Center for Physics Education
Department of Physics
University of Washington
Seattle, Washington

Research as a guide for curriculum development

Selected readings to accompany workshop by
Lillian C. McDermott and the Physics Education Group
at the University of Washington

Workshop for new physics and astronomy faculty
American Center for Physics
November 2011

Table of Contents

UNIVERSITY OF WASHINGTON

Physics Education Group
Department of Physics Box 351560
Seattle, WA 98195-1560

206.685.2046
206.543.9702 (Fax)

The Physics Education Group in the Physics Department at the University of Washington conducts a coordinated program of research, curriculum development, and instruction to improve student learning in physics (K-20). The work of the group is guided by ongoing discipline-based education research. For more than 30 years, we have been deeply involved in the preparation of prospective and practicing teachers to teach physics and physical science by inquiry. In undergraduate physics, we have been engaged in a major effort to improve the effectiveness of instruction at the introductory level and in more advanced courses. These projects provide a context in which we work toward promoting the professional development of teaching assistants and new faculty.

As the director of the Physics Education Group, Professor Lillian C. McDermott shares leadership responsibilities with Professors Paula R.L. Heron and Peter S. Shaffer. The group includes visiting faculty, research associates, graduate students, and a small administrative staff. Graduate students in the group earn a Ph.D. in physics for research on the learning and teaching of physics. Through in-depth investigations of student understanding, the group seeks to identify and analyze specific difficulties that students encounter in studying physics. The findings are used to guide the development of two sets of instructional materials. Ongoing assessment, which is an integral part of this iterative process, takes place at the University of Washington and at pilot sites.

Physics by Inquiry is a self-contained curriculum primarily designed for the preparation of elementary, middle, and high school teachers but also suitable for liberal arts students and for students who aspire to science-related careers but who are underprepared in science and mathematics. The curriculum consists of a set of laboratory-based modules, all of which require active participation by the learner. Experiments and observations provide the basis on which students construct physical concepts and develop analytical reasoning skills. The topics have been chosen to provide teachers with the background needed for teaching K-12 science competently and confidently. Depth is stressed rather than breadth of coverage. Volumes I and II were published by John Wiley & Sons, Inc., in 1996.

Tutorials in Introductory Physics is being developed to supplement the lectures and textbooks through which physics is traditionally taught. The tutorials are suitable for both calculus-based and algebra-based courses in which there is an opportunity for students to work together in small groups. Carefully sequenced exercises and questions engage students in the type of active intellectual involvement that is necessary for developing a functional understanding of physics. Prentice Hall published a Preliminary Edition in 1998, a First Edition in 2002, and an Instructor's Guide in 2003.

In addition to publication of the two curricula, results are disseminated through talks presented at national and international meetings and through papers published in refereed journals, magazines, and conference proceedings. The work of the group, which is supported in part by the National Science Foundation, has contributed significantly to the formal recognition of physics education research as an important field for scholarly inquiry in physics departments.

Physics Education Group Publications

Books

Tutorials in Introductory Physics, First Edition, Lillian C. McDermott, Peter S. Shaffer, and the Physics Education Group at the University of Washington, NJ: Prentice Hall (2002).

Tutorials in Introductory Physics, First Edition, Instructor's Guide, Lillian C. McDermott, Peter S. Shaffer, and the Physics Education Group at the University of Washington, NJ: Prentice Hall (2003).

Physics by Inquiry, L.C. McDermott and the Physics Education Group at the University of Washington, NY: John Wiley & Sons, Inc. (1996).

Video

Physics by Inquiry: A Video Resource, Boston: WGBH Educational Foundation (2000).

Articles in refereed journals and professional publications

A. Boudreaux, P.S. Shaffer, P.R.L. Heron, and L.C. McDermott, "Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system," *Am. J. Phys.* **76** (2) 163 (2008).

L.C. McDermott, Guest Editorial, "Preparing K-12 teachers in physics: Insights from history, experience, and research," *Am. J. Phys.* **74** (9) 758 (2006).

L.C. McDermott, P.R.L. Heron, P.S. Shaffer, and M.R. Stetzer, "Improving the preparation of K-12 teachers through physics education research," *Am. J. Phys.* **74** (9) 763 (2006).

M.J. Cochran and P.R.L. Heron, "Development and assessment of research-based tutorials on heat engines and the second law of thermodynamics," *Am. J. Phys.* **74** (8) 734 (2006).

D.L. Messina, L.S. DeWater, and M.R. Stetzer, "Helping preservice teachers implement and assess research-based instruction in K-12 classrooms," accepted for publication in *Proceedings of the 2004 Physics Education Research Conference,* Sacramento CA, August 2004, edited by J. Marx, S. Franklin, and P.R.L. Heron, AIP Conference Proceedings.

C.H. Kautz, P.R.L. Heron, M.E. Loverude, and L.C. McDermott, "Student understanding of the ideal gas law, Part I: A macroscopic perspective," *Am. J. Phys.* **73** (11) 1055 (2005).

C.H. Kautz, P.R.L. Heron, P.S. Shaffer, and L.C. McDermott, "Student understanding of the ideal gas law, Part II: A microscopic perspective," *Am. J. Phys.* **73** (11) 1064 (2005).

P.S. Shaffer and L.C. McDermott, "A research-based approach to improving student understanding of the vector nature of kinematical concepts," *Am. J. Phys.* **73** (10) 921-931 (2005).

L.C. Ortiz, P.R.L. Heron, and P.S. Shaffer, "Investigating student understanding of static equilibrium and accounting for balancing," *Am. J. Phys.* **73** (6) 545 (2005).

P.R.L. Heron and D.E. Meltzer, Guest Editorial, "Future of physics education research: Intellectual challenges and practical concerns," *Am. J. Phys.* **73** (5) 390 (2005).

L.C. McDermott, "Physics education research: The key to student learning and teacher preparation," *Physics World,* January 2004, pp. 40-41.

L.C. McDermott, "Improving student learning in science," *LTSN (Learning and Teacher Support Network) Physical Science News* **4** (2) pp. 6-10. United Kingdom: University of Liverpool (2003). See also, <http://www.physsci.ltsn.ac.uk/Publications/Newsletter/LTSNnl10.pdf>.

M.E. Loverude, C.H. Kautz, and P.R.L. Heron, "Helping students develop an understanding of Archimedes' principle, Part I: Research on student understanding," *Am. J. Phys.* **71** (11) 1178 (2003).

P.R.L. Heron, M.E. Loverude, P.S. Shaffer, and L.C. McDermott, "Helping students develop an understanding of Archimedes' principle: Research on student understanding, Part II: Development of research-based instructional materials," *Am. J. Phys.* **71** (11) 1188 (2003).

R.E. Scherr, P.S. Shaffer, and S. Vokos, "The challenge of changing deeply held student beliefs about the relativity of simultaneity," *Am. J. Phys.* **70** (12) 1238 (2002).

M.E. Loverude, C.H. Kautz, and P.R.L. Heron, "Student understanding of the first law of thermodynamics: Relating work to the adiabatic compression of an ideal gas," *Am. J. Phys.* **70** (2) 137 (2002).

L.C. McDermott, Response for the 2001 Oersted Medal, "Physics education research: The key to student learning," *Am. J. Phys.* **69** (11) 1127 (2001).

R.E. Scherr, P.S. Shaffer, and S. Vokos, "Student understanding of time in special relativity: simultaneity and reference frames," *Phys. Educ. Res., Am. J. Phys.* **69** (7) S24 (2001).

L.C. McDermott and P.S. Shaffer, "Preparing teachers to teach physics and physical science by inquiry," in *The Role of Physics Departments in Preparing K-12 Teachers*, G. Buck, J. Hehn, D. Leslie-Pelecky, eds., College Park, MD: American Institute of Physics (2000) pp. 71-85.

L.C. McDermott, P.S. Shaffer, and C.P. Constantinou, "Preparing teachers to teach physics and physical science by inquiry," *Physics Education* **35** (6) 411 (November 2000).

S. Vokos, B.S. Ambrose, P.S. Shaffer, and L.C. McDermott, "Student understanding of the wave nature of matter: Diffraction and interference of particles," *Phys. Educ. Res.*, *Am. J. Phys* **68** (S1) S42 (2000).

L.C. McDermott, "A university-based physicist discusses concept formation in the laboratory," an invited article in *Inquiry and the National Science Education Standards--A Guide for Teaching and Learning,* Washington D.C.: National Academy Press (2000), pp. 93-94.

L.C. McDermott and L.S. DeWater, "The need for special science courses for teachers: Two perspectives," an invited chapter in *Inquiring into Inquiry Learning in Teaching and Science,* J. Minstrell and E.H. van Zee, eds., Washington, D.C.: AAAS (2000), pp. 241-257.

B.S. Ambrose, P.R.L. Heron, S. Vokos, and L.C. McDermott, "Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena," *Am. J. Phys.* **67** (10) 891 (1999).

L.C. McDermott and E.F. Redish, "Resource Letter: PER-1: Physics Education Research," *Am. J. Phys.* **67** (9) 755 (1999).

K. Wosilait, P.R.L. Heron, P.S. Shaffer, and L.C. McDermott, "Addressing student difficulties in applying a wave model to the interference and diffraction of light," *Phys. Educ. Res., Am. J. Phys. Suppl.* **67** (7) S5 (1999).

B.S. Ambrose, P.S. Shaffer, R.N. Steinberg, and L.C. McDermott, "An investigation of student understanding of single-slit diffraction and double-slit interference," *Am. J. Phys.* **67** (2) 146 (1999).

P.R.L. Heron and L.C. McDermott, "Bridging the gap between teaching and learning in geometrical optics: The role of research," *Opt. & Phot. News* **9** (9) 30 (1998).

K. Wosilait, P.R.L. Heron, P.S. Shaffer, and L.C. McDermott, "Development and assessment of a research-based tutorial on light and shadow," *Am. J. Phys.* **66** (10) 906 (1998).

T. O'Brien Pride, S. Vokos, and L.C. McDermott, "The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems," *Am. J. Phys.* **66** (2) 147 (1998).

R.N. Steinberg, G.E. Oberem and L.C. McDermott, "Development of a computer-based tutorial on the photoelectric effect," *Am. J. Phys*. **64** (11) 1370 (1996).

D. J. Grayson and L.C. McDermott, "Use of the computer for research on student thinking in physics," *Am. J. Phys.* **64** (5) 557 (1996).

L.C. McDermott, P.S. Shaffer and M.D. Somers, "Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood's machine," *Am. J. Phys.* **62** (1) 46 (1994).

L.C. McDermott, Guest Comment: "How we teach and how students learn—A mismatch?" *Am. J. Phys.* **61** (4) 295 (1993).

L.C. McDermott and P.S. Shaffer, "Research as a guide for curriculum development: An example from introductory electricity, Part I: Investigation of student understanding," *Am. J. Phys.* **60** (11) 994; Part II: Design of instructional strategies," *Am. J. Phys.* **60** (11) 1003 (1992).

L.C. McDermott, Millikan Award 1990, "What we teach and what is learned: Closing the gap," *Am. J. Phys.* **59** (4) 301 (1991).

L.C. McDermott, "A perspective on teacher preparation in physics and other sciences: The need for special courses for teachers," *Am. J. Phys.* **58** (8) 734 (1990).

L.C. McDermott, "Research and computer-based instruction: Opportunity for interaction," *Am. J. Phys.* **58** (5) 452 (1990).

R.A. Lawson and L.C. McDermott, "Student understanding of the work-energy and impulse-momentum theorems," *Am. J. Phys.* **55** (9) 811 (1987).

L.C. McDermott, M.L. Rosenquist, and E.H. van Zee, "Student difficulties in connecting graphs and physics: Examples from kinematics," *Am. J. Phys.* **55** (5) 503 (1987).

M.L. Rosenquist and L.C. McDermott, "A conceptual approach to teaching kinematics," *Am. J. Phys.* **55** (5) 407 (1987).

F.M. Goldberg and L.C. McDermott, "An investigation of student understanding of the real image formed by a converging lens or concave mirror," *Am. J. Phys.* **55** (2) 108 (1987).

F.M. Goldberg and L.C. McDermott, "Student difficulties in understanding image formation by a plane mirror," *The Phys. Teach.* **24** (8) 472–480 (1986).

L.C. McDermott, "Research on conceptual understanding in mechanics," *Physics Today* **37**, 24–32 (July 1984).

L.C. McDermott, M.L. Rosenquist, and E.H. van Zee, "Instructional strategies to improve the performance of minority students in the sciences," *New Directions for Teaching and Learning* **16**, 59–72 (1983).

D.E. Trowbridge and L.C. McDermott, "Investigation of student understanding of the concept of acceleration in one dimension," *Am. J. Phys.* **49** (3) 242 (1981).

D.E. Trowbridge and L.C. McDermott, "Investigation of student understanding of the concept of velocity in one dimension," *Am. J. Phys.* **48** (12) 1020 (1980).

L.C. McDermott, L. Piternick, and M.L. Rosenquist, "Helping Minority Students Succeed in Science: I. Development of a curriculum in physics and biology; II. Implementation of a curriculum in physics and biology; III. Requirements for the operation of an academic program in physics and biology," *J. Coll. Sci. Teach*. **9**, 135–140 (Jan. 1980), 201–205 (March 1980), 261–265 (May 1980).

L.C. McDermott, "Teacher education and the implementation of elementary science curricula," *Am. J. Phys.* **44** (5) 434 (1976).

L.C. McDermott, "Improving high school physics teacher preparation," *The Phys. Teach*. **13** (9) 523 (1975).

L.C. McDermott, "Practice-teaching program in physics for future elementary school teachers," *Am. J. Phys.* **42** (9) 737 (1974).

L.C. McDermott, "Combined physics course for future elementary and secondary school teachers," *Am. J. Phys.* **42** (8) 668 (1974).

Oersted Medal Lecture 2001: "Physics Education Research—The Key to Student Learning"

Lillian Christie McDermott
Department of Physics, University of Washington, Seattle, Washington 98195-1560

Research on the learning and teaching of physics is essential for cumulative improvement in physics instruction. Pursuing this goal through systematic research is efficient and greatly increases the likelihood that innovations will be effective beyond a particular instructor or institutional setting. The perspective taken is that teaching is a science as well as an art. Research conducted by physicists who are actively engaged in teaching can be the key to setting high (yet realistic) standards, to helping students meet expectations, and to assessing the extent to which real learning takes place.
[DOI: 10.1119/1.1389280]

PREFACE

I would like to thank the AAPT for the 2001 Oersted Medal. The accomplishments recognized by this honor are the result of many contributions over many years by faculty, post-docs, graduate students, K–12 teachers, and undergraduates in the Physics Education Group at the University of Washington. We have had many visitors, long and short term, who have enriched our work. In addition to those who have been directly associated with us, there are many others who have helped to build the field of physics education research. They have done so through direct participation in research, through their use of the results, and/or through their support. I want to emphasize that I view this award as one to our entire community and also as recognition of research on the learning and teaching of physics as a useful field for scholarly inquiry by physicists. I deeply appreciate being selected for the Oersted Medal but I am also overwhelmed by the list of previous recipients. Like many of them, I would like to use this opportunity to share some insights drawn from my experience.

I believe that our group's most significant achievement in the last two decades has been to demonstrate the value of discipline-based education research. Our investigation of student understanding of one-dimensional kinematics that began in 1973 led to the publication of research papers on velocity (December 1980) and acceleration (January 1981). These were the first of their kind to appear in the *American Journal of Physics*. The situation has changed greatly since then. Today, there are several groups that conduct research in physics education and there is a substantial literature. Rather than attempt to give a representative overview, I will focus on the work of the Physics Education Group because that is what I know best. Although the data, interpretations, and conclusions presented are drawn from the experience of our group, I shall try to identify the features of physics education research that I believe are the most critical and most universally applicable.

I. INTRODUCTION

Physics education research differs from traditional education research in that the emphasis is not on educational theory or methodology in the general sense, but rather on student understanding of science content. For both intellectual and practical reasons, discipline-based education research should be conducted by science faculty within science departments. There is evidence that this is an effective approach for improving student learning (K–20) in physics. The emphasis in the discussion here is on introductory students and K–12 teachers and, to a lesser extent, on graduate students in their role as teaching assistants. However, insights obtained through research have also proved to be a useful guide for instruction in more advanced physics courses.

II. PERSPECTIVES ON TEACHING AS AN ART AND AS A SCIENCE

Many physics faculty think of teaching solely as an art. This traditional view was clearly expressed in 1933 in the first article in the first journal published by the American Association of Physics Teachers.[1] In *Physics is Physics*, F. K. Richtmyer, who considered teaching very important, argued that it is an art and not a science. He quoted R. A. Millikan in characterizing science as comprising "a body of factual knowledge accepted as correct by all workers in the field." Professor Richtmyer went on to say:

> "Without a reasonable foundation of accepted fact, no subject can lay claim to the appellation 'science.' If this definition of a science be accepted—and it seems to me very sound—then I believe that one must admit that in no sense can teaching be considered a science."

Although this definition of science is somewhat limited, we may challenge the implication that it is not possible to build "a reasonable foundation of accepted fact" for the teaching of physics (and, by extension, other sciences). The Physics Education Group treats research on the learning and teaching of physics as an empirical applied science. We adhere, to the extent possible, to the rules of evidence of experimental physics. To this end, we document our procedures and results so that they can be replicated. Beyond its intrinsic interest to us, we believe that physics education research can provide the key to student learning. We conduct systematic investigations on how well students who have studied physics from the introductory to the graduate level understand important concepts and principles. We use the results to guide the development of instructional materials and assess their effectiveness on the basis of what students have learned. The graduate students in our group earn their

Ph.D.'s in physics for this type of research. As is the practice among scientists, we report our results at professional meetings and in peer-reviewed journals.

Results from our research support the premise that teaching can be considered a science. Students in equivalent physics courses with different instructors are remarkably similar in the way they respond to certain kinds of questions, both before and after standard instruction by lecture, textbook, and laboratory. We have found that there are a limited number of conceptual and reasoning difficulties that students encounter in the study of a given topic. These can be identified, analyzed, and effectively addressed through an iterative process of research, curriculum development, and instruction. Although students vary in the way they learn best, learning is not as idiosyncratic as is often assumed.

Student difficulties and effective strategies for addressing them are often generalizable beyond a particular course, instructor, or institution. When the results are reproducible, as is often the case, they constitute a "reasonable foundation of accepted fact." There is by now a rapidly growing research base that is a rich resource for cumulative improvement in physics instruction.[2] Publicly shared knowledge that provides a basis for the acquisition of new knowledge is characteristic of science. To the extent that faculty are willing to draw upon and to contribute to this foundation, teaching can be treated as a science.

A. Criteria for the effectiveness of instruction

The criteria an individual uses to assess the effectiveness of instruction reflect his or her perspective on teaching. When teaching is considered as an art, the criteria tend to be highly subjective with the personal qualities and style of an instructor having a strong influence on assessments. Instructors frequently judge the success of a new course or innovation by their impression of how much the students have learned or how satisfied they appear to be. An inspiring lecturer can motivate students and kindle their interest. The benefits, however, seldom extend beyond the instructor's own class. Student ratings of a course or instructor are a commonly accepted form of evaluation that is consistent with the view that teaching is an art. In some instances, however, we have found that students whose instructors received low ratings have done better on matched questions than those whose instructors received higher ratings. Moreover, when asked to rate how much they have learned, students are often poor judges. If student learning (as distinct from enthusiasm) is used as the criterion, we have found that effective teaching is not as tightly linked as is often assumed to the motivational effect of the lecturer, to student evaluations of the course and instructor, or to self-assessment of learning by students. Implicit in the perspective of our group that teaching is a science is the belief that the primary criterion for the effectiveness of instruction must be the assessment of student learning in terms of specified intellectual outcomes.

B. Focus on the student as a learner

The focus of our research is on the student as a learner, rather than on the instructor as a teacher. We have conducted investigations among various populations: students enrolled in introductory physics courses, in physics courses for underprepared students, in advanced undergraduate and graduate physics courses, in engineering courses, and in courses for K–12 teachers of physics and physical science. We explore what students can and cannot do and monitor their intellectual state as instruction progresses. We use two primary research methods: individual demonstration interviews that enable us to probe deeply into the way students think and widely administered written tests that provide data on prevalence. We supplement this information through less formal means, such as engaging students in dialogues, examining homework and written reports in detail, and observing in the classroom as students interact with one another and with their instructors. The results are used to guide the development of curriculum. Assessment is an integral part of the process and usually includes a comparison of student performance on post-tests and corresponding pretests.

III. INSIGHTS FROM RESEARCH AND TEACHING EXPERIENCE IN NONSTANDARD PHYSICS COURSES

The Physics Education Group has two major curriculum development projects: *Physics by Inquiry* (Wiley, 1996) and *Tutorials in Introductory Physics* (Prentice Hall, 1998).[3] Both owe much to our research and teaching experience in nonstandard physics courses. For more than 25 years, we have been conducting special courses during the academic year and in NSF Summer Institutes to prepare prospective and practicing teachers to teach physics and physical science by inquiry. Another group whom we have been able to teach in relatively small classes are students who aspire to science-related careers but whose prior preparation is inadequate for success in the required physics courses. Close contact with students in these special courses has provided us with the opportunity to observe the intellectual struggles of students as they try to understand important concepts and principles. We have found that students better prepared in physics often encounter the same difficulties as those who are not as well prepared. Since the latter are usually less adept in mathematics, it is easier to identify and probe the nature of common difficulties. Day-to-day interaction in the classroom has enabled us to explore in detail the nature of specific difficulties, to experiment with different instructional strategies, and to monitor their effect on student learning.

A. Research on student understanding: An example from electric circuits

Below, we briefly illustrate the type of research that underlies the development of curriculum by our group. The context is electric circuits. Our investigation of student understanding of this topic has extended over many years and has included individuals whose background in physics has ranged from the introductory to the graduate level.[4] Since the results are well known by now, only a summary is presented here.

In the question in Fig. 1(a), students are asked to rank the brightness of identical bulbs in three circuits. This question has been used in many different classes over many years. It has been given either before or after the usual treatment of this topic in lecture, textbook, and laboratory. Since the results have been essentially the same before and after standard instruction, they have been combined. As shown in Table I, only about 15% of more than 1000 introductory students have given the correct ranking (A=D=E>B=C). Similar results have been obtained from high school physics teachers and from university faculty in other sciences and mathematics. Only about 70% of the graduate teaching assistants have given a correct ranking. Analysis of the responses has re-

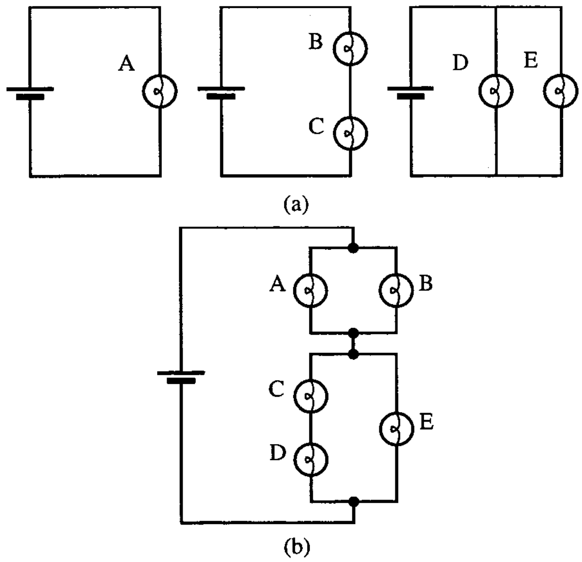


Fig. 1. Circuits used on questions given (a) after standard instruction on electric circuits and (b) after students had studied the material through guided inquiry. Students are asked to rank the bulbs from brightest to dimmest and to explain their reasoning. In both cases, they are told to treat the bulbs as identical and the batteries as identical and ideal.

vealed the widespread prevalence of two mistaken beliefs: the battery is a constant current source and current is "used up" in a circuit. Among all populations, the basic underlying difficulty seems to be the lack of a conceptual model for an electric circuit.

B. Basic instruction by guided inquiry

The nonstandard courses described above have provided the context for the development of *Physics by Inquiry* (PbI). This self-contained, laboratory-based curriculum helps students develop a coherent conceptual framework for important topics. PbI is not like a typical text, in that it does not present information and give explanations. The modules contain carefully structured experiments, exercises, and questions that are intended to engage students actively in the construction of important concepts and in their application to the physical world. The instructional approach can be characterized as guided inquiry. Although expressly designed for the preparation of K–12 teachers, PbI has also proved useful for providing a foundation in physics for underprepared students and nonscience majors.

The *Electric Circuits* module provides an example of how results from research are incorporated in PbI. As the students work through the module, they are guided in constructing a qualitative model for a simple circuit. In the process, specific difficulties identified through research are addressed.

Table I. Results from pretest on electric circuits shown in Fig. 1(a). All percentages are rounded to the nearest 5%.

	Undergraduates $N>1000$	Precollege teachers $N>200$	Faculty in other sciences and mathematics $N>100$	Graduate TAs $N\sim55$
Correct answer	15%	15%	15%	70%

C. Assessment of student learning

The instructional approach in *Electric Circuits* has proved effective with K–12 teachers at all levels. In Fig. 1(b) is an example of a post-test, given after students have worked through the relevant material. Students are asked to rank the brightness of identical bulbs (E>A=B>C=D). Elementary and middle school teachers generally have a weaker mathematical background than students in the introductory calculus-based course. Nevertheless, their post-test performance on this and other relatively complicated resistive circuits has regularly surpassed that of most physics and engineering students.

D. Commentary

We believe that the primary reason for the effectiveness of PbI is that students must go step-by-step through the reasoning needed to overcome conceptual hurdles and build a consistent coherent framework. There are also other features that we think are important. Collaborative learning and peer instruction are integrated into PbI. Students work with partners and in larger groups. Guided by the questions and exercises, they conduct open-ended explorations, perform simple experiments, discuss their findings, compare their interpretations, and collaborate in constructing qualitative models that can help them account for observations and make predictions. Great stress is placed on explanations of reasoning, both orally and in writing. The instructor does not lecture but poses questions that motivate students to think critically about the material. The appropriate response to most questions by students is not a direct answer but a question to help them arrive at their own answers.

IV. INSIGHTS FROM RESEARCH AND TEACHING EXPERIENCE IN STANDARD INTRODUCTORY COURSES

The topic of electric circuits is only one of many in which we have examined student understanding. Our investigations have spanned many topics at several levels of instruction with special emphasis on introductory physics.

A. Need for improvement in student learning

Faculty in introductory courses work hard at preparing lectures in which they give lucid explanations, show demonstrations, and illustrate problem-solving procedures. They expect that, in the process of learning how to solve standard physics problems, students are developing important concepts, integrating them into a coherent conceptual framework, and developing the reasoning ability necessary to apply the concepts in simple situations. It is also assumed that students are learning to relate the formalism of physics to objects and events in the real world. There is ample evidence from research, however, that students do not make nearly as much progress toward these basic goals as they are capable of doing. Few develop a functional understanding of the material they have studied.

The gap between the course goals and student achievement reflects a corresponding gap between the instructor and the students. In teaching introductory physics, many faculty proceed from where they are now or where they think they were as students. They frequently view students as younger

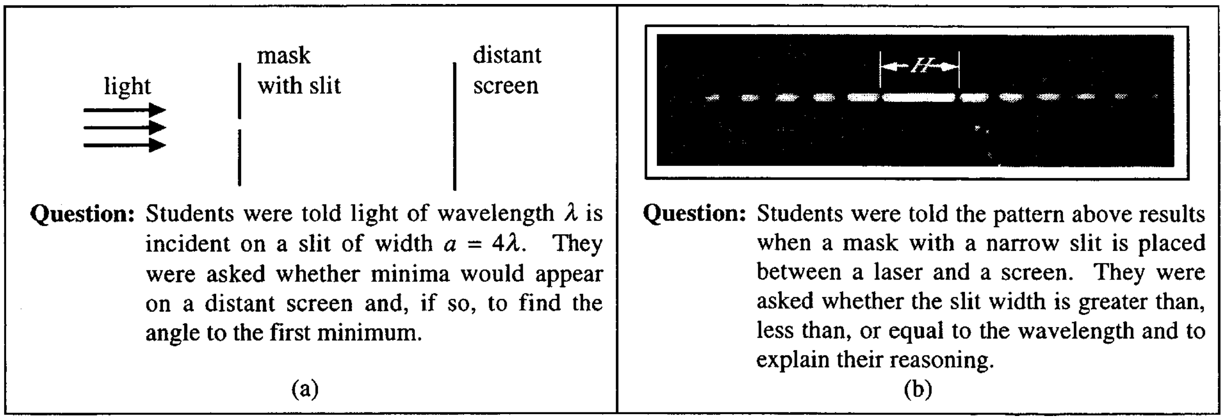


Fig. 2. Questions used to probe student understanding of diffraction after standard instruction in large introductory physics courses: (a) quantitative question and (b) qualitative question.

versions of themselves. This approach is particularly unsuitable for a typical introductory physics course in which fewer than 5% of the students will major in physics. For most, it is a terminal course in the discipline.

A functional understanding of physics connotes the ability to interpret and use knowledge in situations different from those in which it was initially acquired (the degree of difference increasing with educational level). Majors eventually develop this ability. Most students do not. Although faculty hope that they are helping students develop scientific reasoning skills, the type of problem solving that takes place in a typical introductory course is not consistent with this objective. Often the effect is to reinforce the common perception that physics is a collection of facts and formulas and that the key to solving physics problems is finding the right formulas. However, even correctly memorized formulas are likely to be forgotten after the course ends. An understanding of important physical concepts and the ability to do the reasoning necessary to apply them is of greater lasting value.

B. Motivation for tutorials

The success of *Physics by Inquiry* with teachers and other students motivated us to try to provide for students in standard introductory courses a modified version of the intellectual experience that this curriculum provides. However, the challenge of securing the mental engagement of students in a typical calculus-based or algebra-based course is much greater. The large size of these classes, the breadth of material covered, and the rapid pace preclude use of a laboratory-based, self-contained curriculum like *Physics by Inquiry*. Therefore, we decided to try to incorporate some of the important features of *PbI* in a curriculum that could be used to supplement the lectures and textbook of a standard calculus-based or algebra-based course. We wanted to produce materials that would be useful not only at our own university but in a wide variety of instructional settings. *Tutorials in Introductory Physics* has been our response to this challenge. Although this project was motivated by a desire to improve student learning in introductory physics, we and others have found that the same instructional approach also works well in more advanced courses.

V. RESEARCH-BASED GENERALIZATIONS ON LEARNING AND TEACHING

Our experience in research, curriculum development, and instruction has led to several generalizations on learning and teaching.[5] These are empirically based in that they have been inferred and validated through research. The early research and development of *Physics by Inquiry* formed the initial basis for the generalizations. Our later experience with PbI and *Tutorials in Introductory Physics* confirmed their validity and provided additional insights that broadened their applicability. The generalizations serve as a practical model for curriculum development by our group. Below we present several that have proved especially useful. The illustrative examples are from our investigation of student understanding in physical optics.[6] This long-term study involved undergraduates in introductory and more advanced courses, as well as physics graduate students.

A. Research-based generalizations on student learning

Examples from our research are given below as evidence for a few of the generalizations on student learning. Others are supported more broadly from our research base.

1. **Facility in solving standard quantitative problems is not an adequate criterion for functional understanding.**

Although experienced instructors know that there is a gap between what they teach and what is learned, most do not recognize how large the gap can be. The traditional measure for assessing student understanding is performance on standard quantitative problems. Since a significant portion of a typical class receives grades of A or B, instructors may con-

Table II. Results from quantitative and qualitative questions on single-slit diffraction shown in Fig. 2.

Undergraduate students		Graduate TAs
Quantitative question $N\sim130$	Qualitative question $N\sim510$	Qualitative question $N\sim95$
70% correct with correct angle	10% correct with correct explanation	55% correct with correct explanation

clude that students have understood the material at an acceptable level. However, the ability of students to obtain correct answers for numerical problems often depends on memorized algorithms. Liberal awarding of partial credit also may conceal lack of understanding.

Questions that require qualitative reasoning and verbal explanation are essential for assessing student learning. The importance of qualitative questions is demonstrated by all of our research. As illustrations, we consider some examples from physical optics. As part of our investigation, we tried to determine what students who have studied physical optics in a standard course can and cannot do. The two questions below pose essentially the same problem.

a. Quantitative question on single-slit diffraction. The question in Fig. 2(a) was given on an examination to about 130 students. They were told that light is incident on a single slit of width $a=4\lambda$. The students were asked to state if any minima would appear on a screen and, if so, to calculate the angle to the first minimum. Since the slit width is larger than the wavelength, minima would occur. The required angle can be obtained by using the equation $a \sin\theta=\lambda$, which yields $\theta=\sin^{-1}(0.25)\approx 14°$.

Approximately 85% of the students stated that there would be minima. About 70% determined the correct angle for the first minimum. (See the first column in Table II.)

b. Qualitative question on single-slit diffraction. For the question in Fig. 2(b), students were shown a single-slit diffraction pattern with several minima. They were told that the pattern results when a mask with a single vertical slit is placed between a laser (wavelength λ) and a screen. They were asked to decide whether the slit width is greater than, less than, or equal to λ, and to explain their reasoning. They could answer by referring to the equation for the angle θ to the first diffraction minimum. Since minima are visible, the angle to the first minimum is less than 90° and $a \sin\theta=\lambda$. Therefore, since $\sin\theta<1$, $a>\lambda$.

About 510 students, including the 130 who had been given the quantitative question, were asked this question after they had completed standard instruction on single-slit diffraction. Performance was poor. About 45% of the students made a correct comparison. Only 10% gave a correct explanation. (See the second column of Table II.) This same question was also posed in a graduate teaching seminar ($N\sim 95$). About half of the participants responded correctly with correct reasoning. (See the third column of Table II.)

c. Comparison of results from qualitative and quantitative questions. The difference in the way that the introductory students treated the two questions above provides some insight into what they typically can and cannot do. As can be seen from Table II, the success rate on the qualitative question was much lower than on the quantitative question. The 130 students who had previously been given the quantitative question performed at about the same level as those who had not had this experience. Apparently, the ability to solve numerical problems is not a reliable indicator of conceptual understanding.

2. **Connections among concepts, formal representations, and the real world are often lacking after traditional instruction.**

The ability to use and interpret formal representations (algebraic, diagrammatic, and graphical) is critical in physics. The responses to the qualitative question on single-slit diffraction demonstrate that many students could not relate the formula that they had memorized (or had available) for the location of diffraction minima to the diffraction pattern. Two examples that provide additional evidence of a failure to make connections between the phenomena and formalism of physical optics appear under the next generalization.

3. **Certain conceptual difficulties are not overcome by traditional instruction. (Advanced study may not increase understanding of basic concepts.)**

Research has shown that certain conceptual difficulties persist in spite of instruction. The two examples below indicate deep confusion about the different models for light and the circumstances under which a ray, wave, or particle model applies. All the students involved had received explicit instruction on at least the ray and wave models but seemed to have great difficulty in interpreting the information.

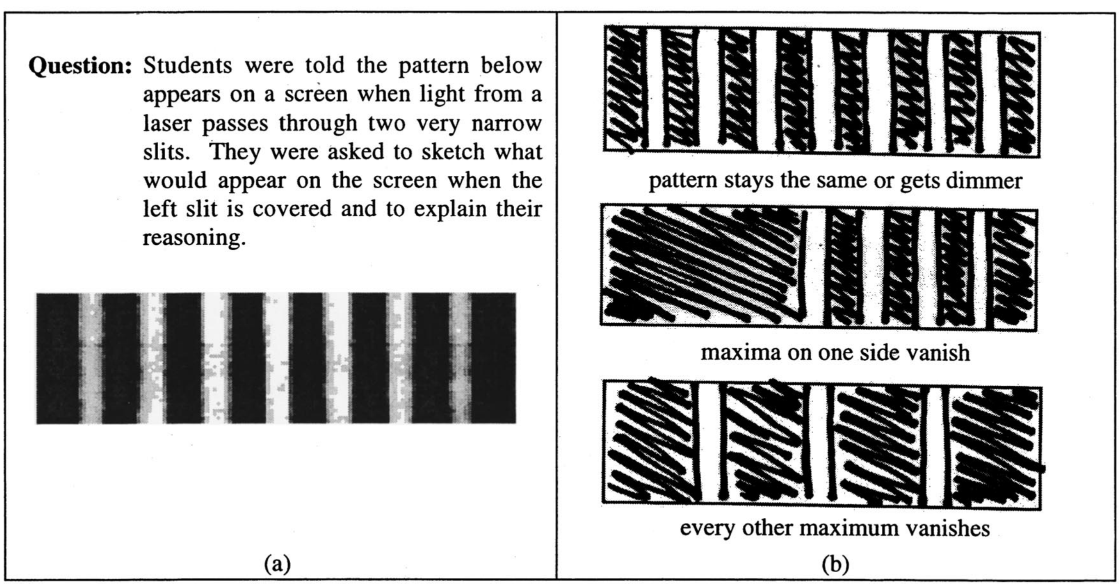


Fig. 3. (a) Question used to probe student understanding of double-slit interference. (b) Common incorrect diagrams drawn by students in response to the written question.

a. Qualitative question on double-slit interference. The students were shown a photograph of the central portion of a double-slit interference pattern in which all the maxima are of similar intensity. [See Fig. 3(a).] They were asked to sketch what would appear on the screen if the left slit were covered. To respond correctly, they needed to recognize that the minima are due to destructive interference of light from the two slits and that each slit can be treated as a line source. After the left slit is covered, the interference minima would vanish and the screen would be (nearly) uniformly bright.

This question was asked in several lecture sections of the calculus-based course ($N \sim 600$) with similar results before and after standard instruction. No more than about 40% of the students answered correctly. Overall, about 45% gave answers reminiscent of geometrical optics. Many claimed that the pattern would be the same, but dimmer. Others predicted that the maxima on one side would vanish, leaving a dark region, or that every other maximum would vanish. [See Fig. 3(b).]

b. Individual demonstration interview on single-slit diffraction. In addition to the written questions on single-slit diffraction, we conducted individual demonstration interviews. Of the 46 students who participated, 16 were from the introductory calculus-based course and 30 from a sophomore-level modern physics course. All were volunteers and had earned grades at or above the mean in their respective courses.

During the interviews, students were shown a small bulb, a screen, and a small rectangular aperture. They were asked to predict what they would see on the screen as the aperture is narrowed to a slit. Initially, the geometric image of the aperture would be seen. Eventually, a single-slit diffraction pattern would appear.

In responding to this and other questions, students from both courses often used hybrid models with features of both geometrical and physical optics. For example, some students claimed that the central maximum of the diffraction pattern is the geometric image of the slit and that the fringes are due to light that is bent at the edges. Another difficulty of both introductory and more advanced students was the tendency to attribute a spatial extent to the wavelength or amplitude of a wave. Many considered diffraction to be a consequence of whether or not light would "fit" through the slit. Some of the introductory students claimed that if the width of the slit were greater than the amplitude of the wave, light would be able to pass through the slit, but that if the slit width were less, no light could emerge. [See Figs. 4(a) and 4(b).] Some modern physics students extended these same ideas to photons distributed along sinusoidal paths. (See Fig. 5.) Their diagrams indicated that the photons would not get through the slit if the amplitude were greater than the slit width. In physical optics and other topics, we have found that study beyond the introductory level does not necessarily overcome serious difficulties with basic material. Unless explicitly addressed in introductory physics, these difficulties are likely to persist.

4. **A coherent conceptual framework is not typically an outcome of traditional instruction.**

Many students emerge from introductory physics without having developed a coherent conceptual framework for important basic topics. As has been discussed, our research on student understanding of electric circuits supports this generalization. The examples from physical optics that have been used as illustrations provide additional evidence.

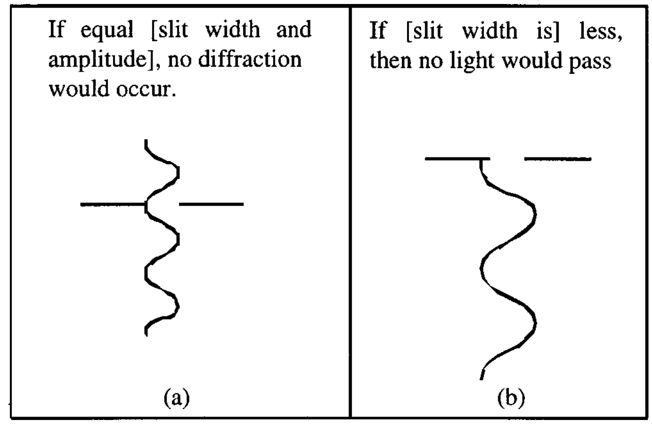


Fig. 4. Diagrams drawn by introductory students during interviews on single-slit diffraction to illustrate their belief that diffraction depends on the amplitude of the light wave: (a) amplitude less than or equal to the slit width and (b) amplitude greater than slit width.

Analysis of the results from the written questions and interviews on physical optics revealed the presence of a number of conceptual difficulties. Among these were: (1) the use of a hybrid model with features of both geometrical and physical optics, (2) a tendency to attribute to the amplitude or wavelength a spatial extent that determines whether light can "fit" through a slit, and (3) lack of recognition that an interference pattern results from two or more slits. Underlying these and other specific difficulties was one of fundamental importance: the failure of students to relate diffraction and interference effects to differences in path length (or phase). They had not developed a basic wave model that they could use to account for the diffraction and interference of light in the far-field limit.

Having a wave model for light would seem to be a prerequisite for understanding the wave nature of matter. Thus, there are clear implications for reform efforts directed toward introducing topics from modern physics into the introductory course. Results from research indicate that difficulties with advanced physics often have their roots in elementary material.

5. **Growth in reasoning ability often does not result from traditional instruction.**

An important factor in the difficulties that students have with certain concepts is an inability to do the qualitative reasoning that may be necessary for applying these concepts. Students often do not recognize the critical role of reasoning, nor understand what constitutes an explanation in physics. Our research has provided many illustrations. For example,

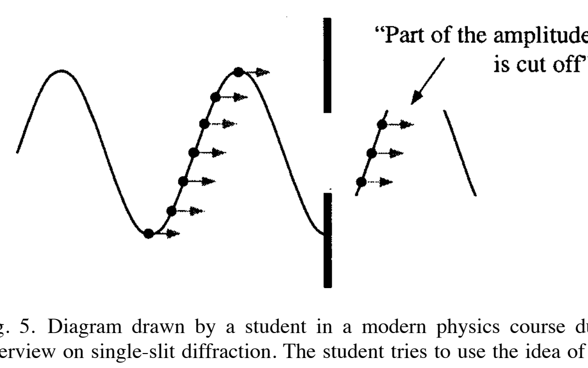


Fig. 5. Diagram drawn by a student in a modern physics course during interview on single-slit diffraction. The student tries to use the idea of photons to account for diffraction.

on the quantitative question on single-slit diffraction discussed earlier, many students used the single-slit diffraction formula to give a correct response for the location of the first diffraction minimum. Yet on the qualitative problem, many of these same students could not do the reasoning necessary to conclude that the presence of diffraction minima in the photograph implied that the slit width must be greater than the wavelength.

6. **Teaching by telling is an ineffective mode of instruction for most students.**

This generalization is based on results from our investigations of student understanding in mechanics, electricity, magnetism, electromagnetic waves, geometrical and physical optics, hydrostatics, and thermodynamics. In all of these topics, we have found that on certain types of qualitative questions student performance is essentially the same: before and after standard instruction by lecture and textbook, in calculus-based and algebra-based physics, with or without demonstrations, with or without a standard laboratory, in large and small classes, and regardless of the popularity of the instructor as a lecturer.

B. Research-based generalizations on teaching

The generalizations on student learning have implications for teaching. Our experience in developing curriculum and testing its effectiveness with students has led to a corresponding set of research-based generalizations on teaching. Below, the generalizations on student learning are repeated. Each is followed by one on teaching [in bold italics].

1. Facility in solving standard quantitative problems is not an adequate criterion for functional understanding. ***Questions that require qualitative reasoning and verbal explanation are essential for assessing student learning and are an effective strategy for helping students learn.***

As has been discussed, the traditional forms of instruction seem to be inadequate for helping most students develop a functional understanding of basic topics in physics. Hearing lectures, reading textbooks, solving quantitative problems, seeing demonstrations, and doing experiments often have surprisingly little effect on student learning. We have found that an effective instructional approach is to challenge students with qualitative questions that cannot be answered through memorization, to help them learn how to respond to such questions, and to insist that they do the necessary reasoning by not supplying them with answers.

2. Connections among concepts, formal representations, and the real world are often lacking after traditional instruction. ***Students need repeated practice in interpreting physics formalism and relating it to the real world.***

Most instructors recognize that students need help in relating the concepts and formal representations of physics to one another and to physical phenomena. However, illustrative examples and detailed explanations are often ineffective. Analogies obvious to instructors are often not recognized by students. For example, in developing our curriculum on physical optics, we found that many students needed explicit guidance in transferring their experience with two-source interference in water to double-slit interference in light.

3. Certain conceptual difficulties are not overcome by traditional instruction. (Advanced study may not increase understanding of basic concepts.) ***Persistent conceptual difficulties must be explicitly addressed in multiple contexts.***

Some difficulties that students have in learning a body of material are addressed through standard instruction or gradually disappear as the course progresses. Others are highly resistant to instruction. Some are sufficiently serious that they may impede, or even preclude, development of a functional understanding. For example, the belief that the amplitude of a light wave has a spatial extent or that the wave is a carrier of photons makes it impossible to develop a correct wave model for light. (See Figs. 4 and 5.)

Our experience indicates that warning students not to make particular errors is ineffective. For most students, assertions by an instructor make no difference. Avoiding situations likely to evoke errors by students, or providing algorithms that they can follow without thinking, may conceal latent difficulties that will surface at some later time. If faulty reasoning is involved, merely correcting an error is useless. Major conceptual change does not take place without a significant intellectual commitment by students.

An instructional strategy that we have often found effective for securing the mental engagement of students can be summarized as: *elicit*, *confront*, and *resolve*. The first step is to create a situation in which the tendency to make a known common error is exposed. After the students have been helped to recognize a resultant inconsistency, they are required to go through the reasoning needed to resolve the underlying difficulty. Since single encounters are seldom sufficient for successfully addressing serious difficulties, it is necessary to provide students with additional opportunities to *apply*, *reflect*, and *generalize*.

A word of caution is necessary because frequent use of the terms "misconceptions" and "misconceptions research" has trivialized the intellectual problem. The solution is not a matter of identifying and eradicating misconceptions. The intellectual issues are much deeper. Misconceptions are often symptoms of confusion at a fundamental level.

4. A coherent conceptual framework is not typically an outcome of traditional instruction. ***Students need to participate in the process of constructing qualitative models and applying these models to predict and explain real-world phenomena.***

Among the goals of a physics course is the development of physical concepts and an understanding of their relationships to one another and to the real world. Helping students develop a sound conceptual understanding is not simply a matter of defining concepts, presenting models, and illustrating their application. Often students cannot identify the critical elements or recognize inconsistencies with their ideas. A spiral approach in which models are continually refined is helpful but may not necessarily lead to coherence. Serious conceptual difficulties that preclude development of a consistent model must be addressed.

We have found that an effective strategy for helping students understand the relationships and differences among concepts is to engage them actively in the model-building process. As has been discussed in the context of electric circuits, this approach also provides some direct experience with the nature of scientific inquiry.

5. Growth in reasoning ability often does not result from traditional instruction. ***Scientific reasoning skills must be expressly cultivated.***

Conceptual models in physics are often inseparably linked with particular lines of reasoning. Hence, instruction should address both concurrently. The *Electric Circuits* module in PbI is an example. The physical optics tutorials to be discussed later are another. In both instances, students go through the reasoning necessary for developing the concepts.

6. Teaching by telling is an ineffective mode of instruction for most students. ***Students must be intellectually active to develop a functional understanding***.

All of the generalizations on learning and teaching support this last set. The extent to which these hold is often not adequately appreciated by faculty. Meaningful learning requires the active mental engagement of the learner. The role of the lecturer is clearly important. He or she is the one who motivates the students and the one to whom they look for guidance about what they need to learn. The lecturer, however, cannot do their thinking for them. The students must do it for themselves. Some are reluctant to do so; others do not know how. For most students, the study of physics is a passive experience.

It seems to be a natural instinct for instructors to believe that if the explanations they give are sufficiently clear and complete, students will learn. To this end, lecturers work at perfecting their presentations. Our experience has been, however, that the effort involved does not result in significant gain for most students. If they learn, it seems to be primarily because they have been willing and able to tackle the material with intellectual intensity. Both *Physics by Inquiry* and *Tutorials in Introductory Physics* are designed to engage students at a sufficiently deep intellectual level for meaningful learning to occur.

VI. APPLICATION OF RESEARCH-BASED GENERALIZATIONS TO THE DEVELOPMENT OF CURRICULUM

The development of all instructional materials by our group is the result of an iterative cycle that has three components: research on student understanding, use of the findings to guide the development of curriculum, and assessment of student learning. Research and curriculum development for PbI and for the tutorials are mutually reinforcing. Research motivated by one of the projects enriches the other. Similarly, instructional strategies that work well in one curriculum often, with some modification, work well in the other. To ensure applicability beyond our own university, all of our instructional materials are also tested at pilot sites. Some have environments similar to ours; others have different instructional settings. Experience at our university and at pilot sites has shown that certain conditions are necessary for the successful implementation of curriculum. The discussion here is limited, however, to those intellectual aspects that bear directly on student learning.

A. Description of the tutorials

Tutorials in Introductory Physics is designed for use in the small-group sections often associated with large lecture courses. The word *tutorial* was chosen to distinguish the type of instruction in the tutorials from more traditional recitation, discussion, quiz, or problem-solving sections. The usual procedure in such sections is for the instructor or TA to work problems, ask students to solve problems, or respond to questions (often with a mini-lecture). The tutorials are very different in purpose and in structure. They incorporate some of the critical features that we believe have contributed to the effectiveness of PbI.

The tutorials provide a context for our ongoing research and curriculum development at the introductory level and beyond. They address the questions: Is the standard presentation of an important topic in textbook and lecture adequate to develop a functional understanding? If not, what can be done? The emphasis in the tutorials is on constructing concepts, on developing reasoning skills, and on relating the formalism of physics to the real world, not on transmitting information and solving standard problems. The tutorials provide experience in learning through guided inquiry. Less detailed and thorough than PbI, they are better able to fit the constraints of large-scale instruction. The tutorials target critical concepts and skills that are essential for developing a functional understanding of important topics and that are known through research and teaching experience to present difficulty to students.

Each tutorial consists of four components: pretest, worksheet, homework, and post-test. The sequence begins with a pretest (so named because it precedes the tutorial although the material has usually been covered in lecture). The pretests have several purposes that include: to alert students to what they need to know and be able to do, to set the stage for the associated tutorial, and to inform the course lecturers and tutorial instructors about the intellectual state of the students. Pretests are not returned to the students. They are expected to be able to answer the questions by working through the tutorials and related homework.

During the tutorial sessions, about 20–24 students work collaboratively in groups of three or four. The structure is provided by tutorial worksheets that contain questions that try to break the reasoning process into steps of just the right size for students to stay actively involved. If the steps are too small, little thinking may be necessary. If too large, the students may become lost unless an instructor is by their side. The tutorial instructors do not lecture or give answers but assist students by posing questions to guide them through the necessary reasoning. Tutorial homework assignments help reinforce the ideas developed during the tutorial. A significant portion of every course examination requires the kind of qualitative reasoning and verbal explanations that characterize the tutorials.

B. Preparation of tutorial instructors

The tutorials require ongoing preparation in both the subject matter and instructional method of the tutorial instructors (mostly graduate Teaching Assistants but also undergraduates and volunteer post-docs). Although they can provide assistance with end-of-the-chapter problems, TAs generally have not thought deeply enough about the concepts nor gone carefully enough through the required chain of reasoning to be able to help introductory students develop a functional understanding of the material. Results from research indicate that study beyond the introductory level does not necessarily lead to a deeper understanding of basic topics. We have found that advanced students not only have conceptual difficulties with special relativity and quantum mechanics but also with topics in introductory physics.

Like most teachers, TAs tend to teach as they were taught. If they are to help undergraduates learn physics by guided inquiry, they need to experience this instructional approach and reflect upon the rationale. This opportunity is provided on a weekly basis in a required graduate teaching seminar led by our group. The seminar is conducted on the same material and in the same manner that the tutorial instructors are expected to teach. The TAs take the same pretests as the introductory students. Their performance provides us with a measure of their level of understanding and helps set a reasonable goal for a tutorial. We consider a tutorial to be successful when the post-test performance of the introductory students matches, or exceeds, the pretest performance of the TAs.

C. Supplementary instruction by guided inquiry: Example from physical optics

The research-based generalizations discussed above and others drawn from experience have proved valid and useful for our continuing development of curriculum. We illustrate their application in the context of physical optics. Other topics could serve equally well.[7]

Underlying the specific conceptual difficulties in physical optics was the failure of students to recognize the role of the difference in path length (or phase) in determining the maxima and minima of diffraction and interference patterns. To address this fundamental difficulty and others that are more specific, we developed a series of tutorials that guide students through the development of a simple wave model that they can use to account for diffraction and interference effects. A more complete discussion of these tutorials and of the rationale that guided their development can be found in previously published papers.

The series begins with interference in the context of water. Waves in a ripple tank are much less abstract than light waves. This environment forms a visual representation of wave fronts and provides a framework in which students can derive the mathematical relationships for locating the maxima and minima of an interference pattern. We knew from previous research that students often do not apply the principle of superposition properly. By investigating what happens when water waves combine under different conditions, we hoped that they might be better able to apply superposition to light. We found, however, that the analogy often eludes students. Consequently, the tutorials were modified to provide explicit help in making the connection between water waves and light waves. In later tutorials, the students extend their wave model to interference from more than two slits, single-slit diffraction, and combined interference and diffraction.

D. Assessment of student learning

As mentioned earlier, our primary means of assessment of student learning is through comparison of student performance on post-tests and corresponding pretests. These also provide the detailed feedback needed for the development of curriculum. The pretests and post-tests consist mostly of qualitative questions for which explanations are required. As has been illustrated, such questions are often a better test of student understanding than more difficult problems that can be solved by manipulation of formulas. Moreover, the feedback provided by numerical problems is often not very useful for improving instruction. Multiple-choice and true–false questions (whether quantitative or qualitative) have this same disadvantage.

The post-tests may or may not be similar to the pretests. Our research has shown that prior experience with a pretest has virtually no effect on student performance on a post-test. The post-tests require an understanding of the concepts and are designed so that (like the pretests) they cannot be answered on the basis of memorization.

The pretest and post-test below have been used in assessing the tutorial on multiple-slit interference. However, since learning is cumulative, the effect of each tutorial cannot be isolated from the preceding ones in the series.

1. Pretest on multiple-slit interference

On the pretest, the students are shown the central portion of the pattern formed by light incident on a mask with two very narrow slits separated by a distance d. [See Fig. 6(a).] A point on the first interference maximum, B, is marked. The students are told that the two-slit mask is replaced by a three-slit mask with the same separation d between adjacent slits. They are asked whether point B would still be a point of maximum constructive interference. This question requires application of the ideas of path length difference and superposition. From the pattern, it can be seen that light from two slits a distance d apart is in phase at point B. Since the distance between adjacent slits in the three-slit mask is also d, light from all three slits is in phase at point B. Thus point B will still be a point of maximum constructive interference and will be brighter than before. [See Fig. 6(b).]

This question was given to about 560 students, either before or after lecture instruction. Since the results were similar, the data have been combined in the first column of Table III. About 30% of the students have responded correctly with fewer than 5% using correct reasoning. Most students have failed to consider path length differences and superposition. About 60% of the participants in the graduate teaching seminar have answered correctly with about 25% giving correct explanations. (See the third column of Table III.)

2. Post-test on multiple-slit interference

In one post-test question, students are shown the same double-slit interference pattern as was used for the pretest. [See Fig. 6(a).] In this case, however, they are asked how the intensity at point B changes when a third slit is added a distance $d/2$ to the right of the rightmost slit. The students need to recognize that the waves from the original two slits are in phase at point B. When the third slit is added, the waves from this slit are 180° out of phase with the waves from both of the other slits. Therefore, the intensity at point B decreases. [See Fig. 6(c).] This question requires students to extend their thinking to a situation beyond their experience, i.e., when the slits are not evenly spaced.

The results of the post-test question are shown in the second column of Table III. About 80% of the students ($N=405$) have stated that the intensity at point B decreases when the third slit is added. About 40% have given correct reasoning. The improvement indicates that the tutorial helps students learn how to take into account the path length (or phase) difference in a situation in which they cannot resort to a formula. As shown in Table III, the introductory students did better on the post-test than the teaching assistants on the pretest, a criterion that we have set for a successful tutorial.

E. Effectiveness of the tutorials

The tutorials have had a very positive effect on the ability of students to solve qualitative problems of the type illustrated. For most students, the post-tests have shown marked improvement over the corresponding pretests. The post-test performance of the undergraduates has often matched (and sometimes surpassed) that of the graduate students on the pretests. In spite of less time devoted to quantitative problem solving, students who have worked through the tutorials do

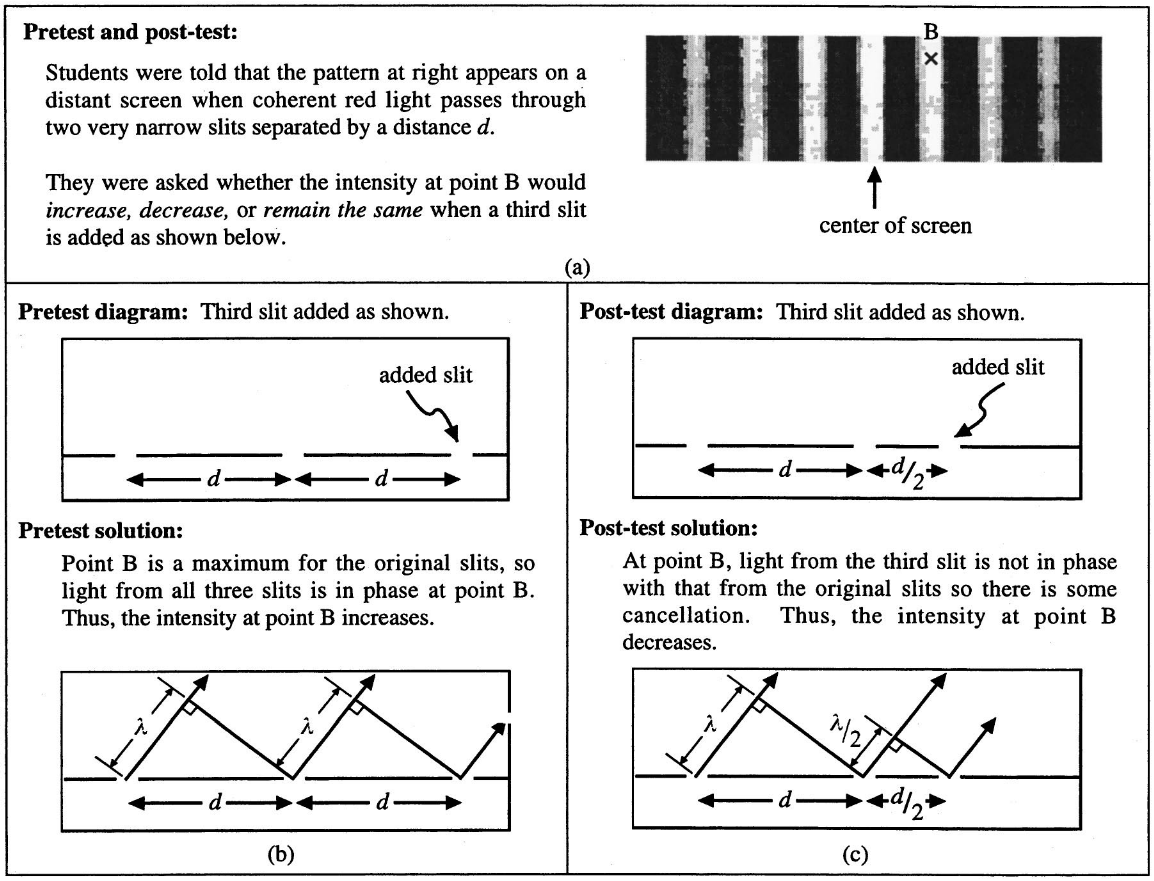


Fig. 6. (a) Basic question for pretest and post-test on multiple-slit interference. (b) Pretest diagram and solution. (c) Post-test diagram and solution.

somewhat better on standard numerical problems than those who have not had this experience. On quantitative problems that require understanding of the concepts, tutorial students have done much better than similar nontutorial students. Moreover, there is evidence that the type of intellectual effort demanded by the tutorials leads to a higher retention rate than that from standard instruction.

The particular instructional approach incorporated in the tutorials is only one of several that can be used to engage students actively in learning physics. *Physics by Inquiry*, in which all instruction emphasizes conceptual understanding and reasoning ability, is even more effective. The tutorials, however, require relatively little modification of the traditional mode. They have proved to be practical, flexible, and sustainable.

Table III. Results from pretest and post-test for tutorial on multiple-slit interference shown in Fig. 6. In both cases, a third slit was added to a mask containing two slits a distance d apart. On the pretest, the third slit was added a distance d to the right of the rightmost slit; on the post-test the third slit was added a distance $d/2$ to the right.

	Undergraduate students		Graduate TAs
	Pretest (d) $N\sim560$	Post-test ($d/2$) $N\sim405$	Pretest (d) $N\sim55$
Correct without regard to reasoning	30%	80%	60%
Correct with correct reasoning	<5%	40%	25%

F. Commentary

Careful assessment of student learning should be an integral part of the development of all printed and computer-based materials. It is difficult to develop curriculum that yields reliable results when used by different instructors. Therefore, unless instructors can devote a long-term effort to the design, testing, and refinement of new materials, it is best to take advantage of existing curriculum that has been thoroughly evaluated. It is important to know what has been accomplished and not expend resources in recreating what has been done well.

VII. CONCLUSION

Research in physics education can provide a guide for setting standards for student learning that are more rigorous than the generally accepted criterion of success in solving quantitative problems. It is possible to help students meet higher standards than most instructors often tacitly accept. As already mentioned, there is considerable evidence that time spent on developing a sound qualitative understanding does not detract from, and often improves (sometimes significantly), the ability to solve quantitative problems. Students should be expected to develop a coherent conceptual framework that enables them to determine in advance the type of answer that they should obtain in a quantitative problem. Therefore, the types of intellectual goals that have been set forth, both explicitly and implicitly, do not represent a "dumbing down" of standards, a charge often levied at attempts to modify traditional physics instruction. On the contrary, an increased emphasis on qualitative reasoning means that we are setting *much higher* standards.

Research can be the key to student learning. Without a sound base for informing the development of curriculum, we lack the knowledge necessary to make cumulative progress in improving instruction. We need to increase our understanding of how students think about traditional and contemporary topics. This information can provide a basis for designing instruction to achieve the specific goals of physics courses. Research on how students learn can also lead to insights about how to promote the development of some more general intellectual goals. We would like to help students understand the nature of scientific models and the scientific method through which they are developed. We want them to know the difference between what is and what is not a scientific explanation and to be able to distinguish between explanations based on scientific reasoning and arguments based on personal belief or popular opinion. Students need to recognize the kinds of questions that they must ask themselves to determine whether they understand a concept or line of reasoning and, if they do not, to formulate questions that can help them improve their understanding. Being able to reflect on one's thinking and to learn on one's own is a valuable asset that transcends the learning of physics. The study of physics offers many opportunities to cultivate the ability to engage in scientific, critical, and reflective thinking. Thus, research can be the key to setting higher (yet realistic) standards, to helping students meet expectations, and to assessing the extent to which the goals for student learning are met.

We can be greatly encouraged by the positive change that has occurred in the physics community within the last decade. Research in physics education has had an increasing influence on the way physics is taught. Faculty have drawn upon the results in producing new textbooks and revised versions of established texts. Research has also had a direct impact on the development of innovative instructional materials that have been shown to be effective. The results have been reported at professional meetings and in readily accessible journals. At meetings of professional organizations, sessions on research are well attended.

Many departments currently devote seminars and colloquia to physics education and, in particular, to research on the learning and teaching of physics. Faculty have been receptive and interested. Today, there are several universities in which graduate students can earn a Ph.D. in physics for research in this area. The rate of publication is increasing. The evolution in climate is reflected in the actions of physics-related professional organizations. In May 1999, the Council of the American Physical Society passed a resolution in support of physics education research as an appropriate field for scholarly inquiry by faculty in physics departments. In December 1999, the American Institute of Physics, the American Physical Society, and the American Association of Physics Teachers (along with others) endorsed a statement urging physical science and engineering departments to become actively engaged in the preparation of K–12 teachers. We have come a long way and, with research as a guide, can look forward to continued progress in physics education.

ACKNOWLEDGMENTS

This Oersted Medal is the result of my collaboration with many members of the Physics Education Group, past and present. Special thanks are due to the current faculty in the group: Paula R. L. Heron, Peter S. Shaffer, and Stamatis Vokos. I would like to recognize the early intellectual influence of Arnold B. Arons and the invaluable support of Mark N. McDermott. I also want to express my appreciation to the Department of Physics and the University of Washington. I am grateful to the National Science Foundation for enabling our group to conduct a coordinated program of research, curriculum development, and instruction.

[1]F. K. Richtmyer, "Physics is Physics," Am. Phys. Teach. **1** (1), 1–5 (1933).

[2]For an overview, see L. C. McDermott and E. F. Redish, "Resource Letter: PER-1: Physics Education Research," Am. J. Phys. **67**, 755–767 (1999). The Resource Letter emphasizes research conducted among university students and contains relatively few entries below the college level.

[3]L. C. McDermott and the Physics Education Education Group at the University of Washington, *Physics by Inquiry* (Wiley, New York, 1996); L. C. McDermott, P. S. Shaffer, and the Physics Education Group at the University of Washington, *Tutorials in Introductory Physics*, First Edition (Prentice Hall, Upper Saddle River, NJ, 2002).

[4]L. C. McDermott and P. S. Shaffer, "Research as a guide for curriculum development: An example from introductory electricity. I. Investigation of student understanding," Am. J. Phys. **60**, 994–1003 (1992); Printer's erratum to Part I **61**, 81 (1993); and P. S. Shaffer and L. C. McDermott, "Research as a guide for curriculum development: An example from introductory electricity. II. Design of instructional strategies," *ibid*. **60**, 1003–1013 (1992). In addition to the examples given in these papers, reference is made to some of the research reported in L. C. McDermott, P. S. Shaffer, and C. P. Constantinou, "Preparing teachers to teach physics and physical science by inquiry," Phys. Educ. **35** (6), 411–416 (2000).

[5]L. C. McDermott, "Guest Comment: How we teach and how students learn—A mismatch?," Am. J. Phys. **61**, 295–298 (1993). The discrepancy between teaching and learning is also discussed in L. C. McDermott, "Millikan Lecture 1990: What we teach and what is learned—Closing the gap," *ibid*. **59**, 301–315 (1991), as well as in other articles by the Physics Education Group.

[6]B. S. Ambrose, P. S. Shaffer, R. N. Steinberg, and L. C. McDermott, "An investigation of student understanding of single-slit diffraction and double-slit interference," Am. J. Phys. **67**, 146–155 (1999); K. Wosilait, P. R. L. Heron, P. S. Shaffer, and L. C. McDermott, "Addressing student difficulties in applying a wave model to the interference and diffraction of light," Phys. Educ. Res., Am. J. Phys. Suppl. **67**, S5–S15 (1999). In addition to the specific examples in these papers, reference is made to some of the research reported in B. S. Ambrose, P. R. L. Heron, S. Vokos, and L. C. McDermott, "Student understanding of light as an electromagnetic wave: Relating the formalism to physical phenomena," Am. J. Phys. **67**, 891–898 (1999); S. Vokos, B. S. Ambrose, P. S. Shaffer, and L. C. McDermott, "Student understanding of the wave nature of matter: Diffraction and interference of particles," Phys. Educ. Res., Am. J. Phys. Suppl. **68**, S42–S51 (2000).

[7]See Ref. 2. The Resource Letter includes references that report on work done by our group in the context of a broad range of topics. These articles provide evidence that the results for other topics are consistent with those in the cited articles on electric circuits and physical optics.

Research in Physics Education

APS News 7 (1) January 1998

by Lillian C. McDermott, Department of Physics, University of Washington

Frequent references in the media to a national crisis in science education have focused public attention in the U.S. on the need to improve instruction at the elementary, middle, and high school levels. In the physics community, this concern extends upward to undergraduate and graduate physics courses, in which there have been steady declines in enrollment. This situation has created an environment at universities that is highly supportive of educational reform, especially in introductory courses. As physicists who care about the future of our profession, we may feel that we cannot afford to forego the present opportunity for improving instruction, but we are also likely to be skeptical. In light of past experience, we may question the possibility of significant, sustainable change. There is reason to believe, however, that the chance for success is now better than before.

During the past 20 years, research in physics education has emerged as a new field for scholarly inquiry by physicists. There are now a number of physics departments in which there are faculty and graduate students whose primary research interest is in physics education.[1] Their work is greatly expanding our knowledge of how students learn physics and has the potential for making significant contributions to the improvement of instruction. The purpose of this article is to provide an overview of this discipline-specific research.[2,3]

Role for Research

Perhaps the most important contributions that research can make to physics education is to investigate the relationship between teaching and learning and to strengthen the link. Systematic investigations have demonstrated that many students emerge from introductory physics without a functional understanding of basic physical concepts. Similar conceptual and reasoning difficulties have been identified among student populations that differ greatly in educational background. Analysis of the data indicates that many difficulties are more pervasive than is commonly realized. Highly resistant to instruction, some of these persist to the graduate level. Studies have been conducted with widely administered multiple-choice tests, as well as through in-depth investigations involving smaller numbers of students. The results, which are consistent and replicable, suggest that there is a need for change in the way that physics has been traditionally taught.

Research has been used to guide the design of instructional strategies that are better matched to the needs and abilities of students. Some of the curriculum that has been developed has been shown to lead to significant, reproducible gains in student understanding. Below are some generalizations that are consistent with results from investigations conducted before, during, and after instruction.[4]

Generalizations from Research on Teaching and Learning

Teaching by telling is an ineffective mode of instruction for most students.

In recalling how they were inspired by introductory physics, there is a tendency among many instructors to view students as younger versions of themselves. In actual fact, this description fits only the small minority who become physics majors (<5%). The vast majority of the students, for whom this is a terminal course, are not inherently motivated to confront the challenges presented by physics. Most listen passively to lectures without becoming intellectually engaged. Unlike the majors, they will have no further opportunity to learn the material. Among these students are future high school teachers who can be expected to teach what they believe they learned in the same way that they were taught.

Facility in solving standard quantitative problems is not an adequate criterion for functional understanding.

This criterion is the one most often used in physics instruction as a measure of mastery of the subject. As course grades attest, many students who complete a typical introductory course can solve such problems satisfactorily. However, they are often dependent on formulas that they are unable to apply to situations not previously memorized. Questions that require qualitative reasoning and verbal explanation provide a much better indication of student understanding. It has been shown that an emphasis on concept development does not detract from, and may even improve, the ability of students to solve quantitative problems. Less time to practice on such problems seems to have no adverse effect. Paying more attention to qualitative problems, however, does not remove the need for helping students learn how to solve quantitative problems.

Certain conceptual difficulties are not overcome by traditional instruction.

Some student difficulties disappear during the normal course of instruction. Others seem to be highly resistant to change. If sufficiently serious, they may preclude meaningful learning, even though performance on standard quantitative problems may be unaffected. There is a great deal of evidence that deep-seated difficulties cannot be overcome through assertion by the instructor. Significant conceptual change does not occur unless the intellectual involvement of students is at a sufficiently deep level.

Growth in reasoning ability does not usually result from traditional instruction.

Most courses taught by lecture tend to reinforce a perception of physics as a collection of facts and formulas. Students often do not recognize the critical role of reasoning, nor understand what constitutes a physical explanation. Most cannot do the qualitative reasoning necessary to apply the concepts taught to situations not expressly memorized. It has been shown, however, that this skill can be developed if students are given practice in solving qualitative problems and in explaining their reasoning. This outcome could be the most important benefit that non-majors could derive from a physics course.

Connections among concepts, formal representations and the real world are often lacking after traditional instruction.

Many instructors tend to underestimate this problem and believe that it can be successfully addressed through standard laboratory experiments and lecture demonstrations. Results from research indicate that this is not the case. Students need repeated practice in explicitly making these connections themselves.

A coherent conceptual framework is not typically an outcome of traditional instruction.

Perhaps the most serious difficulty among introductory students is the failure of many to integrate related concepts. The lack of a coherent framework may pass undetected because mathematical manipulation often suffices for the solution of standard problems. To be able to apply a concept in a variety of contexts, students must be able not only to define the concept but also to recognize its relevance to a given physical situation. They are unlikely to develop this facility, however, unless they themselves have gone through the steps necessary to construct the concept.

Implications for Physics Education

The generalizations above may not be surprising to experienced instructors, many of whom may have made serious efforts to address the issues that have been raised. The problem has been that in many cases the difficulties identified by instructors and the successes and failures of the instructional strategies that they have tried have not been well documented and readily accessible to others. The information has tended to be largely anecdotal. Often insufficient detail is given for replication of an apparently successful technique outside of the environment in which it was developed. The research-based approach to instruction that has been described reflects a perspective in which physics teaching is viewed as science as well as art. It is necessary to break the traditional pattern in which instructors develop their own intuition about effective teaching by practicing on students, often repeating the same mistakes that colleagues have made before. The goal is not a static curriculum since response to new knowledge and technologies is essential. Curriculum development should be ongoing; but it should also be iterative, building on past accomplishments. Rigorous assessment and careful reporting of the results should be an integral part of the development process.

Too often the quality of instruction is judged on the basis of student and teacher enthusiasm. This is not a valid indicator. There is a need to examine carefully what students have actually learned. To make this assessment, physics instructors should draw on findings from research. Unless we are willing to apply the same rigorous standards of scholarship to issues related to learning and teaching that we regularly apply in more traditional research, the present situation in physics education is unlikely to change.

Footnotes

1. The Physics Education Group at the University of Washington is one example. The graduate students in the group are admitted to the Department of Physics through regular channels, take all of the same courses as all other students, pass the same examinations required of all other graduate students, and receive the Ph.D. in physics. The employment record for students who have graduated from this program is excellent. They are currently working in physics departments, in museums, and in industry. Other physics departments in which graduate students do research in physics education include the University of Maryland, North Carolina State University, The Ohio State University, Kansas State University, the University of Nebraska, and the University of Maine at Orono.

2. Some of the ideas included in this paper have been expressed in two previously published articles. For a more extended discussion, see *Millikan Lecture 1990: *What we teach and what is learned *Closing the gap,** L. C. McDermott, Am. J. Phys. 59, 301-315 (1991) and **How we teach and how students learn ? A mismatch?**L.C. McDermott, Am. J. Phys. 61, 295-298 (1993).

3. Listed below are a few articles published during the last five years in the American Journal of Physics and The Physics Teacher that relate to some of the issues discussed in this paper. The list is not intended to be comprehensive but to provide an entry point to the literature for physicists who are not familiar with this area of research. D. Hestenes, M. Wells and G. Swackhamer, **Force Concept Inventory,**The Physics Teacher 30:3, 141-158 (1992); R. Beichner, **Testing student interpretation of kinematics graphs,**Am. J. Phys. 62, 750-762 (1994); E. F. Redish, **Implications of cognitive studies for teaching physics,**Am. J. Phys. 62, 796-803 (1994); F. Reif, **Millikan Lecture 1994: Understanding and teaching important scientific thought processes,** Am.J. Phys. 63,17-32 (1995); R. R. Hake, **Interactive-engagement Vs traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,** Am. J. Phys. (to be published); and R. Thornton and D. Sokoloff, **Assessing and improving student learning of Newton*s laws: The Force and Motion Conceptual Evaluation and the evaluation of active learning laboratory and lecture curricula,** Am. J. Phys. (to be published in).

4. Supporting evidence for these generalizations can be found in reports of research by the Physics Education Group. Articles published during the last five years include:

L.C. McDermott and P. S. Shaffer, **Research as a guide for curriculum development: An example from introductory electricity. Part I: Investigation of student understanding,**Am. J. Phys. 60, 994-1003 (1992); P.S. Shaffer and L.C. McDermott, **Research as a guide for curriculum development: An example from introductory electricity. Part II: Design of an instructional strategy,**Am. J. Phys. 60, 1003-1013 (1992); Printer*s erratum to Part I, Am. J. Phys. 61, 81 (1993); L.C. McDermott, P.S. Shaffer and M.D. Somers, **Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood*s machine,**Am. J. Phys. 62, 46-55 (1994); R. N. Steinberg, G. E. Oberem and L. C. McDermott, **Development of a computer-based tutorial on the photoelectric effect,**Am. J. Phys. 64, 1370-1379 (1996); and T. O*Brien Pride, S. Vokos and L.C. McDermott, **The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems,**Am. J. Phys. (to be published 1997).

Student understanding of the work-energy and impulse-momentum theorems

Ronald A. Lawson[a] and Lillian C. McDermott
Department of Physics, FM-15, University of Washington, Seattle, Washington 98195

(Received 4 September 1986; accepted for publication 17 November 1986)

Student understanding of the impulse-momentum and work-energy theorems was assessed by performance on tasks requiring the application of these relationships to the analysis of an actual motion. The participants in the study were undergraduates enrolled in either the honors section of a calculus-based introductory physics course or in the regular algebra-based course. The students were asked to compare the changes in momentum and kinetic energy of two frictionless dry-ice pucks as they moved rectilinearly under the influence of the same constant force. The results of the investigation revealed that most of the students were unable to relate the algebraic formalism learned in class to the simple motion that they observed.

I. INTRODUCTION

This paper reports the results of an investigation of student understanding of the concepts of impulse and work and the relationship of these concepts to changes in momentum and kinetic energy.[1] Research over the past several years has provided a substantial amount of detail on the difficulty students have in making the proper connection between force and motion. Considerably less information is available about their ability to relate force to more complex concepts.[2] The present study is part of the ongoing effort by the Physics Education Group at the University of Washington to identify specific difficulties encountered by students in various topics in physics and to use these findings as a guide in designing instruction.[3]

The aspect of understanding emphasized in the investigation is the ability to apply the impulse-momentum and work-energy theorems to the analysis of an actual motion. We wanted to determine if students who had studied the relevant concepts could make a correspondence between an observed motion and the algebraic formalism. As in much of our research, the method used is the individual demonstration interview. Because of its focus on real objects and events, we have found this technique to be particularly effective for examining the ability of students to make connections between the physical world and its algebraic and graphical representations.

A typical interview begins with a simple demonstration that serves as the basis for a set of tasks to be performed by the student. The tasks are accompanied by questions that have been structured to reveal the meaning the student ascribes to a particular concept or relation. The questions become part of a dialogue in which the investigator attempts to probe the student's thinking. In addition to those that are prescribed, the investigator may ask additional questions to clarify a reply or follow up on a comment. The student's actions and other nonverbal responses are noted. The entire discussion is audiotaped and transcribed. The transcripts, together with the investigator's notes from the interviews, are later analyzed in detail.

The 28 students who participated in the investigation were volunteers from two introductory physics courses at the University of Washington. Sixteen students were enrolled in the noncalculus physics course and 12 were in the honors section of calculus-based physics. For each group, the average of the final course grades of the participants in the study was somewhat higher than the average for the respective groups as a whole.

II. DESCRIPTION OF THE TASKS

In the tasks used in this investigation, students are asked to compare the changes in momentum and kinetic energy of two dry-ice pucks that move on a glass table, as shown in Fig. 1. The table is about 2 m long and 1 m wide. Two parallel lines, labeled (a) and (b), mark off a distance of about 30 cm on the table. The two pucks differ greatly in mass but are subjected to the same constant force. One puck is made of brass and has a mass of about 3500 g; the other, made of plastic and aluminum, has a mass of only about 350 g. The diameter of the base is about 15 cm for the brass puck and about 10 cm for the plastic one. The height of both pucks is about 15 cm and the diameter of both dry-ice container sections is about 7 cm.

During the demonstration, each puck is started from rest just behind line (a), moves rectilinearly under a constant force applied between lines (a) and (b), and then moves freely beyond line (b). Although the motion of each puck is observed separately, it is readily apparent that the brass puck traverses the distance between lines (a) and (b) much more slowly than the plastic one.

The force on the pucks is supplied by a steady stream of air blown through the hose of a reversed vacuum cleaner of the type used in air track experiments. The magnitude of the applied force can be varied by moving the end of the

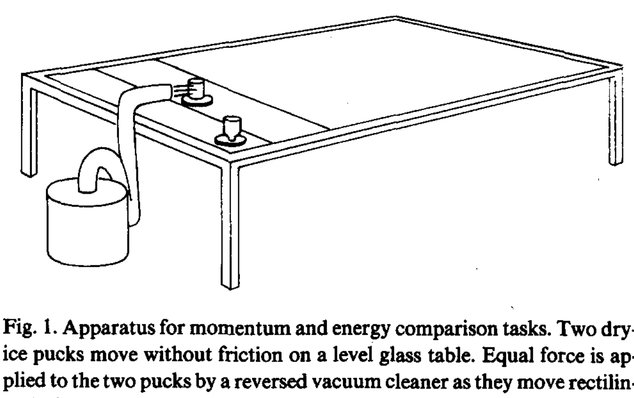


Fig. 1. Apparatus for momentum and energy comparison tasks. Two dry-ice pucks move without friction on a level glass table. Equal force is applied to the two pucks by a reversed vacuum cleaner as they move rectilinearly from line (a) to line (b).

hose toward or away from the puck. Attached to the opening are small strips of paper that, when blown out by the air, serve as spacers for maintaining a constant separation between the hose and the puck. By placing a hand in front of the opening, the students can feel that the force is constant as long as the distance from the hose remains unchanged.

Each of the two tasks presented to the students consists essentially of a single question. After observing the motion of both pucks, the student is asked whether the pucks have the same or different momentum during their free-motion after crossing line (b). Regardless of the response to this first question, the student is then asked if the two pucks have the same or different kinetic energy during their free-motion beyond line (b). If the student does not reply correctly to either or both questions, the investigator begins a dialogue that provides an increasing amount of help to the student in analyzing the motion.

To make a correct comparison of the changes in momentum (Δp) and kinetic energy (ΔT) for the two pucks, it is necessary to know how these quantities are related to the concepts of impulse and work. The honors students were familiar with the impulse-momentum and work-energy theorems in integral form. For motion in one dimension,

$$\int_i^f F_x \, dt = \Delta p_x = mv_f - mv_i$$

and

$$\int_i^f \mathbf{F} \cdot d\mathbf{x} = \Delta T = \frac{1}{2} mv_f^2 - \frac{1}{2} mv_i^2,$$

where $\mathbf{F}$ is the total force acting on the puck, m is the mass, and v_i and v_f are the initial and final velocities.

The noncalculus physics students had encountered these relationships in terms of $F \Delta t$ and $F \Delta x$, the forms to which they can be reduced when the force is constant: the situation in the demonstration.

Successful performance on the tasks requires only qualitative reasoning. Before the students are asked to compare the momenta and kinetic energies of the two pucks, they are led to assume (as is approximately the case) that the air stream exerts the same force on the two pucks. Since equal constant forces are applied to both pucks, the change in momentum is proportional to the time each takes to traverse the distance between the lines. Because of its greater mass, a smaller acceleration is imparted to the brass puck. During the longer time it spends between the lines, it receives a greater impulse. Hence the brass puck experiences a greater change in momentum than the plastic puck.

Comparing the kinetic energies requires fewer steps than comparing the momenta. Since the pucks move almost without friction and do not rotate, the total change in kinetic energy of the center of mass is equal to the work done on the puck. Since the same constant force is applied to each puck for the same distance, the change in kinetic energy is the same.

In order for a response to be considered correct, it was necessary for the student both to make the right comparison and to give the proper reasoning. Students who concluded that the brass puck had the larger momentum or that the kinetic energies were the same, but who did not give adequate justification, are not included in the category of students who made correct comparisons. If a student could not decide if the momenta or kinetic energies were the same or different, or if the student's justification of his comparison could not be deciphered or adequately clarified, the response was considered indeterminate. Students who gave indeterminate responses generally appeared more perplexed by the tasks than students who gave incorrect or inadequately justified responses.

Although the students who participated in the investigation had all completed the parts of introductory mechanics on momentum and energy, it was not really expected that many would be able to make a correct analysis on observing the demonstration for the first time. Therefore, as the interview progressed, the students were given an increasing amount of guidance toward noting the important features of the motion that were needed to make the comparison. If the initial response of a student on either task was incorrect or inadequately justified, the investigator would draw the student's attention to the way in which the pucks had been set in motion, i.e., the same constant force applied for the same distance. It was expected that once the students had taken note of these deliberately orchestrated conditions they would make use of the concepts of impulse and work in analyzing the motion. If, with this amount of prompting, a student still seemed at a loss about how to approach the task, the student was asked directly if he was familiar with the terms "impulse" and "work" and if the words represented ideas that could be applied to the demonstration at hand. If at this point the student was still unable to make a proper comparison of the momenta or kinetic energies of the pucks, the interview was terminated.

III. PERFORMANCE ON THE TASKS

The results of the momentum comparison task are summarized in Tables I and II and of the energy comparison task in Tables III and IV. The data for the noncalculus physics students are contained in Tables I and III and the data for the honors students are in Tables II and IV. The columns in the tables are labeled by the three levels of investigator intervention: (i) no intervention, (ii) attention drawn to the starting procedures, and (iii) explicit mention

Table I. Results of the momentum comparison task in the noncalculus physics group ($N = 16$).

Level of intervention	Correct comparison	Comparison incorrect or inadequately justified	Indeterminate
Initial comparison: No interviewer intervention	0	100%	0
Comparison after discussion of starting conditions	0	81%	19%
Comparison after explicit discussion	6%	25%	69%

Table II. Results of the momentum comparison task in the honors calculus physics group ($N = 12$).

Level of intervention	Correct comparison	Comparison incorrect or inadequately justified	Indeterminate
Initial comparison: No interviewer intervention	25%	75%	0
Comparison after discussion of starting conditions	58%	25%	17%
Comparison after explicit discussion	67%	25%	8%

of impulse or work. The columns are labeled by a description of the adequacy of the comparisons made by the students. Comparisons have been divided into three groups: (i) correct comparisons, (ii) incorrect or inadequately justified comparisons, and (iii) indeterminate comparisons.

As can be seen from the data, a greater number of the honors students solved the energy comparison task than solved the momentum comparison task. Furthermore, those who compared the energies correctly were able to do so with less help from the investigator than was necessary for successful completion of the momentum comparison task. With intervention by the instructor, almost all the honors students were eventually able to conclude that the two pucks had equal kinetic energy. However, only about two-thirds were able to use sound physical reasoning to decide that the brass puck had the larger momentum.

Almost none of the noncalculus physics students was able to apply the concepts of impulse or work to a comparison of either the momenta or kinetic energies of the two pucks. Intervention by the investigator did not seem to help these students as it had the honors students. In fact, there was an increase in the number of "indeterminate" responses on both tasks after investigator intervention.

IV. REASONING ON THE TASKS BEFORE INVESTIGATOR INTERVENTION

At least as important as the correctness of the comparisons made by the students are the explanations they gave in support of their responses. The reasoning used reveals a great deal about the nature of the difficulties.

There were four possible responses on both the momentum and energy comparisons tasks. Students could say that the momentum or kinetic energy of the brass puck was larger or smaller than the momentum or kinetic energy of the plastic puck, or that the momenta or kinetic energies of both pucks were equal. The fourth possible response was that one could not tell from the given information whether there was a difference in these quantities. All four responses were obtained on each task.

A. Momentum comparison task

The data in Table V show the percentage of students in each group who gave each type of response on the momentum comparison task before any investigator intervention. The initial response of most students in both the honors

Table III. Results of the kinetic energy comparison task in the noncalculus physics group ($N = 16$).

Level of intervention	Correct comparison	Comparison incorrect or inadequately justified	Indeterminate
Initial comparison: No interviewer intervention	0	88%	13%
Comparison after discussion of starting conditions	0	69%	31%
Comparison after explicit discussion	0	69%	31%

Table IV. Results of the kinetic energy comparison task in the honors calculus physics group ($N = 12$).

Level of intervention	Correct comparison	Comparison incorrect or inadequately justified	Indeterminate
Initial comparison: No interviewer intervention	50%	50%	0
Comparison after discussion of starting conditions	75%	17%	8%
Comparison after explicit discussion	83%	8%	8%

Table V. Student comparisons of the momenta of the ice pucks before investigator intervention (numbers are percent of each sample). P_B = momentum of brass puck, P_P = momentum of plastic puck.

	Honors calculus physics ($N = 12$)	Noncalculus physics ($N = 16$)
$P_B = P_P$	58%	50%
$P_B > P_P$	25%	19%
$P_B < P_P$	0	19%
Cannot tell	17%	12%

and the noncalculus physics courses was that the momenta of the two pucks were equal. One-fourth of the honors students responded correctly that the brass puck had the greater momentum. None concluded that the brass puck had a smaller momentum than the plastic puck. However, the same number of noncalculus physics students said that the momentum of the brass puck was greater as said that it was smaller.

In Table VI are the reasoning schemes commonly used by students in justifying their comparisons and the percentages of students who gave these different types of explanations before investigator intervention. When Tables V and VI are examined together, it can be seen that students who made the proper comparison often did not arrive at this conclusion by correct reasoning. As stated earlier, unless the comparisons and reasoning were both correct, the response was not considered correct. Although almost 20% of the noncalculus physics students stated that the momentum of the brass puck was greater than that of the plastic puck, none was able to justify this response with an argument involving impulse. The lack of a correct explanation, together with the fact that an equal number of students made the converse claim, suggests that the noncalculus physics students who chose correctly were simply guessing.

Of the students in both groups who concluded incorrectly that the momenta were the same, by far the most common justification given was what might be described as a "compensation argument." The following excerpt is illustrative.

(I, investigator; S, student.)

I: Do the brass puck and plastic puck have the same momentum or different?

S: I think they have the same [momentum]...because, well, momentum is mass times velocity...so the brass puck has more mass, [and] a slower velocity...I'm not sure if they are exactly equal, but with the same force they should be equal because the smaller puck has less mass [and] a higher velocity.

This student reveals the key element of his analysis when he describes momentum as "mass times velocity." He reasons that the larger mass of the brass puck is probably compensated for by its lower velocity when the quantity mv is compared. The student confirms his belief in this analysis by going through the complementary argument for the plastic puck, i.e., he points out that the plastic puck has less mass but a larger speed. He further supports his conclusion that the momenta should be the same by noting that the two pucks are each subject to the same force.

The quality of the argument illustrated in the last interview excerpt indicates that the student has the requisite mathematical capability to deal with the material. It is clear, however, that he has not been able to connect force and time with change in momentum in a way that could be useful in analyzing the motion he has observed. The student seems to think of momentum simply in terms of a definition rather than as a concept that can be applied to account for what happens in the physical world.

The vast majority of students who stated that the momenta were equal justified their response with some variation of the compensation argument. There was one student in each group, however, who gave an explanation based solely on the equality of the forces on the brass and plastic pucks.

Although initially only 25% of the honors students stated that the brass puck had the larger momentum, all who made this correct comparison reasoned correctly that it received a larger impulse. Furthermore, as can be seen in Table II, many students in the honors course who initially used a compensation argument, were with the help of the investigator, eventually able to respond correctly. In contrast, even with assistance, the noncalculus physics students were not very successful in using an impulse argument in comparing the momenta.

B. Energy comparison task

The data in Table VII show the percentage of students who gave each type of response on the energy comparison task before any discussion with the investigator. Table VIII lists the reasoning schemes identified and the corresponding percentages of students who used them before the investigator intervened. About one-half of the honors students answered correctly that the kinetic energies of the two pucks were equal. Slightly fewer than one-third of the

Table VI. Common reasoning schemes on the momentum comparison task before investigator intervention (numbers are percent of each sample). P_B = momentum of brass puck, P_P = momentum of plastic puck.

Comparison	Explanation	Honors calculus physics ($N = 12$)	Noncalculus physics ($N = 16$)
$P_B > P_P$	Larger impulse received[a]	25%	0
$P_B = P_P$	Compensation argument	50%	44%
$P_B = P_P$	Equal applied force	8%	6%
No specific comparison	Confused discussion	17%	50%

[a] Correct response.

Table VII. Student comparisons of the kinetic energies of the ice pucks before investigator intervention (numbers are percent of each sample). T_B = kinetic energy of brass puck, T_P = kinetic energy of plastic puck.

	Honors calculus physics ($N = 12$)	Noncalculus physics ($N = 16$)
$T_P = T_B$	50%	31%
$T_P < T_B$	0	25%
$T_P > T_B$	33%	38%
Cannot tell	17%	6%

noncalculus physics students gave this response. As can be seen from a comparison of Tables VII and VIII, all of the honors physics students who stated that the kinetic energies were equal recognized that the work done on both pucks was the same, but none of the noncalculus physics students who made this choice used this type of argument. Furthermore, as Table III indicates, none of the noncalculus physics students was able to give the proper explanation even after a considerable amount of prompting by the investigator. It is interesting to note, both on this task and on the momentum comparison task, that if explanations had not been required, a significantly different impression of student understanding would have resulted.

As was the case in comparing the momenta, there appeared to be a systematic tendency to make one particular incorrect comparison of the kinetic energies. The initial response of about one-third of the honors students, and more than one-third of the noncalculus physics students, was that the plastic puck had greater kinetic energy than the brass puck. The same type of compensation argument students used to conclude falsely that the momenta were equal appeared to underlie the incorrect conclusion that the plastic puck had a greater kinetic energy. The following interview excerpt illustrates the reasoning typically used.

I: Do the brass puck and the plastic puck have the same kinetic energy or different?

S: I think the smaller puck would have a larger kinetic energy...because kinetic energy is $mv^2/2$ and since the v is squared, the one with the larger velocity would probably have a larger kinetic energy.

This student clearly exhibits some ability to reason mathematically. The argument that speed is a more important variable than mass in comparing kinetic energies, since speed appears quadratically rather than linearly in the definition, is a relatively sophisticated type of analysis. Again, as in the case of the momentum comparison task, the essential physics of the problem has been missed. The student's reasoning is based soley on the definition of kinetic energy and lacks any reference to the way in which the work done on the puck is related to the change in kinetic energy.

The reasoning used in the compensation argument is mathematically nontrivial, if physically inadequate. Among the noncalculus physics students, there were other examples of incorrect reasoning that were much less sophisticated. The following excerpt is one such example taken from a discussion after the investigator had intervened.

I: What does that term kinetic energy mean? What do you think of when you hear the term kinetic energy?

S: I think of the formula...isn't it one-half mass times velocity squared?

I: Yes, that's the definition of kinetic energy. So what does that imply about which one had the greatest kinetic energy?

S: Actually if I reasoned that the momenta are the same, then I would have to say the kinetic energies are the same because the quantities involved are the same ones.

I: Can you be more specific? How are you thinking here?

S: Well, they both incorporate mass and velocity.

This student had earlier defined momentum as mv and here defines kinetic energy as $mv^2/2$. The conclusion that the momenta and kinetic energies are equal seems based only on the idea that they are both combinations of mass and velocity. This level of reasoning was common in the noncalculus physics group, but was not encountered at all among the honors physics students.

V. REASONING ON THE TASKS AFTER INVESTIGATOR INTERVENTION

Students who failed to make a correct comparison on one or both of the tasks often initially used compensation arguments to justify their responses. After the investigator focused their attention on the starting conditions of the pucks and explicitly mentioned the terms "work" and "impulse," specific difficulties with these concepts began to emerge. The two excerpts below come from discussions that took place relatively late in two separate interviews. Although both focus on the energy comparison task, they illustrate the type of questioning and student response that occurred after investigator intervention.

In the first interview, the student failed to analyze the motion correctly even after being reminded to consider the

Table VIII. Common reasoning schemes on the kinetic energy comparison task before investigator intervention (numbers are percent of each sample). T_B = kinetic energy of brass puck, T_P = kinetic energy of plastic puck.

Comparison	Explanation	Honors calculus physics ($N = 12$)	Noncalculus physics ($N = 16$)
$T_P = T_B$	Same work done[a]	50%	0
$T_P > T_B$	Compensation argument	33%	25%
$T_P = T_B$	Equal applied force	0%	19%
No specific comparison	Confused discussion	17%	56%

[a] Correct response.

starting conditions. He is now being asked explicitly about the term "work."

I: Have you ever heard the term work? Do you remember what that word means in physics?

S: Work was...the change in kinetic energy...or, um, let me think here...I think it might have been the force times...I'm not sure, I think I recall the formula R, F, the cosine of the angle between the two. But we just did problems on that and I can't remember exactly.

This student remembers, or at least is able to repeat, some key ideas. He is even able to state that "work was...the change in kinetic energy." The understanding of the concept of work, however, seems to be limited to repeating the elements of a formula. He is not able to connect the symbols with the features of the demonstration. There is no evidence that he understands kinetic energy at that level either. Although he states that work is equivalent to the change in kinetic energy, he does not seem to understand the relationship at a level that would allow him to apply his knowledge to the task at hand. This type of response was fairly typical of students from the noncalculus physics class.

The next excerpt demonstrates that even when there was some understanding of work and kinetic energy considered separately, it did not follow that a student understood the connection between these concepts. The transcript comes from the portion of the interview in which the student was asked specifically if she was familiar with the term "work." As in the interview quoted above, this question was asked after explicit reference to the starting conditions had failed to elicit a proper comparison of kinetic energies.

I: What ideas do you have about the term work?

S: Well, the definition that they give you is that it is the amount of force applied times the distance.

I: Okay. Is that related at all to what we've seen here? How would you apply that to what we've seen here?

S: Well, you do a certain amount of work on it for the distance between the two green lines: You are applying a force for that distance, and after that point it's going at a constant velocity with no forces acting on it.

I: Okay, so do we do the same amount of work on the two pucks or different?

S: We do the same amount.

I: Does that help us decide about the kinetic energy or the momentum?

S: Well, work equals the change in kinetic energy, so you are going from zero kinetic energy to a certain amount afterwards...so work is done on each one...but the velocities and the masses are different so they [the kinetic energies] are not necessarily the same.

In her first three statements this student indicates a satisfactory understanding of the concept of work. She associates the applied force and the distance over which it acts and points out that in the demonstration the same amount of work is done on each puck. In her last statement she mentions that the work done equals the change in kinetic energy. Nevertheless, she still is unable to conclude that the kinetic energies must be the same. In other words, although there appears to be a satisfactory understanding of work and an ability to state the work-energy theorem, the student seems to be distracted by the observed dissimilarity of the masses and speeds.

It should be noted that had the interview been terminated any earlier than it was, the impression would have been that the student's understanding was adequate. After all, almost everything said was correct. It was only by continuing to probe her responses that the investigator was able to determine that the student did not actually make the connection between the work-energy theorem and the moving pucks. Unlike a physicist, the student did not see the demonstration in terms of a direct application of the formula to the real world.

VI. CONCLUSION

Many of the students who participated in this study demonstrated by their performance in introductory physics that they were able to answer examination questions covering the relevant course material. The honors students, in particular, seemed to have little trouble in applying the concepts of impulse, momentum, work, and energy in the solution of some rather sophisticated problems. Yet, almost all of the students in the noncalculus physics course and many of the honors students experienced considerable difficulty in a straightforward application of the impulse-momentum and work-energy theorems to the actual one-dimensional motion of an object under a constant force.

A. Discussion of results

It is evident from the discussion of student performance during the interviews that success or failure on the impulse-momentum and work-energy tasks requires more than memorization of the relevant theorems and the definitions of the quantities that are involved. To be able to apply these relationships to real world situations requires knowledge at a deeper level.

In order to analyze the demonstration correctly, the student must understand the operational definitions for work and impulse. He must further recognize that the definition of kinetic energy and momentum as particular functions of v is not an arbitrary choice. The changes in these quantities are related in a very precise way to the integrals over time and distance of the net force applied on a body. It is because of their connection to impulse and work that the quantities mv and $\frac{1}{2}mv^2$ have special significance. Moreover, it is important that students understand that the impulse-momentum and work-energy theorems express physical laws that relate two different, precisely defined quantities. Moreover, in each case there is a cause and effect relationship. This subtle, but nevertheless, critical point seemed to elude many of the participants in the study. It was apparent during the interviews that students often thought of the symbol $=$ as representing simply a mathematical identity with no dynamic quality.[4] For example, they did not interpret the work-energy theorem as a statement that doing work on a body produces an increase in its kinetic energy.

To make a correct comparison of the observed motions, the raw intuition that pushing on an object makes it go faster is not sufficient. Both the brass puck and the plastic puck visibly increase in speed during the demonstration and changes in both kinetic energy and momentum result from this increase. It is, in fact, this similarity between momentum and kinetic energy for motion in one dimension that requires a comparison task to help make clear the distinction.

B. Implications for instruction

The results of this investigation showed that many of the participants in the study had failed to recognize the significance of the impulse-momentum and work-energy theorems. Most of the noncalculus physics students seemed to have emerged from their course unable to recognize the critical elements in either of these relations. (There are, of course, more subtle aspects of interpretation that this research did not address.[5]) There is no reason to believe that the instruction these students received was inferior to the usual presentation of this, or any other, topic in introductory physics.

It has been our experience that fundamentally important features of concepts that are not easily visualized will be missed if they are presented only verbally, whether by textbook or in lecture. The impulse-momentum and work-energy theorems are a case in point. The demonstrations used in the comparison tasks provide a particularly simple set of circumstances in which these relations can be readily applied. Yet most of the participants in the study failed to make the appropriate connection. Incorporated into instruction, either as demonstrations or as problems, tasks such as those discussed in this paper can help provide the practice needed in applying the impulse-momentum and work-energy theorems to real world events.

The presentation of both theorems may suffer from premature emphasis on their application to systems of objects for which momentum or kinetic energy is conserved. A deeper understanding of these relations may result if their application in single particle dynamics is stressed. Before encountering the conservation laws, students should be given many opportunities to use the impulse-momentum and work-energy theorems to find the kinematical and dynamical quantities that they have previously obtained through the application of Newton's laws and the kinematical equations. For example, the final speed or duration of motion for a single body under the influence of an external force is readily found from impulse and momentum considerations. Exposure to this alternative approach can help students recognize that the theorems have a meaning that is independent of their role in the derivation of the conservation laws.

To develop a functional understanding, most students need experience in interpreting the formal relationships of physics in a variety of different contexts and under different conditions. A deep conceptual understanding is unlikely to be achieved, however, if students passively observe the instructor perform a demonstration or work a problem. In our own teaching, we have found it necessary to engage students actively in making explicit the connections between the algebraic formalism and real world applications.

ACKNOWLEDGMENT

This research has been supported in part by National Science Foundation grant numbers SED81-12924 and DPE84-70081.

[a] Present address: Shell Oil Company, 200 North Dairy Ashford, Houston, TX 77079.

[1] R. A. Lawson, Ph.D. thesis, Department of Physics, University of Washington, 1984 (unpublished).

[2] For a summary of some of the research prior to 1984 on conceptual understanding in mechanics, see L. C. McDermott, Phys. Today, **37** (7), 24 (1984). Among papers that have appeared since then are P. W. Hewson, Am. J. Phys. **53**, 684 (1985); L. Viennot, Am. J. Phys. **53**, 432 (1985); I. A. Halloun and D. Hestenes, Am. J. Phys. **53**, 1043 (1985) and Am J. Phys. **53**, 1056 (1985); D. Maloney, J. Res. Sci. Teach. **22**, 261 (1985); J. Aguirre and G. Erickson, J. Res. Sci. Teach. **21**, 439 (1984).

[3] See, for example, D. E. Trowbridge and L. C. McDermott, Am. J. Phys. **48**, 1020, (1980) and Am. J. Phys. **49**, 242 (1981); F. M. Goldberg and L. C. McDermott, Am. J. Phys. **55**, 108 (1987).

[4] A discussion of difficulties in interpreting the equality sign appears in A. B. Arons, *Handbook of Introductory Physics Teaching* (unpublished), Chap. 3.

[5] For a discussion of complications involved in applying the work-energy theorem in situations that are less idealized than the demonstration used in this study, see B. A. Sherwood, Am. J. Phys. **51**, 597 (1983).

The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems

Tara O'Brien Pride, Stamatis Vokos, and Lillian C. McDermott
Department of Physics, University of Washington, Seattle, Washington

(Received 3 December 1996; accepted 2 July 1997)

The issue of how to assess learning is addressed in the context of an investigation of student understanding of the work-energy and impulse-momentum theorems. Evidence is presented that conceptual and reasoning difficulties with this material extend from the introductory to the graduate level and beyond. A description is given of the development of an instructional sequence designed to help students improve their ability to apply the theorems to real motions. Two types of assessment are compared. The results demonstrate that responses to multiple-choice questions often do not give an accurate indication of the level of understanding and that questions that require students to explain their reasoning are necessary. Implications for the preparation of teaching assistants are discussed.

I. INTRODUCTION

During the past two decades, there has been a steadily increasing amount of research on the learning and teaching of physics.[1] Investigations conducted among introductory physics students indicate that the difference between what is taught and what is learned is much greater than most instructors realize.[2] We can think of the role that research can play in helping to bridge this gap as having three interrelated components. (See Fig. 1.) The first consists of investigations of student understanding and includes most of the studies that have been conducted to date. A second component in which there has been considerable progress is the application of research findings in curriculum development.[3,4] Relatively little attention has been directed toward the third component,

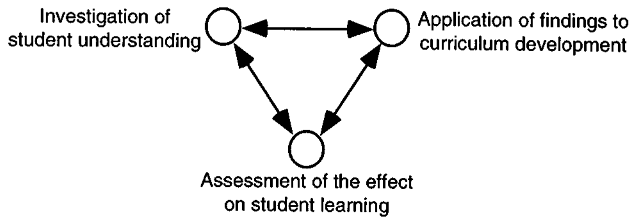


Fig. 1. The role of research in physics education.

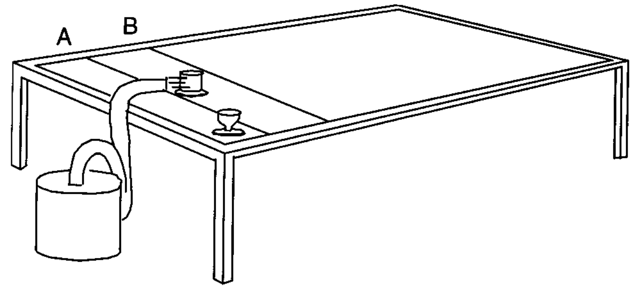


Fig. 2. Apparatus used in individual demonstration interviews on work-energy and impulse-momentum tasks. Students are asked to compare the final momenta and kinetic energies of two dry-ice pucks (one brass and one plastic) that move on a glass table. A constant force is applied by a steady stream of air in a direction perpendicular to the two parallel lines. Each puck starts from rest at line A and moves, without rotating and essentially without friction, to line B.

assessment of the effect on student learning. Efforts to develop innovative curriculum consistent with findings from research do not ensure that the end product will be effective. It is necessary to examine the intellectual impact on students and to ascertain in a rigorous manner whether the use of a particular curriculum or instructional strategy brings about a real gain in student understanding.

The means used to assess student learning should be consistent with the instructional goals. There are some general objectives for an introductory physics course that most instructors would agree are important. Having completed such a course, students should have acquired a sound understanding of some basic physical concepts and their mathematical representations and have developed the reasoning skills necessary to apply the concepts and representations of physics to the analysis and interpretation of simple phenomena. Students should also be able to make explicit the correspondence between a concept or representation and an actual object or event in the real world. Success in solving physics problems, the usual measure of effectiveness of instruction, does not necessarily indicate that other important goals have been achieved.[5]

In an earlier small-scale study, the Physics Education Group examined the ability of introductory students to apply the work-energy and impulse-momentum theorems to the analysis of actual motions.[6] This paper describes how we have extended the scope of the research to include the development and assessment of a tutorial to address some of the difficulties identified.[7,8] The scale has been greatly expanded through the participation of many more students, ranging from the introductory to the graduate level. We compare findings from our in-depth examination of student learning with results obtained from the administration of a broad assessment instrument, for which there are nationally reported scores. Our analysis of the large discrepancy that we found has implications that extend beyond a particular topic in mechanics. Viewed from a more global perspective, this paper addresses the issue of how the effectiveness of instruction can be meaningfully assessed.

II. INVESTIGATION OF STUDENT UNDERSTANDING

The important features of the tasks that we used to probe student understanding of the work-energy and impulse-momentum theorems are outlined below. A detailed description can be found in Ref. 9.

A. Student performance on the interview task

In the tasks used in the interviews, students are asked to compare the final kinetic energies and momenta of two dry-ice pucks (one brass and one plastic) that move on a glass table. (See Fig. 2.) A constant force (F) is applied by a steady stream of air in a direction perpendicular to the two parallel lines. Each puck starts from rest at line A and moves in a straight line, without rotating and essentially without friction, to line B.

A correct explanation was necessary for a response to be considered correct. The comparisons can be made by direct application of the work-energy and impulse-momentum theorems. Since the force is constant and parallel to the displacement (Δx), these reduce to

$$F\Delta x=\Delta K, \quad F\Delta t=\Delta p.$$

The change in kinetic energy (ΔK) equals the work done by the external force and is the same for both pucks. Since the same constant force is applied to both pucks, the magnitude of the change in momentum (Δp) is proportional to the time (Δt) each takes to traverse the distance between the lines. Because of its greater mass, a smaller acceleration is imparted to the brass puck. During the longer time it spends between the lines, it receives a greater impulse and hence experiences a greater change in momentum than the plastic puck. A correct comparison of the final momenta of the pucks also follows from the equality of the kinetic energies and the algebraic relationship between kinetic energy and momentum.

1. Individual demonstration interviews

In the initial research, the comparison tasks were administered during individual demonstration interviews. The 28 students who participated were volunteers from two introductory physics courses at the University of Washington (UW). There were 16 participants from the algebra-based course and 12 from the honors section of calculus-based physics. The average of their final grades was higher than the average for the classes in which they were enrolled.

Although the students had all completed the study of energy and momentum, it was not expected that many would be able to make a correct analysis on observing the demonstration for the first time. Therefore, as the interview progressed, they were given an increasing amount of guidance. When students could not make a proper comparison on their own, the investigator attempted to guide them through questioning. An example of the type of intervention that took place is given in the following excerpt from an interview transcript. [I, investigator; S, student]

Table I. Student performance on interview tasks and on written questions based on these tasks. Students were asked to compare the kinetic energy and momentum of two pucks of different mass acted upon by equal forces for the same distance. The first two columns indicate student responses during the interviews both initially and after intervention by the investigator. The third column shows results on a written test based on the interview questions.

		Results from interviews and written test *Correct explanation required for answer to be considered correct*		
		Interviews		Written test
		Students in calculus-based honors physics ($N=12$)	Students in algebra-based physics ($N=16$)	Students in calculus-based physics ($N=985$)
Correct kinetic energy comparison	***before* intervention**	**50%**	**0%**	**15%**
	after intervention	85%	0%	...
Correct momentum comparison	***before* intervention**	**25%**	**0%**	**5%**
	after intervention	65%	5%	...

I: What ideas do you have about the term work?
S: Well, the definition that they give you is that it is the amount of force applied times the distance.
I: Okay. Is that related at all to what we've seen here? How would you apply that to what we've seen here?
S: Well, you do a certain amount of work on it for the distance between the two green lines. You are applying a force for that distance, and after that point it's going at a constant velocity with no forces acting on it.
I: Okay, so do we do the same amount of work on the two pucks or different?
S: We do the same amount.
I: Does that help us decide about the kinetic energy or the momentum?
S: Well, work equals the change in kinetic energy, so you are going from zero kinetic energy to a certain amount afterwards...so work is done on each one...but the velocities and masses are different so they [the kinetic energies] are not necessarily the same.

The interview excerpt above demonstrates that, even if correct, short responses do not necessarily indicate understanding. Probing in depth is necessary for an accurate assessment. Had the questioning been terminated earlier, it would have seemed as if the student understood the relationship between the work done and the change in kinetic energy. It was only by continuing to probe that the investigator was able to determine that the student did not really connect the actual motion of the pucks with the work-energy theorem.

The data in the first two columns of Table I include responses before and after intervention by the investigator. Before intervention, only 50% of the honors students made a correct kinetic-energy comparison and only 25% made a correct momentum comparison. None of the other students made a correct comparison. While there was a marked improvement among the honors students as the interview progressed, the students in the algebra-based course, even with help, were never able to connect the algebraic formalism to the physical situation.

2. Written tests

After the results described above were published in the *American Journal of Physics*, we presented the same comparison tasks in written form to almost 1000 students in 11 regular and honors sections of the calculus-based physics course. The demonstration was shown. To be sure that they made the proper observations, the students were first asked to compare the accelerations and the masses of the pucks. Comparisons of the kinetic energies and the momenta were considered correct only if supported by correct reasoning in words or by equations.

The students were enrolled in sections taught by different instructors in several academic quarters. Lecture instruction on the work-energy theorem had been completed and homework had been assigned. Momentum and impulse had been presented in some but not all of the classes. When these concepts had not yet been covered, the students were told that the momentum of an object is equal to the product of its mass and its velocity.

The third column of Table I shows that the success rate was 15% on the kinetic energy comparison and 5% on the momentum comparison. The outcome was essentially the same whether or not this material had been covered in lecture. Therefore, we have not separated the data shown in Table I into groups. On the kinetic energy comparison task there were small variations among the sections but on the momentum comparison there were virtually none. Almost all students who responded correctly referred to both theorems. Very few used the equality of the kinetic energies and the mathematical relationship between the variables to compare the momenta.[10] The order in which the tasks were presented did not affect the results.

3. Incorrect reasoning used by students

Analysis of the written responses revealed reasoning difficulties similar to those identified during the interviews. Most students did not seem to recognize the cause–effect relationships inherent in the work-energy and impulse-momentum theorems. They did not relate the result of a force acting over a distance or time interval to a change in kinetic energy or momentum. Instead, they seemed to treat the theorems as mathematical identities.

Compensation reasoning was common. For example, students might claim that the momenta were equal because the greater velocity of the lighter puck compensated for its smaller mass. They might also say that the kinetic energy of the lighter puck was greater than that of the heavier puck because kinetic energy depends more on velocity, since it is squared, than on mass. In both of these examples, an incorrect comparison was made. However, faulty reasoning did not always lead to an incorrect comparison. For example, students sometimes argued that the kinetic energies were the same because energy is conserved or because the same force was applied to both pucks (without reference to the displacement). For the kinetic energy comparison, such incorrect reasoning leads to the right answer in this situation.

B. Need for special instruction

The poor performance on the comparison tasks suggested the need for special instruction on the application of the two theorems. The response of the Physics Education Group in such situations is to develop tutorials that address specific conceptual and reasoning difficulties. *Tutorials in Introductory Physics* is intended to supplement, not replace, the lectures and textbooks through which physics is traditionally taught.[11]

The development of the tutorials has been guided by research. The instructional approach is consistent with the following generalization: *Teaching by telling is an ineffective mode of instruction for most students*. The tutorials are expressly designed to engage students in active learning.[12] The emphasis is on the development of concepts and reasoning skills, not on quantitative problem-solving. The tutorial system consists of the following integrated components: pretests, worksheets, homework assignments, course examinations, and a weekly graduate teaching seminar that is required for all tutorial instructors.

The tutorial sequence begins with a pretest that is given in the large lecture section at the beginning of each week. Pretests are usually on material already covered in lecture but not yet in tutorial. They inform the instructors about the level of student understanding and help the students identify what they are expected to learn in the next tutorial. During the tutorial sessions, 20–24 students work together in groups of three or four. The worksheets, which provide the structure for these sessions, consist of carefully structured tasks that guide students through the reasoning needed to develop a sound qualitative understanding of important concepts. The instructors do not lecture but ask questions designed to help students find their own answers. The tutorial homework extends and reinforces what students have learned during the tutorial sessions. Questions based on the tutorials are included on all course examinations.

III. DEVELOPMENT AND ASSESSMENT OF A RESEARCH-BASED TUTORIAL

In this section, we describe the development and assessment of a tutorial on the work-energy and impulse-momentum theorems. The goal of the tutorial sequence is to help students learn to apply the theorems in specific situations, to reflect on the relationships involved, and to begin the process of generalization.

A. Description of the tutorial

We have often found that a good research probe can be transformed into an effective instructional procedure. The improvement among the honors students that occurred during the interviews suggested a basic design for the tutorial entitled *Changes in Energy and Momentum*. Comparison tasks provide the basis for carefully structured questions that guide students through the reasoning involved in the interpretation and application of the theorems.

The tutorial incorporates an instructional strategy often used by our group. It may be summarized as a series of steps: *elicit*, *confront*, and *resolve*.[13] The written test discussed earlier is used as a pretest to elicit the conceptual and reasoning difficulties that have been described. The tutorial worksheet is designed to address these and other difficulties that have been identified through research. There are two parts to the worksheet. In Part I, students confront and resolve the specific difficulties that they encountered in the physical situation presented on the pretest. In particular, Part I helps students relate the two theorems to real motions. In Part II, this process is continued as students apply the theorems in a more complicated context. The second part of the tutorial also helps to sharpen the distinction between work and kinetic energy as scalar quantities, and impulse and momentum as vectors.

In Part I of the worksheet, the students are guided in making a connection between the motion presented on the pretest and its algebraic representation. At this point, the students who answered incorrectly on the pretest recognize the conflict with their earlier response. They are guided through the reasoning that is needed to compare the final momenta and kinetic energies. They are asked to consider fictionalized dialogues in which compensation arguments are used. As they analyze the dialogues, they begin to see that such reasoning is inappropriate.

Difficulties of a serious nature cannot be successfully addressed in a single encounter.[14] Multiple challenges in different contexts are necessary so that students can have additional opportunities to *apply*, *reflect*, and *generalize*. Part II helps them deepen their understanding by applying the theorems in a situation in which more than one dimension is involved. The students use the apparatus in Fig. 3 to examine the motion of a ball that is released from the same height on a starting wedge under two different conditions. In the first case, the ball arrives at the top of the ramp with a velocity perpendicular to the boundary. In the second case, the ball arrives at the ramp with the same speed but at an acute angle with the boundary.

The tutorial worksheet guides the students through the steps in reasoning summarized in Fig. 4. They recognize that, when the ball is on the ramp, the direction of the net

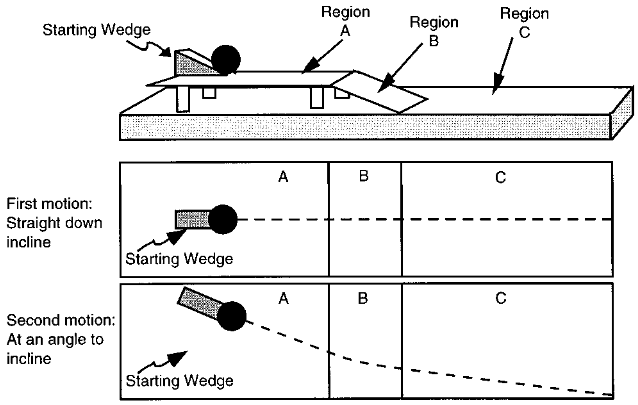


Fig. 3. Apparatus used in the tutorial entitled *Changes in Energy and Momentum* that helps students learn to apply the work-energy and impulse-momentum theorems. The ball is released from the same height on the starting wedge in the two cases.

force (and hence the direction of the change in momentum) is straight down the incline in both cases. From work-energy considerations, the students determine that the final speed of the ball is the same in the two motions. They construct the change-in-momentum vector in the two cases and find that its magnitude is greater in the case in which the initial and final momentum vectors are not collinear. From the impulse-momentum theorem, the students realize that the magnitude of the impulse is greater when the ball enters the ramp at an acute angle. They infer that the ball spends a longer time on the ramp in that case and conclude that this is consistent with kinematical considerations.

Consideration of motion in more than one dimension helps students deepen their understanding of the simpler one-dimensional case. Following the tutorial session, a tutorial homework assignment gives the students additional practice in applying the two theorems and in interpreting the causal relationships involved.

Comparison of Accelerations and Net Forces in Motions 1 and 2

From observed motions, Newton's 2nd Law, and free-body diagrams: $\vec{a}_1 = \vec{a}_2 = \vec{a}$ (down ramp)

$\vec{F}_{net_1} = \vec{F}_{net_2} = \vec{F}$ (down ramp)

Comparison of Changes in Kinetic Energy

Work-energy theorem: $W = \vec{F}\cdot\Delta\vec{x} = \Delta K$

On ramp: $\vec{F}\cdot\Delta\vec{x}_1 = \vec{F}\cdot\Delta\vec{x}_2$

$\Rightarrow$ $\boxed{\Delta K_1 = \Delta K_2}$

At top of ramp: $v_{1i} = v_{2i}$

$\Rightarrow$ $p_{1i} = p_{2i}$

At bottom of ramp: $v_{1f} = v_{2f}$

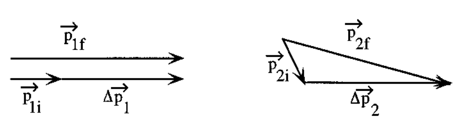
$\Rightarrow$ $p_{1f} = p_{2f}$

Comparison of Changes in Momentum

Impulse-momentum theorem: $\vec{J} = \vec{F}\Delta t = \Delta\vec{p}$

On ramp: $\Rightarrow$ $\vec{J}_1 \parallel \vec{J}_2$

$\Rightarrow$ $\Delta\vec{p}_1 \parallel \Delta\vec{p}_2$



Vector diagrams show: $\boxed{|\Delta\vec{p}_2| > |\Delta\vec{p}_1|}$

Hence, $\Delta t_2 > \Delta t_1$

Fig. 4. Summary of the reasoning in which students engage during Part II of the tutorial. The subscripts "1" and "2" refer to motion straight down the ramp and motion at an angle to the ramp, respectively. The tutorial is designed to help students learn to apply the work-energy and impulse-momentum theorems to real motions. The tutorial also helps students sharpen the distinction between work and energy as scalar quantities and impulse and momentum as vectors.

B. Comparison of pre-tutorial and post-tutorial student performance

In designing questions to assess functional understanding of a concept or principle, it is necessary to determine how different the testing context should be from that in which the ideas were introduced. The degree of transfer that it is reasonable to expect varies with the difficulty of the topic and the academic level of the students. We decided that the students would be sufficiently challenged if we based the post-test on the same physical setup as the pretest (see Fig. 2) but imposed a different condition on the motion. The students were asked to compare the final momenta and kinetic energies when the force was applied for the same time, rather than for the same distance (as on the pretest). They were expected to recognize that since both carts started from rest, at the end of the time interval the momenta would be the same. However, since the lighter puck would traverse a greater distance in the same time, more work would be done on it by the force. Hence, its kinetic energy would be greater.

The post-test question was given to 435 students on midterm or final examinations in three academic quarters. In grading the question, we paid careful attention to the explanations given by the students. In the first column of Table II are the pretest results reproduced from Table I. The pretest performance for the students who took this post-test was the same as for all 985 students for whom pretest data are given. The success rates on this post-test are shown in the second column of Table II. (The heading refers to Post-test #1 because a second post-test was developed later.) As can be seen, performance on the post-test was much better than on the pretest.[15] A correct kinetic energy comparison was given by 35% of the students and a correct momentum comparison by 50%.

To investigate whether students could apply the theorems in a more complicated physical situation, we gave a second version of the post-test on a midterm examination. Post-test #2 was specifically designed so that compensation reasoning would not yield the right answer. (The 320 students who took Post-test #2 had not taken Post-test #1.) For Post-test #2, Cart A and Cart B are at rest on parallel frictionless tracks that terminate in a common finish line. Cart A is behind Cart B. The students are told that Cart A has a greater mass and that a constant force is applied to Cart A. As Cart A passes Cart B, an equal constant force is exerted on Cart B. Both carts reach the finish line simultaneously, at which time Cart B is moving faster than Cart A. The students are asked to compare the final momenta and kinetic energies of the two carts. In this case, neither the final kinetic energies nor the momenta are equal. Since the force is applied to Cart A for a greater distance and for a longer time, Cart A expe-

Table II. Student performance on UW pretest and post-tests. The tests ask for a comparison of the kinetic energy and momentum of two objects of different mass acted upon by equal forces. On the pretest, equal forces act for the same distance. On Post-test #1, they act for the same time; on Post-test #2, the forces act over unequal distances for unequal time intervals.

	Results from UW pretest and post-tests *Correct explanation required for answer to be considered correct*		
	Students in calculus-based course		
	Pretest after lecture but before tutorial (Δx constant) ($N=985$)[a]	Post-test #1 after lecture and after tutorial (Δt constant) ($N=435$)	Post-test #2 after lecture and after tutorial ($\Delta x, \Delta t \neq$ constant) ($N=320$)
Correct kinetic energy comparison	15%	35%	30%
Correct momentum comparison	5%	50%	45%

[a]The column is repeated from Table I for easy reference. In this case, the written test is regarded as a pretest for the tutorial.

riences a greater change in both kinetic energy and momentum. Since both carts are initially at rest, Cart A has a greater final kinetic energy and momentum.

A comparison between the second and third columns of Table II shows that students who took Post-test #2 did almost as well as those who took Post-test #1. Therefore, the two post-tests may be considered roughly equivalent as a measure of conceptual understanding of the two theorems. For each post-test, the results were similar in different lecture sections, varying little from one lecturer to another. This finding is consistent with our experience in other cases. The effectiveness of the tutorial system does not seem to depend as much as some methods on the lecturing skills of individual instructors.

A comparison of pretest and post-test performance indicates that there was a significant improvement in the ability of students to apply both theorems after they had worked through the tutorial. It is clear, however, that students still had considerable difficulty, especially on the work-energy comparison task. There are two plausible reasons for the difference in gain between the two tasks. The greater success rate on the momentum comparison task could have been due to the greater emphasis on the impulse-momentum theorem in the tutorial. There is also an alternative explanation, however, that could account for the disparity in performance on the two tasks. Both post-tests explicitly call attention either to the equality or to the inequality of the time intervals during which the force acts on each cart. We found that many students used $F=ma$ and the definition of acceleration, $a=\Delta v/\Delta t$, to make the momentum comparison. A few students used the relationship $F=\Delta p/\Delta t$. In either case, comparison of the momenta may have been a relatively simple task for some students because they had gone through the reasoning involved in the derivation during the tutorial. However, the failure of most students to refer to the impulse-momentum theorem on the post-tests suggests that they had failed to recognize its generality. They had not developed a functional understanding of the concept that a force acting on an object for an interval of time causes a change in momentum of the object. Instead, they rederived for a specific situation the relationship expressed by the impulse-momentum theorem. In contrast, we found that students did not rederive the work-energy theorem to compare the final kinetic energies.

C. Results from other institutions

We believe that assessment of the effectiveness of instructional materials at institutions other than the one in which they were developed is crucial for the development of effective curriculum. The tutorials are being pilot-tested at other universities and at two- and four-year colleges. *Changes in Energy and Momentum* has been pilot-tested at several sites, including another large research university, where it has been used in a calculus-based course for science and engineering majors, and at a smaller research university in a course for physics majors.

At the large university, the pretest was administered after lecture instruction to about 270 students in three sections of the course during two academic semesters. The success rate was 10% on the kinetic energy comparison task and 5% on the momentum comparison task, results that are very similar to those obtained at the University of Washington. A third version of the post-test, which was constructed at the test site, was given. In this post-test, unequal forces acted on carts of different mass for the same distance in the same time interval. The students were told that the larger force acted on the larger mass. About 70% gave a correct response for the kinetic energy comparison and 75% for the momentum comparison. Analysis of the responses revealed that many students recognized that the final velocities were equal and therefore concluded that the more massive cart had the greater kinetic energy and momentum. Thus a correct comparison could be made quickly without reference to either theorem. Only 20% of the total number of students used the work-energy theorem and about 30% used the impulse-momentum theorem to arrive at correct comparisons. We do not know how many students would have referred to the

The diagram depicts two pucks on a frictionless table. Puck II is four times as massive as puck I. Starting from rest, the pucks are pushed across the table by two **equal** forces.

Question 20: Which puck will have the greater kinetic energy upon reaching the finish line?
(A) I (B) II
(C) They both have the same amount.
(D) Too little information to answer.

Question 21: Which puck will reach the finish line first?
(A) I (B) II
(C) They will both reach the finish line at the same time.
(D) Too little information to answer.

Question 22: Which puck will have the greater momentum upon reaching the finish line?
(A) I (B) II
(C) They both have the same momentum.
(D) Too little information to answer.

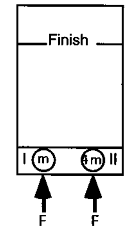


Fig. 5. Questions from the Mechanics Baseline Test (MBT) on the work-energy and impulse-momentum theorems.

theorems if they had not recognized that the velocities were equal. Therefore, we cannot determine, on the basis of this post-test, whether the tutorial was as effective with these students as with our own.

At the smaller university, the pretest was administered after interactive lecture instruction on the work-energy theorem but before instruction on momentum. The success rate for the 34 students who took the pretest was 25% on the kinetic energy comparison and 5% on the momentum comparison. Post-test #2 was included on the final examination. Students did much better on the post-test than on the pretest. About 45% gave a correct response to the kinetic energy comparison task and 50% to the momentum comparison task. The gain in student performance was similar to that at the University of Washington for the same post-test.

The pretest and post-test results discussed above indicate that application of the theorems to actual motions is difficult for students. These difficulties are not readily overcome, but it appears that the types of instructional strategies incorporated into the tutorial *Changes in Energy and Momentum* can be effectively employed by instructors in different institutional settings.

IV. COMPARISON OF IN-DEPTH AND BROAD ASSESSMENTS OF LEARNING

Several multiple-choice instruments designed to assess student understanding in mechanics have been widely disseminated during the past several years.[16–20] The most widely administered and thoroughly tested is the Force Concept Inventory (FCI). The results have increased faculty awareness of the failure of many students to distinguish between Newtonian concepts and erroneous ''common sense'' beliefs, both before and after instruction in physics.[21] Interpretation of the results from the FCI has been the subject of lively debate.[22] The Mechanics Baseline Test (MBT) covers a greater range of topics than the FCI. It is intended for use after instruction. Several questions on the MBT are taken from the research literature, including two derived from the kinetic energy and momentum comparison tasks. Below we compare the results on these questions from our in-depth assessment and from the MBT.

A. MBT version of UW pretest questions

The March 1992 issue of *The Physics Teacher* contained results from the administration of the MBT to eight groups of students at several high schools and universities.[23] Three groups (about one-third of the students) were at universities; the others were from high schools. Questions 20, 21, and 22 on the MBT are shown in Fig. 5. The students are told that two pucks start from rest and are pushed from the same starting line to the same finish line by two equal forces. Question 20 is based on the work-energy task and Question 22 on the impulse-momentum task. Question 21 explicitly calls attention to the time intervals. The format is different from that of the UW pretest in that the MBT is multiple-choice, no demonstration is shown, and the mass of one object is explicitly given as four times that of the other.

The nationally reported results for the eight groups ranged from about 10% to 70% correct for Question 20 on the work-energy task and from about 30% to 70% correct for Question 22 on the impulse-momentum task. The corresponding average scores for the total number of students in the eight groups was 30% for Question 20 and 50% for Question 22. On average, the success rate was higher for the momentum task than for the kinetic-energy task, as on each of the UW post-tests. However, closer inspection of the nationally reported MBT results reveals that this was not the case in all classes. As discussed earlier, we believe that we can account for this difference in the case of the UW students. The wide variation from class to class and from institution to institution in the nationally reported MBT results makes it difficult to draw any general conclusions.

The nationally reported MBT results indicated significantly better student performance than we had found on the UW pretest on essentially the same questions. (See the third column in Table III.) As mentioned earlier, the MBT is intended for post-instruction evaluation. The UW pretest was given after energy and momentum had been introduced but before instruction on these topics had been completed. We did not know details about the instruction the students had received before administration of the MBT. We did know that some of the students had prior experience with the physical setup, whereas the UW students had none.

B. UW student performance on MBT version of pretest questions

We wanted to investigate whether the difference in the format of the questions on the UW pretest and on the MBT could account for the discrepancy in performance. We therefore gave the MBT version of the comparison tasks (Questions 20, 21, and 22) as a pretest to about 400 students after the relevant lectures but before the tutorial. Only 10% answered correctly for the kinetic energy and only 5% for the momentum. These results, which are shown in the first column of Table III, were essentially the same as those obtained on the UW pretest that preceded the tutorial. Therefore, the difference in format between the two tests could not be responsible for the large discrepancy in the results.

C. The right answers for the wrong reasons

The most important difference in the way the MBT version of the comparison tasks was administered at the University of Washington and at other institutions was that we required students to explain their reasoning in addition to choosing an answer. We found that many UW students used the same types of incorrect reasoning on the MBT version as on the UW pretest. For example, some students supported

Table III. Student performance on MBT version of UW pretest. The pretest asks for a comparison of the kinetic energy and momentum of two pucks of different mass acted upon by equal forces for the same distance. The first column shows student performance when a correct explanation was required. The second column shows the results when explanations were ignored. Although the published version of the MBT does not ask for explanations, we required the students to support their answers. The third column shows the nationally reported results on the corresponding questions from the MBT.

	Results from MBT version of UW pretest *Are correct answers without explanations an adequate measure of student understanding?*		
	Students in calculus-based course at UW after lecture instruction		Nationally reported MBT results after instruction
	Correct answer, correct explanations ($N=400$)	Correct answer, explanations ignored ($N=400$)	Correct answer, no explanations ($N=1100$)
Correct kinetic energy comparison	10%	25%	30%[a] (10%–70%)[b]
Correct momentum comparison	5%	30%	50%[a] (30%–70%)[b]

[a]Average scores for the total number of students from eight groups.
[b]Range of the average scores of eight groups.

their correct answer that the kinetic energies are equal by saying that energy is conserved. Although the presence or absence of a demonstration did not affect overall success rates, there was sometimes an effect on the nature of the errors. An example of incorrect reasoning that occurred only on the MBT version was the claim that the heavier puck had a greater momentum because the final velocities of the pucks were equal. This wrong assumption was not made when there was a demonstration.

1. Reassessment of student performance on MBT version of pretest

We suspected that the only significant difference between the UW and MBT versions of the questions was that we required a correct explanation for an answer to be considered correct, whereas the MBT questions were multiple-choice. We decided to re-evaluate the responses of our students on the MBT questions in the same way that the grading had been done at the other institutions, i.e., without regard for the reasons given for the answers. This reassessment yielded a success rate of 25% on the work-energy task and 30% on the impulse-momentum task (see the second column in Table III). Thus, when no explanations were required, we found that our students had scores on these questions within the range of those reported nationally.[24]

2. Reassessment of student performance on UW post-tests

The disparity in student performance on the MBT version of the UW pretest, when correct reasoning was and was not taken into account, suggested that the same situation might prevail for the post-test. Therefore, we decided to reassess the 435 examination responses on Post-test #1, ignoring the reasons that students gave for their answers. As the second column of Table IV shows, the same students had an apparent success rate of 65% for the kinetic energy comparison and 80% for the momentum comparison. Table IV also includes the results from the second column of Table II when reasoning is taken into account. A quick inspection reveals a marked contrast between the two cases. When correct explanations are not required, the results are consistent with the best of the published results obtained with the MBT.[25]

The results for Post-test #2 are similar. The fourth column of Table IV shows the performance on Post-test #2 when credit is given for correct comparisons without regard to reasoning. Correct comparisons were made by 45% of the students on kinetic energy and by 55% on momentum. The results for the same students when reasoning is taken into account are repeated from the third column of Table II. As with Post-test #1, when correct explanations are not required for an answer to be considered correct, the success rate is considerably higher.

V. EFFECT OF ADVANCED STUDY

The effectiveness of the tutorials is heavily dependent on the tutorial instructors. They must have a deep understanding of the material, a knowledge of the intellectual level of the students, and skill in asking appropriate questions that can guide students through the necessary reasoning. The instructional staff of the tutorials is composed primarily of graduate teaching assistants (TA's) but also includes undergraduate physics majors, volunteers who are post-doctoral research associates, and junior faculty in the physics department.

Ongoing participation in a weekly graduate teaching seminar is required for all tutorial instructors. At the beginning of each seminar, the participants take the same pretest as the introductory students. They then examine the pretests taken earlier by the students and try to identify common errors. The participants spend most of the time in working collaboratively step-by-step through the worksheets, just as the students will do later in the week. Experienced tutorial instructors show by example how to conduct the tutorial sessions and how to address the conceptual and reasoning difficulties that are likely to arise. Over a period of several academic quarters, we gave the pretest on the work-energy and impulse-momentum theorems to the participants in the graduate teaching seminar. The results from the 74 seminar

Table IV. Student performance on UW post-tests. The post-tests ask for a comparison of the kinetic energy and momentum of two objects of different mass acted upon by equal forces. On Post-test #1, they act for the same time. On Post-test #2, the forces act over unequal distances for unequal time intervals. The first and third columns show student performance when a correct explanation was required for an answer to be considered correct. The second and fourth columns show results when explanations were ignored.

	Results from UW post-tests re-examined *Are correct answers without explanations on adequate measure of student understanding?*			
	Students in calculus-based course after both lecture and tutorial			
	UW Post-test #1		UW Post-test #2	
	Correct answer, correct explanations (Δt constant) ($N=435$)[a]	Correct answer, explanations ignored (Δt constant) ($N=435$)	Correct answer, correct explanations (Δx, $\Delta t \neq$ constant) ($N=320$)[a]	Correct answer, explanations ignored (Δx, $\Delta t \neq$ constant) ($N=320$)
Correct kinetic energy comparison	35%	65%	30%	45%
Correct momentum comparison	50%	80%	45%	55%

[a]The column is repeated from Table II for easy reference.

participants who were in their first year as tutorial instructors are shown in the first column of Table V. Correct comparisons and explanations were given by 65% for the work-energy task and by 70% for the impulse-momentum task. These results indicate that difficulties in applying the work-energy and impulse-momentum theorems extend to the graduate level. In this topic and in others, we find that advanced study does not necessarily deepen understanding of introductory physics.[26]

There were two noticeable differences between how the tutorial instructors (graduate students and post-docs) and the introductory students approached the comparison tasks. The instructors were much more likely to refer to the theorems than were the students, most of whom did not seem to appreciate the significance of the general principles. The instructors relied more on mathematics. Having arrived at an answer for the comparison of the kinetic energies or the momenta, they frequently used mathematics to make the other comparison.

Similar pretests were given in two national workshops to 137 physics faculty from other colleges and universities. No demonstration was shown, however. The second column of

Table V. Performance of graduate students, volunteer post-docs, and physics faculty on UW pretest and on Post-test #2. The first column shows the results when the pretest was given in the weekly graduate teaching seminar. The graduate students and volunteer post-docs had not yet worked through the tutorial on the work-energy and impulse-momentum theorems. The second column shows the pretest results obtained in two national workshops for physics faculty. The third column shows the results from Post-test #2 after the participants in the graduate teaching seminar had worked through the tutorial and served as instructors in the turorial sessions. Only the results from seminar participants in their first year as tutorial instructors are shown.

	Results from UW pretest and Post-test #2 *Correct explanation required for answer to be considered correct*		
	UW graduate teaching seminar (pretest)	National workshops for physics faculty (pretest)	UW graduate teaching seminar (post-test)
	Pretest before tutorial ($\Delta x=$constant) ($N=74$)	Pretest before tutorial ($\Delta x=$constant) ($N=137$)	Post-test #2 after tutorial (Δx, $\Delta t \neq$ constant) ($N=20$)
Correct kinetic energy comparison	65%	65%	95%[a]
Correct momentum comparison	70%	60%	95%[a]

[a]All graduate teaching seminar participants gave correct comparisons but 1 out of the 20 did not provide explanations.

Table V shows that the average success rate in both workshops taken together was 65% on the kinetic energy comparison task and 60% on the momentum comparison task. The pretests given in the two faculty workshops differed in the order in which the questions were presented. The success rate for the 47 faculty in the first group, who took the same pretest as the graduate students, was 80% on both tasks. In the workshop for the 90 faculty in the second group, the momentum comparison task appeared first on the pretest. The success rate was 55% on the momentum comparison task and 60% on the kinetic energy comparison task.

Analysis of the faculty responses suggested that the discrepancy in performance between the two groups was primarily due to the reversal in the order of the questions on the pretest for the second group. The faculty (like the tutorial instructors) often used the answer to the first comparison task to make the second comparison. The relationship between kinetic energy and work was more often recognized than the relationship between momentum and impulse. Therefore, asking the momentum question first appears to have made the pretest more difficult for the second group of faculty. As mentioned earlier, the order in which the questions were presented did not affect the success rate of the introductory physics students on the pretest.

We have no post-test data for the faculty workshops. However, Post-test #2 was given in the graduate teaching seminar during one academic quarter. (See the third column in Table V.) The post-test was given after the relevant seminar and tutorial session had taken place. Only the results from first-time tutorial instructors are shown. The success rate would probably have been 100% (instead of 95%) if one TA had not failed to give explanations. This improvement is consistent with our experience with other tutorials. After participating in the seminar and in the tutorial sessions, the tutorial instructors demonstrate a sound understanding of the concepts involved and the ability to do the reasoning necessary to apply them in a variety of physical situations. Therefore, it is not only the introductory students but also individuals with a strong background in physics who can benefit from the tutorial approach.

VI. CONCLUSION

Assessments of student learning can be made by a variety of methods. Tests that require only a short response (multiple-choice, true-false, etc.) can be administered to large populations in a relatively brief time period. The statistics obtained can give a general indication of student understanding of a range of topics and a rough measure of the prevalence of known student difficulties. However, broad assessment instruments are not sensitive to fine structure and thus may not accurately reveal the extent of student learning. Moreover, such information does not contribute to a research base that is useful for the design of instructional materials. At the other end of the spectrum are in-depth investigations of student understanding. We have found that testing at this level of conceptual detail is an invaluable guide in the development of curriculum.

The results from this study suggest that the use of broad assessment instruments as a sole criterion for student learning can be misleading. The right answer on a multiple-choice test may be triggered in several ways. A correct guess is always a possibility. The recognition of a clue or the elimination of incorrect choices are strategies often used by students. As has been demonstrated, incorrect reasoning may lead to a correct response. When explanations are not required, it can be difficult to determine whether correct answers indicate a functional understanding of the material. Good performance on a multiple-choice test may be a necessary condition, but it is not a sufficient criterion for making this judgment.[27] To be able to address student difficulties effectively, it is necessary to probe for the reasons behind the answers through detailed examination of student thinking. In addition, to ensure that specific difficulties have been successfully addressed, it is necessary to conduct in-depth assessments not only at the institution in which the materials are developed but at others as well. Feedback from pilot sites helps to improve the effectiveness of the materials locally and increases the likelihood that they will be effective in settings other than the one in which they were originally developed. For cumulative improvement in physics education to occur, it is important to determine and to document under which conditions specific instructional strategies are, or are not, successful.[28]

ACKNOWLEDGMENTS

The work described in this paper was a collaborative effort by many members of the Physics Education Group, present and past. Paula R. L. Heron and Peter S. Shaffer made substantive contributions to the research and curriculum development, as well as to the preparation of this paper. Special thanks are also due to Pamela A. Kraus and Luanna G. Ortiz for their assistance in the analysis of the data and to Fred Reif of Carnegie Mellon University for his critical reading of the manuscript. The cooperation of instructors at the University of Washington and at our pilot sites is deeply appreciated. J. Freericks and A. Liu of Georgetown University, E. F. Redish and R. N. Steinberg of the University of Maryland, and E. Kim and S.-J. Pak of Seoul National University were particularly helpful in providing feedback during the curriculum development described in this paper. In addition, we gratefully acknowledge the role of the National Science Foundation, which has enabled us to conduct our coordinated program of research, curriculum development, and instruction. DUE 9354501, our most recent grant, includes support from the Division of Undergraduate Education, other Divisions of EHR, and the Physics Division of MPS.

[1]A comprehensive list of references on research in physics education will be available in a Resource Letter for the *American Journal of Physics* that is being prepared by L. C. McDermott and E. F. Redish.

[2]For examples of research by the Physics Education Group in support of this statement, see, in addition to Ref. 6, L. C. McDermott, ''Millikan Lecture 1990: What we teach and what is learned—Closing the gap,'' Am. J. Phys. **59**, 301–315 (1991); L. C. McDermott and P. S. Shaffer, ''Research as a guide for curriculum development: An example from introductory electricity. Part I Investigation of student understanding,'' *ibid*. **60**, 994–1003 (1992); Printer's erratum to Part I, *ibid*. **61**, 81 (1993); P. S. Shaffer and L. C. McDermott, ''Research as a guide for curriculum development: An example from introductory electricity. Part II Design of instructional strategies,'' *ibid*. **60**, 1003–1013 (1992); and L. C. McDermott, P. S. Shaffer, and M. D. Somers, ''Research as a guide for teaching introductory mechanics: An illustration in the context of the Atwood's machine,'' *ibid*. **62**, 46–55 (1994).

[3]For an example of a supplementary curriculum that has been developed on the basis of research, see L. C. McDermott, P. S. Shaffer, and the Physics Education Group at the University of Washington, *Tutorials in Introductory Physics* (Prentice-Hall, Upper Saddle River, NJ, 1998).

[4]For an example of a self-contained, laboratory-based curriculum that has been developed on the basis of research, see L. C. McDermott and the Physics Education Group at the University of Washington, *Physics by Inquiry* (Wiley, New York, 1996), Vols. I and II.

[5]See, in addition to the articles in Ref. 2, E. Mazur, *Peer Instruction: A User's Manual* (Prentice-Hall, Upper Saddle River, NJ, 1997), pp. 5–7.
[6]R. A. Lawson and L. C. McDermott, ''Student understanding of the work-energy and impulse-momentum theorems,'' Am. J. Phys. **55**, 811–817 (1987).
[7]See Ref. 3.
[8]Some preliminary results from this investigation were presented in a plenary talk at the International Conference on Undergraduate Physics Education, 31 July–3 August 1996, at College Park, MD. The Proceedings of the Conference will be published in 1997 with E. F. Redish and J. S. Rigden as Editors.
[9]See Ref. 6.
[10]This approach is more common among students with a strong mathematical background. The same pretest was given to 27 students enrolled in one recitation section of the introductory physics course at Seoul National University, one of the most selective universities in South Korea. Of the 33% of the students who gave correct responses to both the kinetic energy and momentum comparison tasks about one-half used the theorems and the rest used algebra [Eunsook Kim and Sung-Jae Pak (private communication)].
[11]See Ref. 3.
[12]For additional discussion of the tutorials and the tutorial system, see the last two articles in Ref. 2.
[13]For further discussion and examples of the use of this strategy, see the articles in Ref. 2.
[14]Evidence from research in support of this statement can be found in the articles in Ref. 2. For a discussion of this issue based on extensive teaching experience, see A. B. Arons, *The Various Language: An Inquiry Approach to the Physical Sciences* (Oxford U.P., New York, 1977); and A. B. Arons, *A Guide to Introductory Physics Teaching* (Wiley, New York, 1990).
[15]The data in Table II include results from some classes that had an extra tutorial on the work-energy theorem. The addition of this tutorial did not substantially alter the results obtained when it was not used.
[16]I. A. Halloun and D. Hestenes, ''The initial knowledge state of college physics students,'' Am. J. Phys. **53**, 1043–1055 (1985). This article contains the Mechanics Diagnostic Test.
[17]D. Hestenes, M. Wells, and G. Swackhamer, ''Force Concept Inventory,'' Phys. Teach. **30** (3), 141–158 (1992). The FCI is included in this article. This test has evolved from the Mechanics Diagnostic Test.
[18]D. Hestenes and M. Wells, ''A mechanics baseline test,'' Phys. Teach. **30** (3), 159–166 (1992). This article contains the MBT.
[19]R. Beichner, ''Testing student interpretation of kinematics graphs,'' Am. J. Phys. **62**, 750–762 (1994).
[20]R. K. Thornton and D. R. Sokoloff, ''Assessing student learning of Newton's laws: The Force and Motion Conceptual Evaluation and the Evaluation of Active Learning Laboratory and Lecture Curricula,'' Am. J. Phys. (to be published).
[21]An analysis of a survey in which the FCI was used with 6000 high school and university students is described in a paper by R. R. Hake, entitled ''Interactive engagement vs traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses,'' Am. J. Phys. **65** (1997).
[22]D. Huffman and P. Heller, ''What does the Force Concept Inventory actually measure?'' Phys. Teach. **33** (3), 138–143 (1995); D. Hestenes and I. Halloun, ''Interpreting the Force Concept Inventory—A response to March 1995 critique by Huffman and Heller,'' *ibid*. **33** (8), 502, 504–506 (1995); and P. Heller and D. Huffman, ''Interpreting the Force Concept Inventory—A reply to Hestenes and Halloun,'' *ibid*. **33** (8), 503, 507–511, 1995.
[23]See Ref. 18.
[24]See Ref. 18.
[25]See Ref. 18.
[26]See, for example, the last article in Ref. 2.
[27]This issue is also addressed in R. N. Steinberg and M. S. Sabella, ''Performance on multiple-choice diagnostics and complementary exam problems,'' Phys. Teach. **35** (3), 150–155 (1997).
[28]See the articles in Ref. 2.