CURRICULUM DEVELOPMENT AND THE ASSESSMENT OF STUDENT LEARNING: AN EXAMPLE FROM THE WORK-ENERGY AND IMPULSE-MOMENTUM THEOREMS[1]

(Desenvolvimento curricular e aferição da aprendizagem do aluno: um exemplo dos teoremas trabalho-energia e impulso-momentum)

Lillian C. McDermott
Department of Physics
University of Washington Seattle, Washington, USA

Abstract

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. The results from this study demonstrate the need to probe for the reasons behind the answers that students give. Questions that require students to explain their reasoning are necessary. Implications for the preparation of teaching assistants are discussed.
Key-words: student understanding, work-energy, impulse-momentum.

Resumo

A questão do como aferir a aprendizagem é enfocada no contexto de uma pesquisa sobre a compreensão do aluno em relação aos teoremas trabalho-energia e impulso-momentum. Apresenta-se evidência de que as dificuldades com este assunto vão desde os cursos introdutórios até a pós-graduação, e além dela. Descreve-se o desenvolvimento de uma seqüência instrucional planejada para ajudar os alunos a melhorar sua habilidade de aplicar os teoremas a movimentos reais. Os resultados deste estudo demonstram a necessidade de questionar as razões que estão por detrás das respostas que os estudantes dão. São necessárias questões que exijam que os alunos expliquem seu raciocínio. Discutem-se também implicações para a formação de professores.
Palavras-chave: compreensão do aluno, trabalho-energia, impulso-momentum.

I - INTRODUCTION

During the past two decades, there has been a steadily increasing amount of research on the learning and teaching of physics[2]. 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[3]. We can think of the role that research can play in

---

[1] This paper is based on an article to be published in the American Journal of Physics in 1997. See 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.

[2] Invited talk, VI Interamericam Conference on Physics Education, Córdoba, Argentina. June 29- July 4, 1997. 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.

[3] 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

helping to bridge this gap as having three interrelated components. (See Fig. I.) 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. However, relatively little attention has been directed toward the third component, 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.

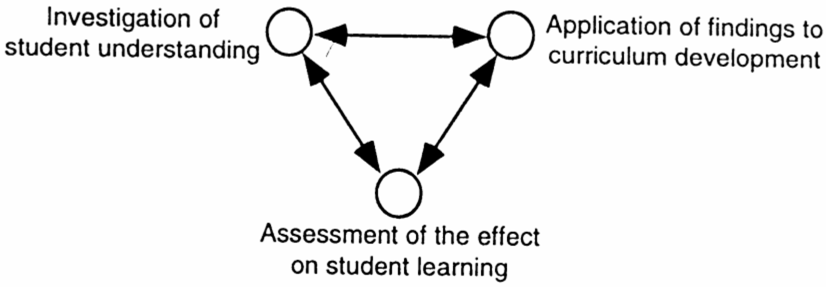


Figure 1

Figure 1: The role of research in physics education.

The means used to assess student learning should be consistent with the instructional goals. There are some basic 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 that they can define operationally and link in a meaningful manner to important principles. They should have developed facility with formal representations and be able to describe in detail the relationship between a concept and the formalism that is used to represent it. They should have developed sufficient proficiency in scientific reasoning to apply the concepts and representations of physics to the analysis and interpretation of simple phenomena. They should be able to make explicit the correspondence between a concept or representation and an actual object or event in the real world. It is, of course, also necessary that students learn how to solve physics problems but the ability to do so does not necessarily indicate that other important goals have been achieved[4].

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," Am. J. Phys. 60, 994-1003 (l992); Printer's erratum to Part I, ibid. 61, 81 (l993); 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," Am. J. Phys. 60, 1003-1013 (l992); 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," Am. J. Phys. 62, 46-55 (l994).

[4] See, in addition to the articles in Ref. 2, E. Mazur, Peer Instruction: A User's Manual (Prentice Hall, Inc., Upper Saddle River, NJ, 1997) pp. 5-7.

Assessments of student learning can be made by a variety of methods. Tests that require only a short response (true-false, multiple-choice, "fill-in-the-blanks", etc.) can be administered to large populations in a relatively brief time period. Results from large-scale testing of this type can give a general indication of student understanding of a range of topics and provide a rough measure of the prevalence of known student difficulties. At the other end of the spectrum are investigations that probe the thinking of individual students. We have found that testing at the level of conceptual detail is an invaluable guide in the development of curriculum[5],[6].

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[7]. 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[8],[9]. The scale has been greatly expanded through the participation of many more students, ranging from the introductory to the graduate level. 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 the article cited[10].

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.

---

[5] 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, to be published in a preliminary version in 1997.

[6] 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, Vols. I and II, (John Wiley & Sons, Inc., New York, NY, 1996).

[7] R.A. Lawson and L.C. McDermott, "Student understanding of the work-energy and impulse-momentum theorems," Am. J. Phys. 55, 811-817, 1987.

[8] See Ref. 4.

[9] 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.

[10] See Ref. 6.

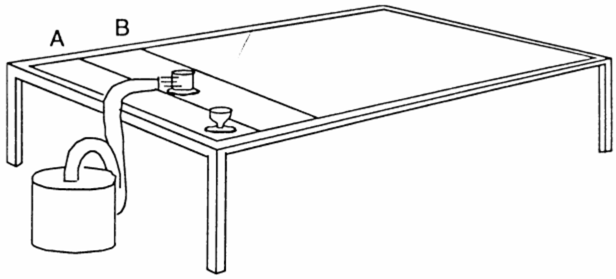


Figure 2

Figure 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.

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 (DX), these reduce to:

$$F \Delta x = \Delta K \text{ and } F \Delta t = \Delta p .$$

The change in kinetic energy (DK) 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 (Dp) is proportional to the time (Dt) each takes to traverse the distance between the lines. Because of its greater mass, a smaller acceleration is impacted 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]

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 had 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.

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%	—

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[11]. The order in which the tasks were presented did not affect the results.

[11] 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.)

3. Incorrect reasoning used by students

Analysis of the written responses revealed reasoning difficulties similar to those identified during the interviews. Most students dad 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 conserve 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[12].

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[13]. 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.

[12] See Ref., 4.

[13] For additional discussion of the tutorials and the tutorial system, see the last two articles in Ref. 2.

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[14]. 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[15]. 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.

---

[14] For further discussion and examples of the use of this strategy, see the articles in Ref. 2.

[15] Evidence from research in support of this statement can be found in the articles in Ref. 2. For discussion of this issue based on extensive teaching experience, see A. B. Arons, The Va tious Language: An Inquiry Approach to the Physical Sciences (Oxford University Press, New York, NY, 1977); and A. B. Arons, A Guide to Introductory Physics Teaching (John Wiley & Sons Inc., New York, NY, 1990)

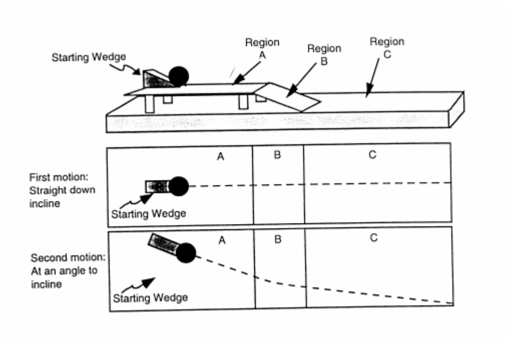
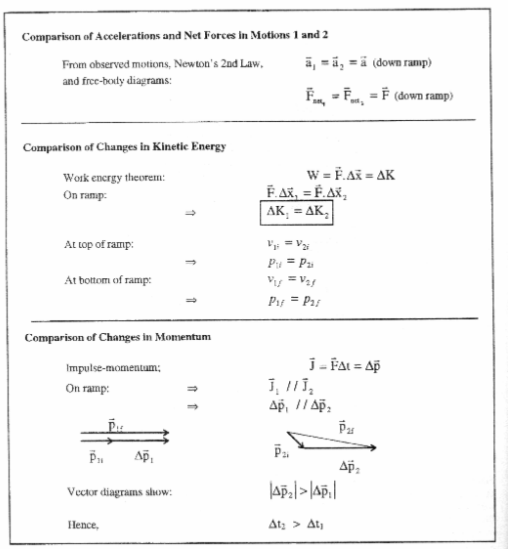


Figure 3

Figure 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.

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 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. Figure 4.: Summary of 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.



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: $W = \vec{F}.\Delta\vec{x} = \Delta K$

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

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

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; $\vec{J} = \vec{F}\Delta t = \Delta\vec{p}$

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

$\Rightarrow$ $\Delta\vec{p}_1 \;//\; \Delta\vec{p}_2$

$\vec{p}_{1f}$ $\vec{p}_{1i}$ $\Delta\vec{p}_1$ $\vec{p}_{2f}$ $\vec{p}_{2i}$ $\Delta\vec{p}_2$

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

Hence, $\Delta t_2 > \Delta t_1$

Figure 4: Summary of 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.

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.

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.

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 set-up 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 #I because a second post-test was developed later.) As can be scan, performance on the post-test was much better than on the pretest[16]. A correct kinetic energy comparison was given by 35% of the students and a correct momentum comparison by 50%.

	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)*	Post-test #1 after lecture and after tutorial (Δt constant) (N = 435)	Post-test #2 after lecture and after tutorial (Δx, $\Delta t \neq$ constant) (N = 320)
Correct kinetic energy comparison	15%	35%	30%
Correct momentum comparison	5%	50%	45%

* Column is repeated from Table I for easy reference. In this case, the written test is regarded as a pretest for the tutorial.

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

[16] 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.

forces. On the pretest, equal forces act for the same distance. On Post-test #I, they act for the same time; on Post-test #2, the forces act over unequal distances for unequal time intervals.

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 #I.) 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 experiences 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 #l. 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 = Dv/Dt, to make the momentum comparison. A few students used the relationship F =Dp/Dt. 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 a time interval causes a change in its momentum. Instead, they re-derived for a specific situation the relationship given by the impulse-momentum theorem. In contrast, we found that students did not re-derive the work-energy theorem to compare the 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 very similar to those 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 all the 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 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.

IV - REASSESSMENT OF POST-TESTS WITH REASONING IGNORED

Several multiple-choice instruments designed to assess student understanding in mechanics have been produced during the past several years[17],[18],[19],[20],[21]. 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 "commonsense" beliefs, both before and after instruction. The Mechanics Baseline Test (MBT), which is intended for use after instruction, covers more topics than the FCI.

---

[17] 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.

[18] D. Hestenes, M. Wells and G. Swackliammer, "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.

[19] D. Hestenes and M. Wells, "A mechanics baseline test," Phys. Teach. 30 (3), 159-166, 1992. This article contains the MBT.

[20] R. Beichner, "Testing student interpretation of kinematics graphs," Am. J. Phys. 62, 750-762, 1994.

[21] 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 for the first and second laws," to be published in 1997.

Two of the questions on the MBT are derived from the kinetic energy and momentum comparison tasks. There was a large disparity between the nationally reported results and the performance of the UW students on these questions[22]. Therefore, we decided to reassess the 435 examination responses on Post-test #I, ignoring the reasons that students gave for their answers. As the second column of Table III shows, the same students had an apparent success rate of 65% for the kinetic energy comparison and 80% for the momentum comparison. Table III 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 on the MBT.

	Results from UW post-tests re-examined *Are correct answers without explanations an 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)*	Correct answer, explanations ignored (Δt constant) (N = 435)	Correct answer, correct explanations (Δx, $\Delta t \neq$ constant) (N = 320)*	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%

* Column is repeated from Table II for easy reference.

Table III: 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 #I, 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.

The results for Post-test #2 are similar. The fourth column of Table III 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 #I, when correct explanations are not required for an answer to be considered correct, the success rate is considerably higher.

In many instances, we found that right answers were given for wrong reasons[23]. A correct answer on a multiple-choice test can be triggered in several ways. A good guess -is always a possibility. The recognition of a clue or the elimination of incorrect choices are strategies often used by students. When explanations are not required, it can be difficult to determine if a correct answer indicates a functional understanding. For example, some

[22] See Ref. 18.

[23] For a more detailed discussion of the differences between the results obtained from an in-depth examination of student learning and from administration of the MBT, see T. O'Brien Pride, S. Vokos, and L.C. McDemott, "The challenge of matching learning assessments to teaching goals: An example from the work-energy and impulse-momentum theorems," to be published in the American Journal of Physics.

students supported their correct answer that the kinetic energies are equal by saying that energy is conserved.

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 participants who were in their first year as tutorial instructors are shown in the first column of Table III. 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[24].

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 III 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 impulse 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

[24] See, for example, the last article in Ref.2.

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 III.) 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.

	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 (Δx = constant) (N = 74)	Pretest before tutorial (Δx = constant) (N = 137)	Post-test #2 after tutorial (Δx, $\Delta t \neq$ constant) (N = 20)
Correct kinetic energy comparison	65%	65%	95%*
Correct momentum comparison	70%	60%	95%*

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

Table IV. 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 tutorial sessions. Only the results from seminar participants in their first year as tutorial instructors are shown

VI - CONCLUSION

The results from this study demonstrate the need to probe for the reasons behind the answers that students give. To ensure that new curriculum will be effective, it is necessary to conduct ongoing in-depth assessments of student learning not only at the institution where the development is taking place but at others as well. Feedback from pilot-sites increases the likelihood that instructional materials will be useful in other settings. 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[25].

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 research, curriculum development, and instruction. DUE 9354501, our most recent grant, includes support from the Division of Undergraduate Education, other Division of EHR, and the Physics Division of MPS.

[25] See the articles in Ref. 2.