovations already mentioned, Malcolm's research included a careful evaluation of actual results in the classroom. To that we turn next.

III. EVIDENCE FOR EFFECTIVENESS OF THE MODELING METHOD

In creating his version of the modeling method, Malcolm incorporated every good idea he could find—some from his own long experience, some from educational research. When evaluating educational innovation it is important to ascertain what the various factors contribute to improvements. This is difficult, not only because there are so many variables and practical constraints severely limit the possibilities for controlling them independently, but because a significant effect may come from combining separate factors which do not

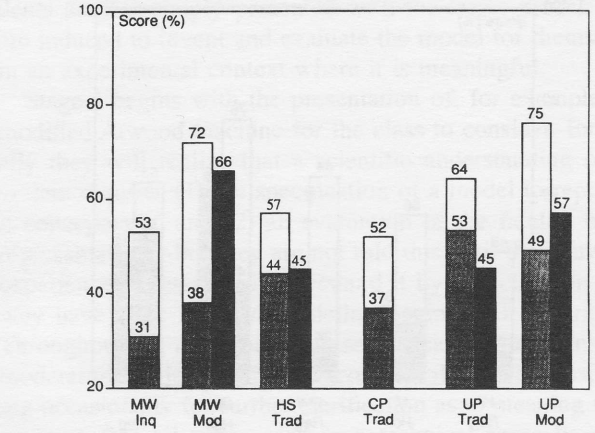

Fig. 1. Mean scores on the *Mechanics Diagnostic* for three high school classes (on the left) and three university classes (on the right). The bar chart shows the pretest score (at the beginning of the course) below the post-test score (at the end of the course). The dark bars are mean class scores on problem solving tests. See the text for explanation.

appear to contribute much alone. Fortunately, the unusual circumstances of Malcolm's doctoral research made it possible to achieve an exceptionally clean separation of the major factors contributing to his instructional results.

Figure 1 shows the impact of Malcolm's teaching in comparison with that of other teachers as measured by the Mechanics Diagnostic. Data on the high school courses come from Malcolm's thesis. The remaining data come from Refs. 1 and 8, which also provide an extensive analysis of the validity and implications of Diagnostic data. To interpret the data in Fig. 1, distinguishing features of the various instructional approaches must be identified. The three high school courses employed distinctly different approaches, which we describe by the terms "cooperative inquiry," "modeling method," and "traditional." We discuss each in turn and then compare their results.

Cooperative Inquiry has become increasingly popular in recent efforts to reform K–12 science education, and it is strongly advocated by educational researchers. The term is generally applied to any method of instruction with the following characteristics: It is student centered, activity oriented, and often laboratory based; students are actively engaged in investigating real phenomena in collaboration with their peers and under guidance by the instructor. Investigations are frequently organized into learning cycles by the teacher. All this fairly describes Malcolm's method in 1982–83—He was ahead of his time in this.

To be more specific about the content of Malcolm's inquiry course: 70% of class time was devoted to lab activities, which were either developed by Malcolm or modified from the Harvard Project Physics handbook. The lab activities targeted concepts involved in Newton's laws. Thirty percent of class time was devoted to in-class study groups utilizing the PSSC fourth-edition textbook. Problems for class and homework were selected from the textbook or designed by Malcolm to reinforce and expand on concepts developed in the lab activities.

Modeling Method. Malcolm's method at the close of his doctoral work (1986–87) can be described as *cooperative inquiry with modeling* structure and emphasis. He retained the general features of his original cooperative inquiry approach, including all the lab activities, to which he still devoted 70% of class time. The instructional difference resided in the systematic emphasis on models and modeling. The learning cycle was elaborated into a modeling cycle. Though it remained unobtrusive, teacher guidance was strengthened by focusing on a modeling agenda informed by the "misconceptions taxonomy." Consequently, student investigations and presentations were more coherently structured. The net result was an increase in the *coherence* of the whole course and its subject.

Traditional Method. The high school teacher who agreed to using his 1986–87 honors physics course as a control for comparison with Malcolm's course was well matched to Malcolm in regard to age, experience, training and dedication. He used a standard textbook [A. W. Smith and J. N. Cooper, *Elements of Physics*, (McGraw-Hill, New York, 1979) 9th ed.]. His course consisted of lectures and demonstrations (80% of class time), with homework questions and problems selected to reinforce important concepts from lecture and to provide practice in problem solving. There was a heavy emphasis on problem solving, with many examples worked out in lecture. Lab activities (20% of class time) were designed and/or selected to emphasize important concepts from lectures and/or to develop laboratory skills. In short, the course was quite traditional.

Comparisons. All three high school courses (inquiry, modeling, traditional) were honors courses with about 24 students in each. By prior agreement between the teachers, all three covered the same topics in mechanics on nearly the same time line (from early September until mid-March), so the total instructional time was the same.

The data in Fig. 1 strongly support the conclusions that Malcolm's modeling method is a considerable improvement over his cooperative inquiry method and clearly superior to the traditional method. In Diagnostic post-test score, the modeling class (MW Mod) surpasses the inquiry class (MW Inq) by 19% and the traditional class (HS Trad) by 15%. This is a large effect, because the standard deviation of student scores does not exceed 16% for any of the classes in Fig. 1. The inquiry class pretest score is exceptionally low for an honors physics class. However, it may be doubted that this accounts for any difference in the post-test scores. The pretest scores for both classes are so low (20% is a random score) that the difference cannot be attributed to more than superficial knowledge. For the same reason, the data do not show much difference between the inquiry and traditional methods, although inquiry produced a 9% greater gain.

These results should serve as a *warning* that the general approach of cooperative learning is not likely to improve student learning by itself. Improvement depends critically on the structure of the activities and the guidance by the teacher, so much so that, even for a superior teacher like Malcolm, results can be greatly improved by careful instructional design.

For comparison with Malcolm's score, Fig. 1 gives Diagnostic scores for traditional (algebra-based) College Physics (CP) and (calculus-based) University Physics (UP) courses. These courses were taught by the traditional lecture-demonstration method to classes with hundreds of students. One of the instructors has many awards for superior teaching. Nevertheless, as measured by the pre–post Diagnostic gains, neither course is more effective than the traditional high school course and both are far less effective than Malcolm's modeling course. Even on the final post-test Malcolm's high school students perform much better than the university students. Only Halloun's experimental modeling class (UP Mod) achieves a comparable result—which should not be surprising.

Problem solving. The modeling course was also compared to the traditional course with respect to student competency in traditional-type problem solving. For this purpose, a test was constructed consisting of 24 mechanics questions and problems from the 1983 NSTA–AAPT standardized examination, and 16 questions from PSSC and Harvard Project Physics tests. The problems were carefully selected to require some reasoning and some understanding of physics concepts, as opposed to being solvable by blind substitution into a formula. In this respect, it could be regarded as a "hard test." Otherwise, physics teachers would regard the test as fairly ordinary.

Since the traditional class had far more conventional problem solving practice, it might be expected to do better on the test. However, as Fig. 1 shows, Malcolm's modeling class outperformed the traditional class by 21%. How could this happen?

We have a definite answer which we can assert with much more confidence than Malcolm could in his thesis, because the result has been replicated many times since and detected with the more refined instruments described below.

The lower post-test score on the Mechanics Diagnostic (Fig. 1) means that the traditional class has a much weaker grasp of basic Newtonian concepts than the modeling class. In fact, at least half the class can be classified as pre-Newtonian (see discussion of Fig. 3). This means that those students are seriously deficient in basic concepts required for effective problem solving. Without those concepts, the students are forced to fall back on rote learning and plug-and-chug problem solving. Therefore, most of their problem-solving practice is a waste of time. Malcolm's approach concentrates on a thorough grounding in basic concepts first. Thereafter problem-solving skill develops more easily and surely. More evidence for this below.

Halloun's results in Fig. 1 support our conclusions about Malcolm's results. Although he was teaching problem solving directly, Halloun concentrated on identifying and correcting weaknesses in student grasp of basic concepts. Halloun's (UP Mod) class surpassed the traditional (UP Trad) class by 12% on a common problem-solving final exam (scores represented by dark bars in Fig. 1). More noteworthy is Halloun's success with underprepared students:[8] All such students in his recitation section passed the course with grade C or better, while 80% of the underprepared students in the traditional class failed to achieve at least a C grade, though there was a common grading system for both. This is comparable to Malcolm's achievement with high school students. It strongly supports the conclusion that traditional instruction fails miserably with underprepared students, though much better results are possible.

D.H. was so impressed with the results of Malcolm's thesis that he collaborated with Malcolm on a NSF grant to continue improving the method and develop workshops to pass it on to other teachers. The high school teacher who had acted as Malcolm's control was equally impressed and eagerly signed up for the first workshop. The experience revolutionized and rejuvenated his teaching, so he postponed his retirement.

The first task on the NSF grant was to improve the evaluation instruments. For this task Malcolm's intensive experience examining and applying the Mechanics Diagnostic and the misconceptions taxonomy was invaluable. The first result was the *Force Concept Inventory*,[10] which can be regarded as an improved version of the Mechanics Diagnostic. The second result was the *Mechanics Baseline* test,[11] which can be regarded as a greatly improved version of the problem-solving test that Malcolm used in his thesis. Details about the tests are given in the references. Here we are only interested in using test results for further documentation of Malcolm's achievements as a teacher.

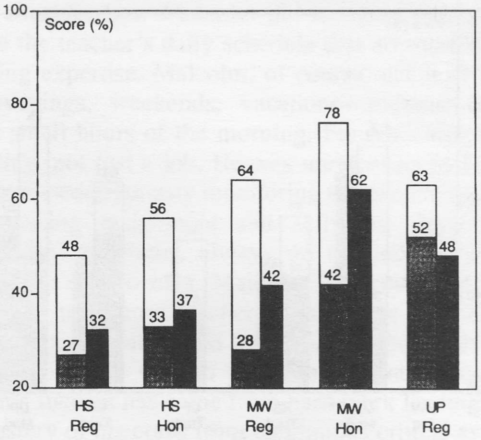

Fig. 2. Mean scores on the *Force Concept Inventory* and the *Mechanics Baseline* test. Pretest scores are displayed below the posttest scores for the *Inventory*. The dark bars represent *Baseline* scores.

The Inventory and Baseline tests provide a thorough and systematic evaluation of *basic conceptual understanding and problem-solving competence in mechanics*. They were published along with extensive data that have made it possible to compare the mechanics competence of physics students at every level from high school into graduate school. An enormous and rich body of data has accumulated since, and efforts are underway to analyze and organize it for informative publication.[12] It can be asserted, however, that the new data are generally consistent with the original data and so support the original conclusions.

Figure 2 is constructed from data in the original Inventory and Baseline papers. The scores for the traditional high school regular and honors physics courses are averages for more than 700 students and 17 different teachers. The dispersion of scores among the teachers is negligible, because it is much smaller than the dispersion among students in a single class. Unpublished data from other teachers give about the same result. We are quite confident in asserting that the scores in Fig. 2 *are typical for traditional* physics courses throughout the nation. Moreover, the small dispersion of scores for different teachers leads to the surprising conclusion that these typical scores are essentially *independent of the teacher's experience and academic background.* Data on university physics lead to much the same conclusion.[1,10] The scores for University Physics in Fig. 2 are for a single course. Again, consistent with our broader knowledge of the data, we regard these scores as typical for traditional University Physics courses at *large state universities.*

To summarize, the scores for traditional classes in Fig. 2 are typical and firm. Moreover, large variations in teacher expertise produce insignificant variations in student performance on the Inventory and Baseline tests. *Results of traditional instruction are uniformly poor for all teachers.* This suggests that instructional methodology is a more serious

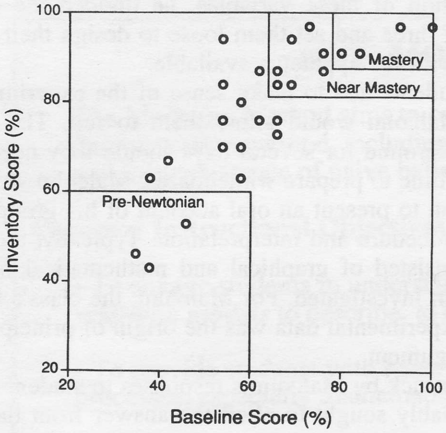

Fig. 3. Student competence after instruction in Malcolm's honors physics course. Post-test *Inventory* score is plotted against *Baseline* score for each of the 27 students.

problem than teacher competence. The good news is that the firm numbers in Fig. 2 provide a reliable baseline from which to measure the success of instructional innovation.

It should now be obvious that the scores in Fig. 2 document a remarkable achievement by Malcolm Wells, fully confirming the results of his thesis. Malcolm's superiority on this measure is so decisive that there is no need to describe the many other virtues of his method to be sure of its overall superiority. Malcolm's scores in Fig. 2 are for a single year, but unpublished data show that he achieved similar scores consistently year after year—with one exception. The scores fell one year when he was spending a lot of time on an experimental course at ASU. On seeing the results, he lamented "I wasn't minding the store!" This is indicative of his intense personal commitment to teaching.

Though Malcolm contributed heavily to the construction of the Inventory and Baseline tests, he scrupulously avoided teaching to the tests in his own courses. His scores were about the same, whether the tests were given immediately after the mechanics portion of the course or at the end of the spring semester. Thus, the retention of his students is strong.

Figure 3 gives the distribution of scores for students in Malcolm's honors course. A comparable figure for the University Physics course at Harvard is published in Ref. 10. Remarkably, the distributions for the two courses are very similar, though the Harvard course has four times as many students. Their mean scores on both tests are also about the same. Even for a group of first year physics graduate students at ASU, the mean scores are about the same as Malcolm's. Malcolm is in very good company indeed! He has given us an existence proof that high school physics students just about anywhere can be competitive with Harvard! There is no reason to believe that Malcolm had a special breed of student in his classes.

The details of Fig. 3 tell us more about Malcolm's impact. First note that all the data points lie above the diagonal. The reason for this is that the basic physics concepts (measured by the Inventory) are necessary but not sufficient for problem solving (measured by the Baseline). We refer to scores below 60% on the Inventory as *Pre-Newtonian*, because they indicate serious conceptual deficiencies, such as inability to discriminate reliably between velocity and acceleration. As data on Fig. 3 suggest, Pre-Newtonians are unable to score better than 60% on the Baseline. Scores in the box at the upper right-hand corner indicate genuine *mastery* of basic Newtonian mechanics. The "mastery box" is contained in a slightly larger *near mastery* box. Near mastery students are likely to be top physics students at any university they attend. More than a quarter of Malcolm's students fall within the near mastery box. Remarkably, this is more than the number of near mastery students from *all* 700 students in the traditional high school physics classes contributing to the data in Fig. 2. Malcolm's regular physics class also has several students in the near mastery box, though the full data will not be presented here. Malcolm's regular physics class differs from his honors class mainly in having a larger number of students stuck in the Pre-Newtonian box.

We have discussed Malcolm's case in such detail because there is a dearth of objective evidence for truly exceptional teaching and a lot of doubters that any such evidence exists. To our knowledge Malcolm's combined Inventory–Baseline scores have never been surpassed by any other high school teacher. But others are getting closer, and a few college teachers have surpassed him in absolute score, though not in fractional gain. Malcolm's mark is worth shooting at. We are sure that no one would be happier than Malcolm to see himself surpassed!!

IV. MALCOLM'S CLASSROOM

G.S. had the unique privilege of observing Malcolm's classroom in action over many months. G.S. had become intrigued with the possibilities of "modeling instruction" from published articles by D.H., so he arranged to spend sabbatical leave from his own high school physics teaching, with D.H. at ASU. He arrived just when Malcolm and D.H. had completed a preliminary version of the *Force Concept Inventory*, whereupon he was invited to join them in completing the job. His main task was to investigate the validity of the test through extensive interviews of high school students. This brought him to Malcolm's classroom for many hours, and he remained there for many more out of fascination. Here are his recollections of Malcolm's classroom, admittedly transmogrified by subsequent reflection and experience.

It was a November morning when I first visited Malcolm Well's classroom. The class was discussing a problem about the motion of an object subject to several forces. One student was holding up a whiteboard with a solution sketched on it. The board displayed clearly drawn diagrams with a few algebraic equations and some numbers. The class was gathered round as he explained his solution. An occasional question from another student was answered crisply. Relations between the diagrams and the algebraic statements were explained clearly. Substitution of the numbers into the algebraic statements was explicit. But Malcolm challenged the student further.

"Why did you do that?"

The student replied that he had identified and added all the forces along one dimension.

"Why did you do that?"

"So I could find the net force."

"Why did you do that?"

"Because $a = F/m$."

"How do you know that?"

"Because that's Newton's Second Law."

It was the first time that I had heard a student account for *everything* he had done in solving a problem, explaining why he had done it, and ultimately appealing to theory developed on the basis of experiments that had been done by the students. These students were explicit in their understanding. Malcolm did not take correct statements for granted. He always pressed for explicit articulation of understanding.

The students in Malcolm's class explained their solutions to problems publicly, and he made sure that they could justify them. He was uncanny in his ability to expose deficiencies in student explanations with questions. Many times I would have joyfully accepted a student's correct answer as sufficient. But Malcolm would again ask one more question, and, much to my surprise, the student would falter. This ability, as I gradually came to understand it, arose from his mastery of modeling in Newtonian physics. His understanding extended beyond the content of Newton's Laws to an acute awareness of the techniques for applying the laws in practice.

Malcolm was alerted when a student failed to mention the procedures required to be faithful to Newtonian physics. He would ask for elaboration at the very point where I was satisfied that the student had achieved the desired result. His deep understanding of scientific explanation and justification enabled Malcolm to be a remarkable Socratic guide. He had clear knowledge of what students had to make explicit to be assured that their understanding is adequate. His line of questioning was unfailingly purposeful. Students were required to present an explicit model to account for the physical situation in question and explain how the model had been obtained from overarching theory and/or experimental data. His students became accustomed to supplying not just answers and clear explanations of how they got them, but also full justification for their approach. The students' solutions to physics problems were superior.

The students were busy in Malcolm's classes. Working in groups of three they performed experiments, solved problems, explored activities. Regularly, Malcolm would assemble them to present accounts of their work orally with the aid of whiteboards or join in questioning the presenters. Whiteboards were new to me. Student groups prepared them with care and pride. With colorful dry markers they dressed the whiteboard with diagrammatic, graphical, and mathematical representations of physical situations from problems or lab activities. By the time I visited the class, students were consistently referring to these representations as *models*. They were using these models to solve problems or interpret experiments, and they could explain how the various representations cohere in their interpretations. The dialog during oral presentations was potent, whether the presentation was consistent with Newtonian physics or not. Students found holes in their understanding and honed their arguments, both by questioning one another and providing answers. Malcolm served as Socratic guide to keep the dialog moving in a profitable direction.

Another feature of Malcolm's teaching that was new for me was the solid experimental underpinning for all theoretical constructions that followed. Malcolm had adapted and designed experiments which were conceptually clean, with equipment enabling students to generate good data reliably. The students were given no instructions for doing these experiments. Rather, Malcolm would introduce the class to the physical system to be investigated and engage the students in describing the system until a consensus was achieved. Malcolm would stealthily elicit from his students the appropriate dependent and independent variables to characterize the system. After obtaining reasoned defenses from the students for the selection of these variables, he divided the class into groups of three and set them loose to design their own procedures with the apparatus available.

The students had to make sense of the experiment themselves. Malcolm would allow them to fail. The apparatus would be around for several days should they need it. After allowing time to prepare whiteboards, Malcolm would select one person to present an oral account of his group's experimental procedure and interpretation. Typically, the interpretation consisted of graphical and mathematical models for the system investigated. For Malcolm, the class's interpretation of experimental data was the origin of principle and the end of argument.

I was struck by Malcolm's responses to student questions. He invariably sought to elicit the answer from the students themselves, and to induce them to assume responsibility for their own explanations. Sometimes, when students were thoroughly nonplused, he would suggest that they find out what other students were doing. Malcolm assiduously avoided the role of authority—this was a matter of principle with him. The belief that learning science is acceptance of what the text or teacher declares was regarded by Malcolm as an obstacle to valid understanding by the students. In this respect he stands with Feynman, who said that "science is a belief in the ignorance of experts." The struggle for understanding was fostered and facilitated by Malcolm, but never mitigated.

Computers played a prominent role in Malcolm's classroom, but that role was defined by Malcolm's pedagogy. Computers became tools for analyzing experimental data and for simulating physical systems when real, clean, and reliable experiments were not available. Computers helped students create good models of physical systems and generalize their results into theoretical statements. They helped provide the physical theory developed in the course with a firm experimental foundation to which the students continually had to appeal to justify their work. Computers were not just a nice addition to the course, they were indispensable. The foundational experiments that Malcolm used to span the desired dimensions of physics could not have been done without them. Never had I seen computers used so effectively and frequently to facilitate the struggle for understanding.

As exhibited in his classroom, Malcolm's method has a clear moral: Teaching by telling is ineffective. Coherent understanding cannot be transferred from teacher to student by lucid explanations or brilliant demonstrations. Students construct their own understanding. The teacher is a facilitator. Malcolm labored to guide students to a coherent and, therefore, lasting understanding of physics. He sought to change their view of learning from collectors of information to expectant creators of this coherent understanding. He was more concerned with what students would think about his course five years later than with what they thought about it during the school year. To Malcolm it must have been the ultimate tribute when one of his former students gave thanks not for teaching him what to think but how to think!

V. MODELING CYCLE

The atmosphere in Malcolm's classroom was not simply the product of a talented teacher doing his stuff. It was the result of careful preparation, planning, and deliberate execution of a definite method. Let us describe his method in more detail.

A synopsis of the *modeling method* is enclosed in Box 2.

Box 2: MODELING METHOD Synopsis

The Modeling Method aims to correct many weaknesses of the traditional lecture-demonstration method, including the fragmentation of knowledge, student passivity, and the persistence of naive beliefs about the physical world.

Coherent Instructional Objectives

- **To engage students in understanding the physical world by *constructing and using scientific models* to describe, to explain, to predict and to control physical phenomena.**
- **To provide students with *basic conceptual tools* for modeling physical objects and processes, especially mathematical, graphical and diagrammatic representations.**
- **To familiarize students with a small set of basic models as the *content core* of physics.**
- **To develop insight into the *structure* of scientific knowledge by examining how *models* fit into *theories.***
- **To show how scientific knowledge is *validated* by engaging students in *evaluating* scientific models through comparison with empirical data.**
- **To develop skill in all aspects of modeling as the *procedural core* of scientific knowledge.**

Student-Centered Instructional Design

- **Instruction is organized into *modeling cycles* which move students through all phases of model development, evaluation and application in concrete situations — thus promoting an integrated understanding of modeling processes and acquisition of coordinated modeling skills.**
- **The teacher sets the stage for student activities, typically with a demonstration and class discussion to establish common understanding of a question to be asked of nature. Then, in small groups, students *collaborate* in planning and conducting experiments to answer or clarify the question.**
- **Students are required to present and justify their conclusions in oral and/or written form, including a *formulation* of models for the phenomena in question and *evaluation* of the models by comparison with data.**
- **Technical terms and concepts are introduced by the teacher only as they are needed to sharpen models, facilitate modeling activities and improve the quality of discourse.**
- **The teacher is prepared with a definite *agenda* for student progress and *guides* student inquiry and discussion in that direction with "Socratic" questioning and remarks.**
- **The teacher is equipped with a *taxonomy* of typical student misconceptions to be addressed as students are induced to articulate, analyze and justify their personal beliefs.**

The instructional objectives are appropriate for any implementation of a modeling approach to instruction. The instructional design is more specific to Malcolm's inquiry approach. The centerpiece of this design is the modeling cycle, which organizes class activities into *coherent* units with similar procedural structure.

The modeling cycle can be regarded as a refinement of the *learning cycle* developed by physicist Robert Karplus for the *Science Curriculum Improvement Study* (SCIS). It greatly elaborates the role of models and modeling in the cycle. We have recently heard from Anton Lawson that there was an unresolved debate among scientists on the SCIS development team as to whether *models* or *theories* should play the central role in the curriculum. Biologist Chester Lawson championed *theory* while Karplus was firmly in favor of *models*, though, in deference to his colleague, he allowed his position to be somewhat diluted in the curriculum. The Karplus view has been keenly described by Victor Pollock.[13] We believe Karplus would come out strongly in favor of the modeling cycle if he were around today.

Before describing the modeling cycle, let us briefly review the three stages of the learning cycle (*exploration, invention, discovery*) from a modeling perspective.

Exploration. Typically, in this stage students are given some physical phenomenon to investigate with hands-on activities. Students are given minimal guidance so they can make their own observations and formulate their own conclusions. The main instructional difficulty with this stage is that it tends to degenerate into aimless "messing about" under too little guidance or become unimaginative under too much. The modeling method resolves this difficulty over several cycles by teaching students a *general method of scientific inquiry*. Students learn that in every investigation it is essential to develop a *model* of the physical system, and they continue to grow in their understanding of what modeling involves. When investigating some general physical concept like "energy conservation," they learn that it cannot be explored experimentally apart from a specific model. The model supplies a *context* for the exploration. Thus, in investigating a new phenomenon, students learn to focus quickly on identifying particular systems to be modeled and on quantitative measures of their properties.

Invention (or concept introduction). This stage recognizes that modeling cannot go beyond simple description without the *invention* of new concepts and symbolic tools to represent them. Chief among these are the inventions of algebra and calculus, which make it possible to formulate quantitative relations among variables. The mathematical tools make it possible to formulate "universal" principles like Newton's Laws, which facilitate mechanics modeling in (nearly) every situation.[4,5]

Students cannot be expected to invent the concepts and notations introduced in this stage. But they must discover for themselves the utility of the concepts for modeling phenomena from the exploration stage. From the modeling perspective, that is the main objective of the invention stage.

The stage name "concept introduction" is usually preferred over "invention," because it is supposed to be more descriptive of what is actually done. However, that very name may encourage the serious pedagogical mistake of introducing concepts piecemeal and out or context, in the misguided belief that complexities are mastered by concentrating on one concept at a time. The very strength of the learning cycle is that *new concepts are introduced within the context of modeling and for the purpose of modeling*. The modeling approach makes this explicit. The emphasis on models rather than single concepts makes instruction more *coherent*, for model construction requires the coordinated use of a whole set of concepts.

The new concepts introduced in this stage are usually nontrivial and fully deserve to be recognized as inventions, often great inventions! Students and teachers need to appreciate the power that such inventions confer on the user. For this reason, we think the stage name invention is well chosen.

Discovery (or concept application). Likewise, we prefer the original name "discovery" for this stage. It is not usually a single concept that is applied in this stage, but the whole model that was developed in the first two stages. The model is abstracted from its original physical context and applied to new situations. The applications often require genuine (though not original) discoveries by the student, so why not celebrate that with the word discovery? Rather than "model applications," we speak of "model deployment" below, to emphasize strategic and tactical aspects of modeling which are not so straightforward as the term "application" suggests.

Now let us turn to the modeling cycle. The modeling cycle has two stages, involving the two general classes of modeling activities: *Model development* and *model deployment* (See Refs. 4 and 5 for more details). Roughly speaking, model development encompasses the exploration and invention stages of the learning cycle, while model deployment corresponds to the discovery stage. It will be noted that the "modeling terminology" is more descriptive of what the students actually do in the cycle.

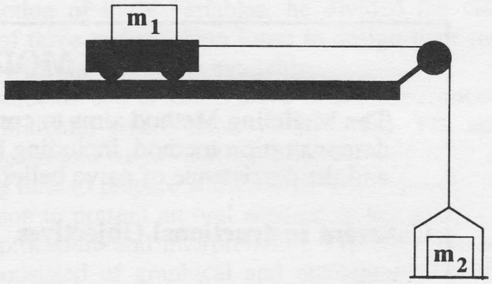

Fig. 4. *Modified Atwood Machine*. A (low friction) cart with mass m_1 is connected to a mass m_2 suspended by a string over a pulley. Values of the masses can be varied. The diagram raises some questions about accounting for the masses.

The *two-stage modeling cycle* has a generic and flexible format which can be adapted to any physics topic. In its high school physics implementation, the cycle is two or three weeks long, with at least a week devoted to each stage, and there are six cycles in a semester, each devoted to a major topic. Each topic is centered on the development and deployment of a well-defined mathematical model, including investigations of empirical implications and general physical principles involved.

Throughout the modeling cycle the teacher has a definite agenda and specific objectives for every class activity, including concepts and terminology to be introduced, conclusions to be reached, issues to be raised, and misconceptions to be addressed. Though the teacher sets the goals of instruction and controls the agenda, this is done unobtrusively. The teacher assumes the roles of activity facilitator, Socratic inquisitor, and arbiter (more the role of a physics coach than a traditional teacher). To the students, the skilled teacher is *transparent*, appearing primarily as a facilitator of student goals and agendas.

To make the present discussion of details in the modeling cycle more concrete, we choose a specific topic which appears in both high school and university physics courses. Accordingly, as major objectives for the instructional agenda in the cycle, we aim to develop student conceptual understanding of the following: *Target model*: Motion of a material particle subject to a constant force. *Physical principle*: Newton's second law of motion. *Experimental context*: Modified Atwood's machine (Fig. 4).

Prerequisite: Before beginning this cycle, the students should have previous experience with kinematic models (two cycles in the high school course), so they have fairly clear concepts of velocity and acceleration. *Many students still have only a shaky grasp of these concepts at this point*, and more experience with the concepts in a variety of contexts is necessary to consolidate them. Conceptual development takes time, and it will be haphazard unless instruction is carefully designed to promote it systematically.

Stage I is designed to lead students systematically through the four main *phases of model development: description, formulation, ramification and validation* (Refs. 4 and 5), though students are not introduced to this fancy terminology. Students are not simply presented with the target model; they are induced to invent and evaluate the model for themselves in an experimental context where it is meaningful.

Stage I begins with the presentation of, for example, the modified Atwood machine for the class to consider. Eventually they will realize that a scientific understanding of the system requires (1) the specification of a model to represent it conceptually, and (2) an evaluation of the fidelity of the representation—but they are not told this until they have the experience necessary to understand it by reflecting on what they have done already. Modeling begins with description. Throughout the *descriptive phase* the teacher functions as a moderator, nonjudgmentally recording all suggestions, asking occasionally for further clarification as to meaning while insisting that all terms used in a technical sense be given valid operational definitions. Technical terms, such as "frame of reference, one-dimensional motion, and system" are introduced by the instructor *only* in situations where they serve to clarify the discussion. Ample opportunity to introduce important technical terms occurs as the course proceeds. Beginning students may state, for example, that an object is accelerating but when asked what they mean by acceleration, they often reply "speeding up." The teacher continues to ask probing questions until the students articulate a satisfactory quantitative characterization of the concept. The teacher strives to remain unobtrusively in control of the agenda throughout the discussion, never acting as an authority or a source of knowledge.

At the conclusion of the descriptive stage, the students are directed, collectively, to identify quantitatively measurable parameters that might be expected to exhibit some cause–effect relationship. A variable under direct control by the experimenters is identified as the *independent variable*, while the effect is identified as the *dependent variable*. This is a critical step in the modeling process. It is at this point that the students learn to differentiate aspects of the phenomenon to which they must attend from those which are distracters. While this issue of *identifying and controlling variables* is critical to modeling, it is scarcely addressed in traditional instruction, where a lab manual typically provides students with the lab purpose, procedure, evaluation of data, and even questions suggesting appropriate conclusions. This critical issue is also missed in conventional homework and test problems, which typically provide only that information necessary to accommodate the *author's choice* of solutions.

Having completed the *descriptive phase* of modeling by settling on a suitable set of descriptive variables, the instructor guides the class into the *formulation phase* by raising the central problem: to develop a functional relationship between the specified variables. A *brief* class discussion of the essential elements of the experimental design (which parameters will be held constant and which will be varied) is pursued at this time. The class then divides into teams of two or three to devise and perform experiments of their own.

Before starting data acquisition, each team must develop a detailed experimental design. Except where the design might pose risk of injury to persons or equipment, the teams are permitted to pursue their own experimental procedures without intrusion by the instructor. For a post-lab presentation to the class, the instructor selects a group which is likely to raise significant issues for class discussion—often a group that has taken an inappropriate approach. At that time, the group members are expected to present a detailed explanation and defense of their experimental design and conclusions.

Each lab team performs its own data analysis cooperatively, using computers and striving to construct graphical and mathematical representations of the functional relationships previously posited. The principal goal of the laboratory activities is to lead students to develop a conceptual correspondence between targeted aspects of the real world phenomenon and corresponding symbolic representations.

Every lab activity is concluded by each lab team preparing, on a whiteboard, a detailed post-lab analysis of the activity and reasoning that led to the proposed model(s). The teacher then selects one or more of the lab groups to make presentations before the class, explaining and defending their experimental design, analysis of data, and proposed model.

Laboratory reports for each activity are written up in a laboratory notebook according to a given format. It is stressed that the purpose of the laboratory report is to articulate a coherent argument in support of their model construction. While each student must prepare and submit a lab notebook, most of the work is done in class in their cooperative study groups. Grading is done by selecting one report at random from each group and selecting different members of the group to defend different aspects of the report. This induces students, during the preparation of reports by the groups, to ensure that every member of the group understands all aspects of the model that they have developed, thus instilling a sense of *shared responsibility* for the knowledge. This concludes Stage I.

The end product of Stage I is a mathematical model together with evidence for a claim that accurately represents the behavior (or structure) of some physical system, in this case the Modified Atwood's Machine. Students have verified that the equation $a=F/m$ accurately describes the acceleration when F and m are varied independently. They are encouraged to consider the possibility that this equation represents a general law of nature, but they should be led to realize that there is *no such thing as an experimental proof of a general law*. At best, experiment can validate specific models which conform to the law, as in the present case.

Stage II is devoted to *deployment* of the model developed in Stage I to a variety of new physical situations in a variety of different ways. This helps free the students' understanding of the model from the specific context in which it was developed. The model may be deployed to describe, to explain, to predict, or to design a new experiment. Though some of the activities in Stage II involve the laboratory, most are more like traditional problem solving, except the work is done cooperatively in small groups. Most of the work is done in class.

Each study group develops solutions for each problem in the study set. Each group is then assigned one of the problems in the set to prepare, on the whiteboards, for class presentation. One member of the group is then selected to make the presentation. The same recitation grade is given to the entire group, and it depends on the quality of the presentation. During the presentation, if questions are asked by fellow students that the selected presenter cannot answer, other members of the group may offer assistance. If however any assistance from other members of the group is required to satisfy the questioner, the recitation grade awarded the group may be reduced. The recitation scores of the groups are enhanced if the members ask valid, well thought out questions during the presentations (*shared responsibility*).

On each pass through the modeling cycle the students' understanding of models and modeling is progressively deepened; students become more independent in formulating and executing tasks and more articulate in presenting and defending their points of view. The ultimate objective is, of course, to have them become autonomous scientific thinkers, fluent in the vicissitudes of mathematical modeling.

VI. CULTIVATION OF TEACHING EXPERTISE

What does it take to become a master teacher like Malcolm Wells? The skill and training required for expert teaching are generally underestimated and undervalued. Accolades and awards for teaching are often based on superficial criteria. Malcolm's example sets a higher standard—one to be emulated if teaching is to be elevated.

An extensive review and analysis of the literature on expert performance has identified essential conditions for the acquisition of expert skill in most domains.[14] The chief condition is *prolonged effort to improve performance* extending for a minimum of 10 years. A striking conclusion of the study is that individual differences, even among elite performers, are primarily due to intense practice rather than innate talent. Music, sports, chess, scientific research and literature are among the several domains examined in the study. Teaching was not included, of course, but there is no reason to doubt that the general conditions for acquisition of expertise apply there as well. Assuming so, we can draw some important conclusions about the professional development of teachers.

Our first conclusion is that *standard teacher preparation and in-service teaching experience is not sufficient to develop a high level of teaching expertise*. Consider what is involved. Even assuming that a physics teacher has acquired adequate "content knowledge" from a B.A. or even an M.A. in physics, the relevant pedagogical training is practically nil. After landing a teaching position, the tyro teacher may scramble for a couple of years to organize lab materials and activities, problem sets and homework, grading procedures and the rest into a smoothly running course. By this time the teacher has adopted a personal style and a teaching routine which makes it possible to cope with the perpetual exigencies of everyday teaching.

Most physics teachers are dedicated to their job and care deeply about their students. But caring and dedication are not enough! The experience of routine teaching over many years, even when conducted with dedication and enthusiasm, will not contribute significantly to the development of teaching expertise—just as plug-and-chug practice does little to promote problem solving skill! There is strong empirical support for this kind of assertion from the domain of chess.[5,14] Tournament chess players are assigned numerical performance ratings which are extremely reliable predictors of their tournament results. The fact is that, after an initial increase when learning the game, the average rating of an avid amateur scarcely changes over the years no matter how many games are played. Thus routine chess playing does not improve chess competence. Likewise, we conclude, *routine teaching does not improve teaching competence*. Most teachers become trapped in a routine that prevents them from coming close to realizing their true potential.

How to rise above it?

First consider how Malcolm did it. The schools have so crowded the teacher's daily schedule that no room is left for cultivating expertise. Malcolm, of course, did it on his own time—evenings, weekends, vacations—routinely working into the small hours of the morning. For Malcolm, teaching is a calling, not just a job. He was unrelenting in his efforts to improve—continuously monitoring the progress of his students, revising assignment and activities, designing and building new apparatus, always on the lookout for some other teacher's good idea. Malcolm is a counterexample to the myth of the "born teacher." Unlike the typical award-winning teacher, Malcolm is not a master showman. Rather, he goes out of his way to give the students center stage. Malcolm's success has come from hard work leading to technical mastery of his craft, from continuous critical evaluation of his own teaching performance, and from meticulous attention to every detail, large and small. "The devil is in the details!"

Few can match the prolonged and dedicated effort of Malcolm Wells, but many can aspire to his level of teaching expertise, because Malcolm has prepared the way. This paper aims to pass on some of Malcolm's hard won insights. However, most of Malcolm's expertise is bound up in teaching skills. Such skills cannot be transmitted verbally; they can only be passed on through personal interaction and deliberate practice in the classroom.

To develop a practical means for training teachers in the modeling method, we joined Malcolm in designing and conducting a series of NSF summer workshops for in-service teachers. A brief account of the experience provides some background for future action.

Two groups of high school physics teachers participated in the project. In teaching experience they ranged from novice to state teacher of the year, and in academic background, from one year of College Physics to a Masters in physics education. The first group of 17 teachers attended five-week workshops in the summers of 1990 and 1991 with a follow-up one-week workshop in the summer of 1992. They were also brought together for half-day workshops at regular intervals during the school year to discuss progress and problems with implementing the new method. All the teachers employed the new method in their regular high school physics teaching during their two years with the project, and they have continued using it since.

After initial hesitancy in the first workshop, teacher enthusiasm for the new modeling method grew to a stupendous level by the middle of their first year of teaching with it, and all teachers reported big improvements in student interest and activity. By the usual anecdotal measures the program was a great success. However, the *Force Concept Inventory* gave us an objective measure of gain in teaching effectiveness by comparing the score of each teacher's class just before the workshop with the one just after. The result was a sobering 4%—barely significant! We could identify several reasons for the limited gain: (1) The written curriculum materials tailored to the new method were inadequate; (2) the teachers were so caught up in the mechanics of the computer-based laboratory activities that they overlooked crucial pedagogical features that make the method effective, and (3) too much lecturing about the method (shame!).

In the second summer workshop, the teachers were involved in developing the necessary curriculum materials, and this gave them a satisfying sense of ownership in the program as well as rich experience collaborating with their peers. Also, pedagogical techniques were given renewed emphasis. This contributed to a clearly significant 22% average gain on the *Inventory* for all teachers. That, however, is still well short of the results consistently achieved by Malcolm Wells. Moreover, though there was some improvement on our other measure of student competence, the *Mechanics Baseline*, it is not worth reporting.

In the summer of 1992, a new group of 14 teachers attended a single five-week "Modeling Workshop." With the printed curriculum materials available, this workshop proceeded more smoothly and quickly than the previous ones. Most important, the workshop design was improved to enable the teachers to practice the new methods on their colleagues almost every day. From our personal observations, we are confident that this new group made as much progress in one summer as the original group did in two. Unfortunately, we were unable to validate this conclusion with an objective follow-up evaluation.

Overall, we regard the workshops as moderately successful. The teachers were unanimous in high praise for the experience. As a consequence, all of them have radically and permanently changed their teaching methods. As far as we know, their teaching is now laboratory based, computer enhanced, student centered, and activity oriented. They report that their students are more engaged and enthusiastic than ever. They are especially delighted with the enhanced student participation stimulated by the whiteboards. In short, the workshops succeeded fully in getting teachers to adopt a *cooperative inquiry* method of teaching. They were less successful in leading teachers to understand the rationale for the

Box 3: MODELING WORKSHOP Description

Participants will be introduced to the Modeling Method as a systematic approach to the design of curriculum and instruction.

- **They will collaborate on the redesign of the high school physics course to enhance learnability and exploit technology.**
- **They will learn how to use computers and electronic networks as an integral part of their teaching practice.**
- **They will implement a student-centered instructional strategy which engages students in active *scientific inquiry, discourse and evaluation of evidence.***
- **They will examine implications of educational research for physics teaching.**

CURRICULUM

- **Standard** topics will be covered (including mechanics, optics, electricity and magnetism), but they will be organized into a systematic and coherent curriculum.
- **Flexible** curriculum design will facilitate future upgrades of computers and software and incorporation of new topics or activities.
- **Structured** curriculum for the introductory physics course will be supplemented by a **project-oriented** curriculum for an advanced course or extracurricular activity.

INSTRUCTION

- Since "teachers teach as they have been taught," workshops will include extensive practice in implementing the curriculum as intended for high school classes.
- Participants will rotate through roles of student and instructor as they practice techniques of guided discovery and cooperative learning.
- Plans and techniques for raising the *level of discourse* in classroom discussions and student presentations will be emphasized.

modeling method. For example, a video of one teacher's class shows enthusiastic students in intense and animated discussion over a whiteboard, but the teacher failed to focus the discussion, so it went nowhere. Another teacher inadvertantly subverts the objectives of guided-inquiry lab experiments by summarizing the findings instead of requiring the students to do so. On the other hand, the *Inventory* scores show that the teachers have been greatly sensitized to student misconceptions and are learning to address them; although only a few of them have learned to appreciate the deeper aspects of the modeling method. This is reflected in the minimal improvements of *Baseline* scores. Considerable advances in workshop design and execution will be needed to achieve a satisfying outcome along this dimension.

We are now prepared to draw some strong conclusions about what is most needed to improve high school physics. *Teacher expertise is the critical factor*. The teacher, above all, determines the quality of student experience in the classroom. Equipment and school environment are secondary factors. To reach and maintain his/her full potential, the *teacher must be engaged in lifelong professional development*. It will take at least ten years to reach the teacher's highest level of competence. Mere accumulation of academic credits and hours of classroom teaching count for little, unless the teacher is consistently engaged in deliberate effort to improve.

Teacher commitment is essential, and individual teachers, like Malcolm, can go far in designing and executing their own programs for personal development. However, even Malcolm needed help to reach his peak, so the *ultimate success of every teacher depends on opportunities to draw on the resources of the physics community*. Teachers need a support system in the physics community to nourish their professional development. The infrastructure for such support is in terrible shape across the nation.

From many quarters, especially the National Science Foundation, we hear a clarion call for nationwide systemic reform of science and math education. It signals widespread recognition of a need to rebuild the educational infrastructure. But systemic reform will fail unless it focuses on developing and sustaining teacher expertise. This is a problem of immense proportions, but we need not wait for someone else to attack it. The physics community must assume responsibility for establishing and maintaining an infrastructure for high school physics reform. To be fully successful it must be a collaborative effort involving all segments of the physics community—in high schools, colleges, universities, and professional societies. Here is how we propose to attack the problem.

We have recently been awarded a NSF grant to conduct a nationwide program of *Modeling Workshops* for in-service high school physics teachers beginning in summer 1995. Besides the authors, the Project team includes Larry Dukerich, Ibrahim Halloun, and Jane Jackson. The Workshop is described in Box 3.

It builds on the design pioneered by Malcolm Wells, and it is aimed at cultivating Wells-like expertise among teachers. We are *dedicated to using the Modeling Workshop as an instrument for high school physics reform*. We are keenly aware that the impact of the program depends critically on the dedication and local support of the participants. Consequently, participation in the first round of workshops is competitive, with preference to applications showing the most promise for local reform. If the first round is successful, we have plans and funding to expand the program, and we would like nothing better than to make the workshop available to all interested teachers. For further information about the program, write the *Modeling Workshop Project Director*, Dr. Jane Jackson, at D.H.'s address.

MALCOLM WELLS has started something!

ACKNOWLEDGMENT

This work has been supported by grants from the National Science Foundation. Note: Malcolm Wells died on July 20, 1994, two weeks after this paper was completed.

[1] I. Halloun and D. Hestenes, "The initial knowledge state of college Physics students," Am. J. Phys. **53**, 1043–1055 (1985).

[2] R. Karplus, "Science Teaching and the development of reasoning," J. Res. Sci. Teach. **14**, 169–175 (1977).

[3] A. Lawson, M. Abraham, and J. Renner, "A Theory of Instruction: Using the Learning Cycle to Teach Science Concepts and Thinking Skills," NARST (1989).

[4] D. Hestenes, "A modeling theory of Physics instruction," Am. J. Phys. **55**, 440–454 (1987).

[5] D. Hestenes, "Modeling games in the Newtonian world," Am. J. Phys. **60**, 732–748 (1992).

[6] F. J. Rutherford and A. Ahlgren, *SCIENCE for all Americans* (AAAS Project 2061) (Oxford University Press, Oxford, 1990).

[7] L. Steen, *On the Shoulders of Giants* (National Academy, Washington D.C., 1990), especially pp. 61–94.

[8] I. Halloun and D. Hestenes, "Modeling instruction in mechanics," Am. J. Phys. **55**, 455–462 (1987).

[9] I. Halloun and D. Hestenes, "Common sense concepts about motion," Am. J. Phys. **53**, 1056–1065 (1985).

[10] D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," The Phys. Teach. **30**, 141–158 (1992).

[11] D. Hestenes and M. Wells, "A mechanics baseline test," The Phys. Teach. **30**, 159–166 (1992).

[12] R. Hake, "Survey of test data for introductory mechanics courses," AAPT Announcer **2** (2), 55 (1994).

[13] V. L. Pollock, "Science education II: Scientific literacy and the Karplus taxonomy," J. Sci. Ed. Technol. **3**, 89–97 (1994).

[14] K. Ericsson and N. Charness, "Expert performance, its structure and acquisition," Am. Psychologist **49**, 725–47 (1994).