Project Description

Previous NSF Support: David Meltzer has received NSF support for four projects during the past five years. Two of these supported the development of an elementary physics course targeted at education majors and other nontechnical students. (See Biography section for grant numbers and titles.) These projects involved the development and testing of a one-semester activity-centered elementary physics course based on guided-inquiry instructional methods, given at Southeastern Louisiana University (SLU). The bottom-line result of these projects may be succinctly stated: Before 1994, virtually *no* presecondary education majors ever set foot in a physics classroom at SLU. In the 1997-98 academic year, by contrast, *twenty-one* education majors (and eight other nontechnical students) successfully completed this physics course at SLU, with significant learning gains as documented by rigorous assessment and an overwhelmingly positive student response. An independent study prepared for the NSF's Division of Undergraduate Education described the immediate predecessor of these projects as "very successful" (Whalen *et al*., 1996). In June of 2000, Meltzer (along with Co-P.I. Thomas Greenbowe) received a CCLI-EMD award for the development of active-learning curricular materials in thermodynamics. We have carried out preliminary development and testing of sample curricular materials, have made numerous conference presentations (e.g., Meltzer and Greenbowe, 2001; Greenbowe and Meltzer, 2001), published two papers (Meltzer, 2001; Greenbowe and Meltzer, *in press*) and submitted a third (Meltzer, 2002) regarding our investigations of student learning of thermodynamic concepts which will form the basis for our ongoing curriculum development efforts. In July 2002, Meltzer (PI) and Greenbowe (Co-PI) were awarded an NSF grant in the "Research on Learning and Education Program" to investigate the role of representational modes in student learning of physics and chemistry.

Outline of project: The goal of this project is to develop formative assessment materials for the introductory general physics course which are to be used in the context of a large lecture class. Among the project goals will be to (1) analyze the reliability and validity of these materials, (2) evaluate their effectiveness in the process of instruction, and (3) acquire baseline data regarding student performance that will be of value to other instructors who make use of the materials.

It is important to make clear here at the outset that these assessment materials are, in a sense, "100% formative" in that they have the simultaneous function of (a) assessing student learning, (b) guiding the instructor to make necessary alterations and corrections to his or her instruction, and (c) forming an integral part of the instructional activities themselves. Although they may therefore be regarded as curricular materials, they are designed to allow real-time formative assessment by the instructor who may then make direct application of the results with literally no time delay at all. It should also be noted that the materials are designed for use in the context of a class organized along "active-learning" lines, in a format to be described in detail below.

The assessment materials themselves consist of carefully sequenced sets of multiple-choice questions, each focused on a specific topic. The individual items are primarily conceptual questions which downplay algebraic manipulations, and instead make heavy use of diagrammatic, graphical, and pictorial elements. The materials are intended for use in large lecture classes, and they are specifically designed to allow for rapid and reliable assessment of student learning during the course of a single class. The instantaneous feedback they provide will allow instructors to make immediate alterations, as needed, in their presentations and planned instructional activities.

The assessment materials are designed to be integrated into a workbook which also includes non-multiple-choice (free-response) worksheets designed for students working in groups, and which contains as well a supporting set of "Lecture Notes." In fact, the major portion of such a workbook designed for the second semester of the algebra-based general physics course (i.e., Vol. II of the Workbook) has already been written, and much of the material has been repeatedly class tested. The present proposal is intended to support the development and testing solely of the multiple-choice assessment items. More specifically, about two-thirds of the items for Vol. II of the Workbook have already been developed and require additional testing and validation; one-third of the items for Vol. II have yet to be developed. To date, none of the items for Vol. I have been developed.

In the subsections to follow, I will describe these assessment materials in detail and discuss preliminary work and results. First, though, I will outline the motivation that underlies their creation, beginning with a brief overview of general pedagogical issues and then discussing concrete applications in a real instructional setting.

General pedagogical issues: The motivation for the creation of these new materials is the now very extensive research base into student learning in introductory physics at the university level. An increasing body of evidence suggests that instruction utilizing ***only*** lecture classes and standard recitations and labs results in relatively small increases in ***most*** students' understanding of fundamental concepts (e.g., Hestenes, Wells, and Swackhamer, 1992; Hake, 1998; Redish, Saul and Steinberg, 1997). It has been pointed out by many experienced researchers that complex scientific concepts are often not effectively communicated to students simply by lecturing about them – however clearly and logically the concepts may be presented (e.g., McDermott, 1991, 1993, 1997; Redish, 1994; Wells, Hestenes, and Swackhamer, 1995; Arons, 1997). For one thing, students taught exclusively through lecture-based curricula are inclined to short-circuit the highly complex scientific thought process (Maloney, 1994; Reif, 1995). In other words, students do not absorb physics concepts simply by being told (***or*** shown) that they are true. They must be guided continually to challenge their present state of understanding, and to resolve conceptual confusion through some process of "active engagement."

Pedagogical models that actively engage students in a process of investigation and discovery – often oriented around activities in the instructional laboratory – have been found to be effective in improving students' conceptual understanding of physical principles (Thornton and Sokoloff, 1990, 1998; Hake, 1992, 1998; Redish and Steinberg, 1999). The targeted physical concepts are in general not "told" to the students before they have the opportunity to carry out investigations – or follow through chains of reasoning – that might lead them to synthesize the concept on their own. It has been especially challenging to develop effective active-learning materials that do not have the benefit of a simultaneous laboratory component to the instruction. Similarly, the environment of the large lecture class – where there may be 100, 200, or more students facing a single instructor – is an extremely challenging environment in which to establish active learning. This is, in effect, the starting point for this project.

Other approaches to active-learning in large classes. A number of workers in recent years have explicitly addressed the challenge of the large-class learning environment. A pioneering figure has been Alan Van Heuvelen (Van Heuvelen, 1990a,b; 1991a,b), who early on developed "Active-Learning Problem Sheets," consisting of free-response worksheets for use by students during class meetings in the lecture hall. Eric Mazur (Mazur, 1997; Crouch and Mazur, 2001) has achieved spectacular success in popularizing "Peer Instruction," the method he developed for suspending a lecture at regular intervals with challenging conceptual questions posed to the whole class. The students spend several minutes thinking about and discussing the question with each other, then offer a response to the instructor using one form or another of classroom communication system. Sokoloff and Thornton, 1997, have adapted their popular and effective Microcomputer-based Laboratory materials (Thornton and Sokoloff, 1990, 1998, originated in collaboration with Priscilla Laws, 1993) for use in large lecture classes, in the form of "Interactive Lecture Demonstrations." Novak and collaborators (Novak *et al.*, 1999) have developed "Just-In Time Teaching," which makes use of pre-class web-based computer warm-up exercises, and in-class group work by students using white boards. To some extent these incorporate similar methods used and promoted by Hestenes and his collaborators (Hestenes, 1997) who have developed "Modeling Instruction." The Physics Education Group at the University of Washington has experimented with modifications of their "Tutorials in Introductory Physics" (McDermott *et al.*, 1998), adapted for use in large lecture classes (Kraus, 1997). Textbooks and workbooks with a high "interactive" component that have been used in large classes include those by Chabay and Sherwood, 1995, and Knight, 1997. Other implementations of active learning in large classes using classroom communication systems have been described by Dufresne *et al.*, 1996, Shapiro, 1997, and Bernstein and Lederman, 2001. The "Scale-Up" project at North Carolina State University (Beichner and Saul, 2000) also makes use of technology-based systems with similar goals in mind.

Distinctive elements of the current project. A detailed description of the materials proposed for this project will be given below. Here however I pause simply to highlight a number of the key distinctions between this work and those cited immediately above. We have, of course, drawn enormously on the outstandingly creative and original efforts of these other workers. The present project incorporates many of those ideas, with the following unique set of elements: (1) the assessment items in this project are specifically targeted at the conceptual entry level. That is, they are primarily designed with the algebra-based general physics course in mind, and moreover include a large proportion of items suitable for students in those courses with below-average preparation levels. (2) The items are designed with numerous multiple-choice responses suitable for use with the flash-card response system (see below), or other similar systems. (3) The items are organized in discrete sets of tightly-structured sequences, each focused on a specific topic. They are intended to build on and feed off of each other by proceeding in an easy-to-hard conceptual sequence that incorporates different contexts and multiple forms of representation. (4) The conceptual step size between the items is relatively small, thereby increasing their utility for minute-by-minute assessment of learning and allowing multiple fine-turned course adjustments by the instructor during a single class. This also permits the use of a relatively large number of items during each individual class, in comparison with most of the other methods described above.

(It is appropriate to mention at this point that at Iowa State University, and probably at many other institutions, a majority of students in the second semester of the algebra-based general physics course are female, in striking distinction to the demographic composition of almost all other physics courses at the university. Most of these students are life-sciences majors, predominantly majors in biology, microbiology, and animal science. Many are pre-professional students who are planning to attend medical or veterinary school, or other health-science professions such as physical therapy, pharmacy, optometry, chiropractic, etc. For this reason the present project will have a disproportionately large immediate impact on females, a demographic group that is typically underrepresented in physics courses.)

The Challenge: Learning and assessment in a large lecture class. The motivation for this project begins in a classroom where the instructor comes face to face with 50, 100, or more students at once. The instructor begins a carefully prepared presentation, striving to be as clear as possible. Every now and then the instructor asks a question of the class, pauses and waits for someone to answer, and then comments on their response. Repeatedly, the instructor asks if anyone has any questions about what they have just said. It seems that no one does – or maybe just one or two people, and always the same ones. The instructor is a bit uneasy about the lack of questions – surely they're not finding the explanations to be all ***that*** clear, are they?

So, how well ***did*** the students follow the lecture? How can the instructor really know? Instructors can wait until the exam and see how well the students do, but this does not really tell you whether the students actually got anything out of the lecture. Perhaps they learned nothing in class, but figured it out themselves by reading the text. (In fact, if the class is typical, probably 50% or more of the students are not even there on an average day.) The premise of this project is that there is, indeed, a way out of this frustrating dilemma. There is an effective way to do formative assessment of student learning in large lecture classes "real-time," and to implement necessary alterations and corrections right on the spot.

In what follows, I will briefly describe instructional methods designed to address these issues that were developed in close collaboration with Kandiah Manivannan of Southwest Missouri State University (Meltzer and Manivannan, 1996, 2002). I have used these methods primarily in the second semester of the introductory general physics course, taught over the past seven years at Southeastern Louisiana University and Iowa State University. Both institutions are typical in that their large student enrollments result in many large lecture courses. In physics, this means that an instructor teaching an introductory course might face anywhere from 100 to 250 students at one time. Students are often unhappy with the "anonymous" atmosphere of such large classes, where their individual questions may go unasked and answered. Instructors too are frequently dissatisfied with the very limited amount of individual attention they are able to provide in such a situation. Both have a common interest in improving the effectiveness of the learning environment in these large lecture courses, and that is the goal of these methods.

These methods are, in effect, a variant of "Peer Instruction," which was developed by Eric Mazur at Harvard University (Mazur, 1997). The basic strategy is to drastically increase the quantity and quality of ***interaction*** that occurs in class between the instructor and the students, and among the students themselves. To this end, the instructor poses many questions. All of the students must decide on an answer to the question, discuss their ideas with each other, and provide their responses to the instructor using a classroom communication system. The instructor makes immediate use of these responses by tailoring the succeeding questions and discussion to most effectively match the students' pace of understanding. As a model of this learning environment, consider the instructor's office.

When an instructor has one or two students in her office asking for help, it would be surprising indeed if she proceeded to present a 50-minute lecture, pausing every now and then to ask a question. More likely she will speak for just a few minutes, sketching diagrams and writing a few simple equations. Then she will stop and solicit some feedback. Maybe she poses a simple question; perhaps she sketches out a problem for the students to try, or asks them to draw a diagram of some sort. Perhaps the instructor asks one of the students to comment on an answer given by the other.

In the office, the instructor is able to get an ongoing sense of where the students are, "conceptually," and how well they are following the ideas that are being presented. By getting continual feedback from them, the instructor is able to tailor his or her presentation to the students' actual pace of understanding. By asking the students to consider each other's ideas (if there is more than one of them in the room), the instructor helps them to think critically about their own ideas. The question, then, is whether it's ***practical*** to do this in a room filled with over 100 students.

A simple answer is now possible: It ***is*** practical. It is not necessarily easy, but it is very possible to recreate in the lecture hall much of the learning environment that exists in the instructor's office. One can to a large extent transform the environment of the lecture hall into that of a small seminar room in which ***all*** the students are actively engaged in the discussion.

The key is for the instructor to be able to get ***instantaneous*** feedback ***simultaneously from all the students in the class***. I ask questions during class – ***many*** questions – and no longer have to wait for one daring individual to offer a response. Every single student in the class has a pack of six large "flash cards" (5½" × 8½"), each printed with one of the letters A, B, C, D, E, and F. They bring the cards every day, and I always have extras in case someone forgets. Repeatedly during class I will present a multiple-choice question to the students. The questions stress qualitative concepts involving comparison of magnitudes (e.g., "*Which is larger: A, B, or C?"*), direction (*"Which way will it move?"*) and trends (*"Will it decrease, remain the same, or increase?"*). These kinds of questions are hard to answer by plugging numbers into an equation. I give the students some time to consider their response, 15 seconds to a minute depending on the difficulty. Then I ask them to signal their response by holding up one of the cards, everybody at once. I can easily see all the cards from the front of the room. Immediately, I can tell whether most of the students have the answer I was seeking – or if, instead, there is a "split vote," half with one answer, half with another. (Hopefully, one is the right answer!)

I can see whether the class held up their cards quickly, with confidence, or if instead they brought them up slowly, with confused looks on their faces. If there is a split vote, I ask them to talk to each other. I allow about a minute for those who think the answer is, say, "A" to try to persuade those who believe it is "C" to change their views. And, of course, the "C" supporters argue for ***their*** side of the case. Then I ask for another vote. If it is still split, I'll ask for an "A" supporter to stand and present their argument, followed (in alphabetical order) by a proponent of the "C" point of view. Eventually, if necessary, I will step in to – I hope – alleviate the confusion. But by this time, most of the students will have thought through the concept that was causing the problem; they will have thought it through ***hard*** because they will have tried to convince their neighbors that they were right. And, if they haven't already figured things out by themselves, they will now at least be in an excellent position to make ***sense*** out of any argument I offer to them. Before that minute or two of hard thinking, though, I could have made the same argument and then watched as almost every student in the class gave the ***wrong*** answer to some simple question. I know this is true, because I have tried it often enough.

In the course of using these assessment materials I have had many opportunities to ask my students questions during lecture that I would once have considered "trivial." These questions pertain to concepts that I – and most instructors – would have covered in a few seconds or a minute of clear, logical reasoning. ***Impossible*** to think that my students could get these simple questions wrong, or have any difficulty with them. But in fact they do, and now I know it. I pose a question that, I think, is a completely straightforward application of a principle I just presented. For instance: *If a parallel circuit with two identical resistors has a third identical resistor added in parallel, what happens to the total power provided by the battery?* The logic points inescapably toward only one possibility. I wait impatiently as my students study the question, debating the answer with each other, looking around. Slowly, after a minute, the cards come up: half are "A" (decreases), and nearly a third are "B" (remains the same). But the correct answer is "C" (increases), a choice selected by perhaps one student out of five.

I realize that I need to retreat, and I offer another question – perhaps I make it up on the spot – that goes back to a concept discussed last week. Then we work our way through a series of intermediate questions, back to the one that started the trouble. At each step, I get a reading on my class: Do they respond quickly? With confidence? *Mostly* correctly? Then I comment briefly and move forward. Otherwise, I pause for a longer discussion. In the old days I would have disposed of this entire topic in less than two minutes of lecture, and have been well satisfied that I made my points clearly and effectively. Now I take 10 to 15 minutes, and struggle together with my students as they work their way through a conceptual minefield. But this time, I believe, my students really do construct a basis for understanding the material.

A typical class proceeds in three phases:

(1) A brief introduction/review of the basic concepts is presented at the blackboard, a sort of "mini-lecture" lasting three to seven minutes.

(2) A sequence of about a half-dozen multiple-choice questions (sometimes more) is posed to the class; these questions emphasize qualitative understanding, proceed from relatively simple to more challenging, and are closely linked to each other to explore just one or two concepts from a multitude of perspectives. They frequently employ graphs, diagrams, and "verbal" descriptions. Students provide responses to these questions using the flash cards as described above.

(3) Follow-up activities, which – depending on the individual course – may consist either of interactive demonstrations, group work using free-response worksheets, or another sequence of mini-lecture followed by multiple-choice problem set.

Assessment materials. This method is crucially dependent on having at one's disposal a large number of carefully constructed sequences of conceptual multiple-choice questions. The purpose of emphasizing non-numerical questions is to prevent students short-circuiting the thinking process by blindly plugging numbers into poorly understood equations. Although some collections of such problems exist in the literature (e.g., Mazur, 1997; Novak *et al.*, 1999) and at the constantly expanding Project Galileo web site (http://galileo.harvard.edu), we have had to construct our own set to meet part of the needs of a full one-semester course (Meltzer and Manivannan, 2000). It is the preparation and testing of such question sets that is among the most time-consuming prerequisites for this instruction. The question sets that we have created up until now are based, as much as possible, on the physics education research literature. The purpose of this present project is to complete the development of these items for the second semester of the algebra-based physics course, and to begin development of those for the first semester. The topical areas intended for development are magnetism, light and electromagnetic waves, optics, modern physics, kinematics, dynamics, work, and energy. It is important to point out that although the materials are targeted primarily at the algebra-based course, they may be used quite effectively by instructors in the calculus-based course as well. This type of qualitative, conceptual assessment question is a valuable tool in any introductory physics course, and their utility is not at all restricted by the identity of their original target audience. Indeed, these materials can be used effectively by instructors are all types of institutions, including both two-year and four-year colleges, universities, and high schools.

The materials are designed with a guiding theme. This theme is the premise that the solution of even very simple physics problems invariably hinges on a lengthy chain of concepts and reasoning, much of which is often glossed over, or which is simply unstated “tacit” knowledge gained through experience (Reif, 1995). The question sequences guide the student to lay bare these chains of reasoning. They help students construct in-depth understanding of physical concepts through step-by-step confrontation with conceptual “sticking points” and counterintuitive ideas. One has to illuminate in a stark and glaring light, so to speak, the phases in the student’s thought process where the concept is lacking, so that in the student’s own mind the gap to be filled by the missing concept is clearly sensed. Then, the eventual synthesis of the concept by the student becomes ***dramatically*** apparent to them.

This is accomplished through carefully linked sequences of activities that first lead the student to confront the conceptual difficulties, and then to resolve them. This is, in essence, the strategy developed and employed by the Physics Education Group at the University of Washington (McDermott, 1991, 1993, 1997). The strategy is to break down complex physical problems into conceptual elements, allowing students to grapple with each one in turn, and then returning to synthesize a unifying perspective. Frequently, we return to re-explore difficult concepts in varied contexts, in order to reinforce students’ understanding and confidence.

It is important to emphasize that materials of the type proposed here are *not* presently available in significant quantities elsewhere, whether in the test banks that accompany standard textbooks, or even in the very fine research-based curricular materials produced by other groups. Although individual questions on various topics may be found in a variety of sources, the multi-item sets of carefully sequenced, tightly focused multiple-choice questions employing small conceptual step sizes (necessary for use in “fully interactive” lectures) are not available outside of the *Workbook for Introductory Physics* (Meltzer and Manivannan, 2000).

On the following pages I present a sample excerpt from a set of assessment items. This includes parts of the sequence on electrical forces. A number of items are omitted due to space limitations. The entire sequence consists of 18 items, and each item in the sample shown carries its number from the original order.

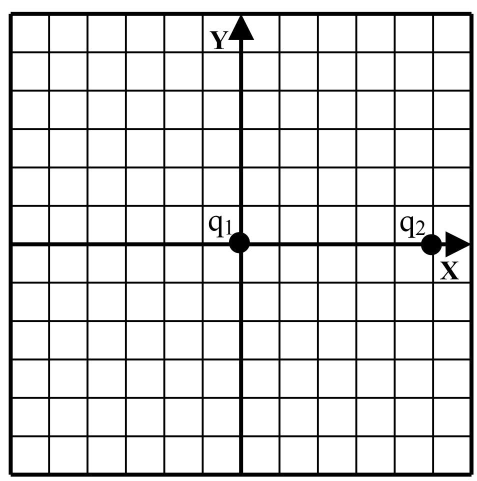
Chapter 1 Electrical Forces

In-Class Questions

Prerequisite Concepts:

- **Positive and negative charges; Coulomb's law: $F = kq_1q_2/r^2$**
- **Protons (+) and electrons (-)**
- **Superposition principle:** $\vec{F}_{net} = \vec{F}_1 + \vec{F}_2 + \ldots + \vec{F}_n$
- **Vector addition: $F_{netx} = F_{1x} + F_{2x} + \ldots + F_{nx}$**
- **Newton's second law,** $\vec{a} = \frac{\vec{F}_{net}}{m}$

Questions #1-2 refer to the figure below. Charge q_1 is located at the origin, and charge q_2 is located on the positive x axis, five meters from the origin. There are no other charges anywhere nearby.



1. If q_1 is positive and q_2 is negative, what is the direction of the electrical force on q_1?

   A. in the positive x direction
   B. in the negative x direction
   C. in the positive y direction
   D. in the negative y direction
   E. the force is not directed precisely along any of the coordinate axes, but at some angle
   F. there is no force in this case

2. If q_1 is positive and q_2 is positive, what is the direction of the electrical force on q_1?
    A. in the positive x direction
    B. in the negative x direction
    C. in the positive y direction
    D. in the negative y direction
    E. the force is not directed precisely along any of the coordinate axes, but at some angle
    F. there is no force in this case

3. In this figure, a proton is located at the origin, and an electron is located at the point (3m, 3m). What is the direction of the electrical force on the proton?

A.

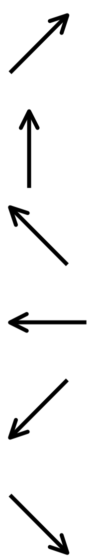
B.

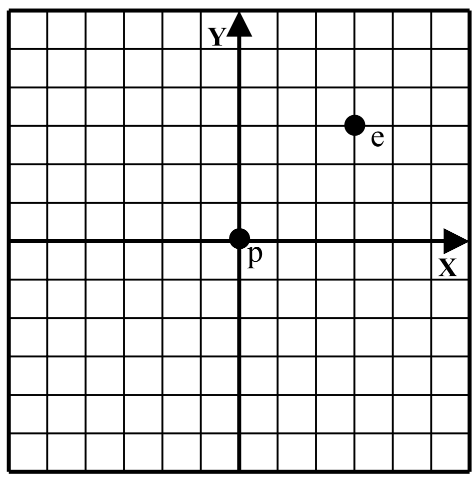
C.

D.

E.

F.





9. Which of these will result in the repulsive force between two identical charged particles *increasing* by a factor of eight:

    A. double one of the charges

    B. double both of the charges

    C. double one of the charges and cut the particle separation in half

    D. triple one of the charges and cut the particle separation in half

    E. triple both of the charges

    F. double both of the charges and double the particle separation

10. A 6-C charge and a 12-C are separated by 2 m; there are no other charges present. Compared to the electrical force on the 6-C charge, the electrical force on the 12-C charge is:

    A. one-fourth as strong

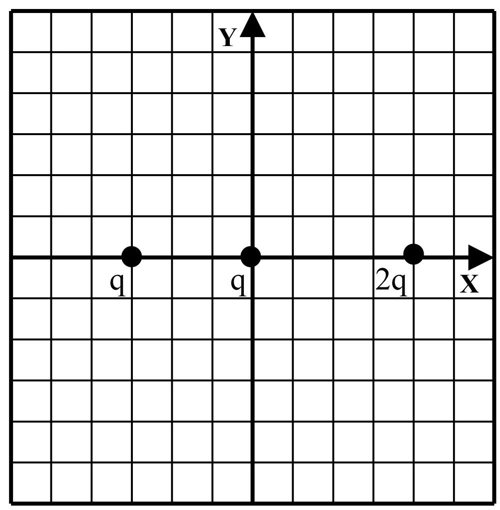
    B. one-half as strong

    C. the same magnitude

    D. two times as strong

    E. four times as strong

13. In this figure, positive charges of magnitude q, q, and 2q are located on the x axis as shown. The direction of the net electrical force on the positive charge at the origin is:



    A. towards positive x

    B. towards positive y

    C. towards negative x

    D. towards negative y

    E. nowhere, since there is no net force on this charge.

Previous work already completed: As has been described above, the in-class assessment materials (i.e., flash-card questions) form part of the "Workbook for Introductory Physics," which also includes free-response worksheets and lecture notes. The following table shows the portion of the work that is complete (√) (currently consisting of over 400 printed pages, including 114 multiple-choice "in-class" items, plus over 200 such items on quizzes and exams), and that which remains to be done (X). (An asterisk [*] indicates a column [shaded] corresponding to the sets of multiple-choice "in-class" assessment items covered under this proposal).

	*	*	*	*			
Chapter	**Written**	**Class tested**	**Validated with interview data**	**Baseline data acquired**	**Worksheets completed**	**Lecture notes completed**	**Quiz & Exam items completed**
1. Electrical Forces	√	√	X	X	√	√	√
2. Electrical Fields	√	√	X	X	√	√	√
3. Electric Potential Energy	√	√	X	X	√	√	√
4. Electric Potential	√	√	X	X	√	√	√
5. Current and Resistance	√	√	X	X	√	√	√
6. Series Circuits	√	√	X	X	√	√	√
7. Electrical Power	√	√	X	X	√	√	√
8. Parallel Circuits	√	√	X	X	√	√	√
9. Magnetic Forces and Magnetic Fields	√	√	X	X	√	√	√
10. Magnetic Induction	X	X	X	X	√	√	√
11. Electromagnetic Waves	X	X	X	X	√	√	√
12. Optics	X	X	X	X	√	√	√
13. Photons and Atomic Spectra	X	X	X	X	√	√	√
14. Nuclear Structure and Radioactivity	X	X	X	X	√	√	√
Vol. I, 1: Kinematics	X	X	X	X	X	X	X
Vol. I, 2: Forces	X	X	X	X	X	X	X
Vol. I, 3: Force and Motion	X	X	X	X	X	X	X
Vol. I, 4: Work and Energy	X	X	X	X	X	X	X

Results of class testing to date. The Workbook in its current form has been used for the past four years at Iowa State University. Although it has been under continuous development, the basic outline of materials has been in place for that whole time. Of course, since the multiple-choice items are merely one part of the entire curriculum project it is not possible to determine what part of the overall learning gains may be ascribable to them (or to any other separate part of the Workbook). Nonetheless it is relevant to cite some results. Among the most dramatic consequences of this instructional method is a very small number of dropouts, typically 1-3% after the first week. Class attendance averages about 90%, no doubt largely due to daily graded quizzes. End-of-course surveys show that most students react favorably to the instructional methods, with approximately 30-40% giving maximum ratings. (Sample comment: *". . . best physics instructor I have ever had. I liked the way he had class interaction and explained things. He makes physics fun and interesting to learn, whereas most physics instructors just babble inanely during lecture."*) Most of the remainder are positive or neutral, but a core of less than 10% ***despises*** these methods, and the instructor as well (sample comment from the same class: *". . . has a new way of teaching he is trying to develop. It doesn't work. He relies too heavily on the students to help each other, when all we want is to learn the material . . . going to lecture was pointless other than to take required quizzes."*)

What is clear, however, is that the overall learning gains by the students are very high in relation to comparable courses nationwide. For the past six years I have given an abridged version of the "Conceptual Survey of Electricity" (O'Kuma *et al.*, 1999), a diagnostic instrument that assesses qualitative understanding. A subset of 14 of these questions is also contained on the "Conceptual Survey of Electricity and Magnetism" (CSEM); national baseline data have recently been published for the CSEM (Maloney *et al.*, 2001). My students' pretest scores on the subset (three-year average of 28%) are nearly identical to those reported in comparable algebra-based courses, and substantially lower than those in a nationwide sample of over 1500 students in calculus-based courses (37%). However, the average post-test score of my students over the 1998-2000 period is 78%, while those of the nationwide sample range from 43% in the algebra-based course to 51% for students in the calculus-based classes (Maloney *et al.*, 2001). Other assessment data are consistent with these results. Moreover, on quantitative problems borrowed from exams given in the calculus-based class at Iowa State, students in my algebra-based course do comparably well, or better. (Detailed assessment data are in Meltzer and Manivannan, 2002.)

Plan of work: There are three main phases yet to be completed: (1) Drafting and initial class testing of multiple-choice items for Vol. II, Chapters 10-14, and Vol. I, Chapters 1-4; (2) Validation of all multiple-choice sets through "think-aloud" problem solving interviews with student volunteers; (3) Acquiring a complete set of baseline data for all multiple-choice questions, by mean of automatic data logging using an electronic classroom response system. In addition, ongoing assessment using the "Conceptual Survey of Electricity and Magnetism" (O'Kuma *et al.*, 1999; Maloney *et al.*, 2001), as well as other assessment items for which we have already acquired baseline data, will aid in evaluation of the impact the new materials will have on student learning. For Vol. I materials, the Force Concept Inventory (Hestenes *et al.*, 1992) will be used in assessment. Because of the structure of the course, it will not be possible to determine with any certainty what weight should be ascribed to the new materials with respect to increased learning gains. However, because essentially all other instructional materials are presently in place in near-final form, it is plausible that any additional learning gains of any significance in the *second*-semester course may be due in sizeable part to the new materials to be developed with this project.

As was the case with the materials already completed, the sets of multiple-choice assessment questions will focus on a limited set of specific concepts within each topic area. These are the concepts that are targeted in the Lecture Notes (already completed for each chapter in Vol. II). For Vol. II, Chapters 10-14, these key concepts are the following: *positive and negative magnetic flux; distinction between flux and changing flux; induced current; period, frequency, wavelength, and velocity of electromagnetic waves; image formation by thin lenses; absorption and emission of electromagnetic wave photons by transitions between atomic electron energy levels; radioactive decay and half-life; radioactive dating.* For Vol. I, Chapters 1-4, it would be premature to give a very specific list of topics. However, the broad areas include *velocity, acceleration, forces, vector addition and force superposition, Newton's laws of motion, work, work-energy theorem, kinetic and potential energy, conservation of energy*.

Validation plan and reliability check: Up until this time, the assessment items have been drafted based on extensive instructional experience, knowledge of the results of physics education research, and experience in the use of previous, related assessment items. The items for Chapters 1-9 have undergone extensive class testing (including trials at other institutions) and have been revised accordingly several times. However, what is still lacking is a systematic validation process that can determine to a still higher degree of certainty that the items both test the knowledge they are intended to test, and catalyze students' reasoning process in the manner intended. The most effective method for achieving this type of validation is through patient and time-consuming "think-aloud" interviews with individual students, recorded on audiotape.

In this type of process, students are asked to work through the sequence of questions, explaining their reasoning as they go, while the interviewer examines the details of the student's thinking with gently probing questions. This process can be very effective in (1) uncovering confusing or ambiguous language and word usage; (2) confirming that the students interpret the meaning of the question in the manner intended; (3) determining whether the students make any tacit assumptions intended by the question (e.g., no external electric field), and do not impose any unintended assumptions (e.g., a need to consider small-magnitude interactions which the instructor had assumed would be ignored). In our previous class-testing experience, all of these problems, as well as others, have occurred repeatedly. Experienced instructors learn that they are virtually unavoidable in the "first draft" process. This proposal therefore includes a plan to (a) carry out student interviewing to aid validation of all newly drafted assessment items, and (b) conduct interviews on a selected basis for items that have previously been written, class tested, and revised. The outcome of this process should be to substantially strengthen the quality and utility of the collection of assessment items as a whole.

As part of the interview process, it will also be possible to obtain a check on the reliability of the assessment items. This is not a proposal to create an actual assessment "instrument," and so standard reliability measures such as KR-20, split-half, etc. are not appropriate. Rather, it is important to gauge whether (and to what degree) immaterial alterations in wording or context will significantly affect students' answers to a question. Using a variant of the "alternate forms" technique, I will include in the student interviews two or more slightly varied forms of selected items in each topical sequence. This will provide a measure of how robust are the student's responses to the item, and therefore on how reliable the item is in reflecting the student's thinking. This will help pinpoint instances where revisions may be necessary.

Collection of baseline data. Although the flash-card system is efficient and inexpensive, it does have one major drawback in the context of materials development: it does not allow for rapid and accurate recording of student responses. This is not normally a problem when considering solely the instructional function of the system. However, a major goal of this project is to record student responses to each of the assessment items, including those items already developed and class-tested, as well as the items that will first be developed as a result of the present project. For this reason I propose to purchase an electronic classroom communication system. This consists of individual handheld wireless keypads for each student which allow them to signal a response to a multiple-choice question. The signals are received, logged and tabulated by a central computer. ISU already owns such a system with 100 keypads, but this is insufficient for our lecture classes in the mechanics portion of the algebra-based course. Therefore I propose to purchase 100 additional keypads and the associated hardware and software.

The database to be created will have several very important uses: (a) It will provide a baseline for comparison when other instructors make use of the assessment materials. The student population enrolled in physics courses at Iowa State University is one that is very much characteristic of a large segment of physics students nationwide. A bank of "typical" student response rates for each assessment item will provide a useful benchmark for other instructors regarding the performance of their own students. (b) The database will allow detailed analysis of assessment items to help pinpoint possible anomalies in the response patterns. These may indicate items that need revision or rewording, or which merit special attention by testing during student interviews. These data may also provide insights into student thinking that will allow for drafting of additional or revised questions that sharpen the focus of instruction and assessment.

Dissemination: In collaboration with K. Manivannan, I have already given four workshops at AAPT national meetings (e.g. Meltzer and Manivannan, 1998) in which other physics instructors were guided in the use of the instructional methods and materials described in this proposal. That initial cycle of workshops has now ended. As the complete version of the assessment items and other curricular materials is put into final form, it will be time to initiate a new cycle of workshops to begin the process of disseminating the materials and accompanying baseline data.

Several other instructors have already done initial testing of materials, and have expressed interest in continuing such in-class testing in the future. In addition, Prof. Kandiah Manivannan (Southwest Missouri State University) and Prof. Dan McCarthy (Southeastern Louisiana University) have expressed an ongoing commitment to do extensive in-class testing of the new materials in the courses they teach at their institutions. ***Please see section I for commitment letters from Drs. Manivannan and McCarthy.***

An early version of the Workbook was sent out for review by a major textbook publisher and received positive, and in some cases enthusiastic reviews. One issue, however, is that publishers are strongly inclined toward publishing a complete two-volume set to cover the entire general physics course. Since it will take some time for Volume I of this Workbook to be completed, this could lead to delays. We have therefore adopted an interim dissemination method that is proving to be quite effective. The entirety of the preliminary version of the Workbook has been burned onto a CD-ROM, and copies of this CD-ROM version of the materials are being distributed to physics instructors nationwide. Initially, both instructors who have specifically indicated an interest in the materials, as well as others who are known to use similar instructional methods, have been targeted to receive the preliminary version. The CD-ROM includes PDF files of every item, to insure faithful reproduction, but also includes Microsoft Word versions of most items as well. This ensures that individual instructors will be able to modify and adapt the materials to their own local circumstances. Because of the extremely low unit dissemination cost of the CD-ROM version, we have simply been mailing out free copies of the disc in order to promote widespread use and class testing; we are presently absorbing the reproduction and shipping costs ourselves at ISU. More than 50 CD's have been distributed so far. We have placed an announcement on our web site inviting instructors to request a free copy of the CD-ROM. This method should be an extremely viable alternative for dissemination of the new materials to be created by this project, pending commercial publication. We are also exploring the possibility of posting selected portions of the materials on our web site, accessible via password to physics instructors. In addition, we may be able to post materials on the FLAG web site (i.e., the ("Field-tested Learning Assessment Guide") at http://www.wcer.wisc.edu/nise/CL1/flag/default.asp.

Broader Impacts: As has been detailed above, this project has among its central goals (1) develop research-based educational materials and creation of a database (of student response frequencies) useful in teaching; (2) involve graduate researchers in undergraduate teaching activities; (3) participate in developing new approaches (e.g., use of interactive lecture instruction) to engage underserved individuals and groups (i.e., female physics students); (4) make data available in a timely manner by means such as CD-ROMs; (5) publish in diverse media (e.g., websites and CD-ROMs) to reach broad audiences; (6) integrate research (on teaching and learning) with education activities to order to communicate in a broader context, and (7) benefit society by increasing the effectiveness of undergraduate physics instruction.

The tentative timeline for this project is as follows:

July 2003-December 2003: Write first draft of assessment materials for remaining Chapters of Vol. II (10-14). Initial class test of newly drafted materials; first phase of baseline data acquisition for all assessment items using electronic classroom response system. Develop and administer exam questions related to new flash-card questions; analyze student responses to questions. Recruit student volunteers and begin interviews to assess validity of all items, beginning with Chapter 1.

January 2004-June 2004: Write first draft of assessment materials for initial chapters of Vol. I (1-4). Analyze assessment data for Vol. II materials, use to revise and correct these materials. Initial testing of Vol. II materials at off-campus sites including Southwest Missouri State University (SMSU).

July 2004 – December 2004: Revise and correct all materials, begin to put Vol. II materials into final form. Class test Vol. I materials at Iowa State University. Carry out interviews to assess Vol. I materials. Initial testing of Vol. I materials at off-campus sites including SMSU. Initiate workshops for other instructors in use of materials.

January 2005 – June 2005: Final phase of class testing at Iowa State. Completion of interviews and analysis of all assessment data. Class testing at SMSU and Southeastern Louisiana University. Carry out workshops for other faculty at AAPT national meeting.