CHAPTER 4

Extending Our Assessments: Homework and Testing

If I had to reduce all of educational psychology
to just one principle, I would say this:
The most important single factor influencing learning
is what the student already knows.
Ascertain this and teach him accordingly.
D. P. Ausubel [Ausubel 1978]

The two-level cognitive model we develop in chapters 2 and 3 has implications for the way we evaluate our students and our instruction. First, our understanding of the context dependence of the cognitive response leads us to appreciate the importance of students developing functionality and coherence in their knowledge, not just learning isolated facts or memorizing a small set of problem solutions. Second, our understanding of the associative character of the cognitive response leads us to appreciate the importance of building a good conceptual base and helping students link their conceptual understanding to their problem-solving and analysis skills. Third, our understanding of student expectations and the stability of well-established mental models leads us to appreciate that changing what our students value and what they choose to do to learn is going to require a negotiation between instructor and students.

This negotiation has two important components: (1) We need to offer our students activities that promote the kind of thinking we want to encourage (seeking coherence, sense-making, effective analysis, creativity and so on), and (2) we need to set up mechanisms that permit two-way feedback.

What students actually "do" in our classes usually consists of a variety of activities including both in-class activities, such as listening to a lecture or doing a lab, and out-of-class activities, such as reading the text or doing homework.[1]

---

[1] In the education literature, activities that are done by all students together in a class are referred to as *synchronous*. Activities that are done by students out of class independently at their own time and pace are called *asynchronous*.

In a negotiation, there must be feedback in both directions. We, the instructors, need to know what our students are doing and how much they have learned. Students in our classes need to know when they are meeting the goals we have for them and how to modify their activity when things aren't working.

In this chapter, I discuss the role of homework and testing in our classes. (Other components of the class are discussed in chapters 7–9.) These activities play critical roles in helping our students understand what it means to learn and do physics. I begin with a general discussion of the goals of assessment and evaluation. Next I specifically discuss homework and testing and present some of the methods I have found useful. I then review some of the different kinds of questions we can use both in homework and on exams to help our students understand what it is we want them to get out of our course and to help us understand the extent to which they have done it.

ASSESSMENT AND EVALUATION

If we are interested in probing how well our instruction is working, we can think about the answer along two axes, depending on our answer to the questions "when?" and "what?"

- *When?* Are we asking while the class is still in session or at the end?
- *What?* Are we asking about a student or about the class as a whole?

Probes of the success or failure of our instruction that occur during the class can help both the student and the instructor correct things that are going wrong and identify problems that still have a chance to be fixed. Such probes are called *formative.* Probes that occur at the end of a class and that serve to provide a final certification of student success are called *summative.* There is also a difference in whether we are probing individual students or our instruction as a whole. I refer to probes of an individual student's learning as *assessment* and to probes of our instruction as a whole as *evaluation.*

In this chapter, I focus on looking at assessment through the lens of the question, "What has a particular student learned?" In the next chapter, I consider evaluation and the question, "How effective was my instruction for the group of students in my class?" This separation is somewhat artificial, since when we assess our individual students' learning it tells us something about how successful our instruction has been overall. When we try to evaluate our instruction overall, the only way we can do it is by probing individual students. We will see, however, that some tools that are useful for evaluation of an overall result may not do an adequate job of assessing an individual's learning.

To get feedback on how our class is doing and to measure the success of our interventions, we must first decide what we mean by "success." The discussions in chapters 2 and 3 show that this can be a complex issue. Our choice of goals plays an important role in determining our approach to assessment. What we mean by success is, in turn, determined by our model of student understanding and learning.

To evaluate a student's learning, we have to take account of a number of factors, including conceptual learning, problem-solving skills, coherence and appropriateness of their patterns of association/mental models, and so forth. We must recall that a student may "have" an item of knowledge, that is, be able to recall it in response to a narrow range of cues, but be unable to recall and apply it in a wide range of appropriate circumstances. As a consequence, our probes need to be varied and diverse. Furthermore, we need to watch out for "wishful thinking" such as "filling in the blanks." Often, a student will offer a statement or an equation, and we will assume that they have come to that result in the same way that we would have. Unfortunately, that is often not the case.

GIVING FEEDBACK TO YOUR STUDENTS

As we learned in chapters 2 and 3, students can have ways of looking at and analyzing physics that are quite different from those expected by the teacher. Significant conceptual difficulties may be buried in a student's "wrong" answer to a homework or exam problem, and the student may not have the knowledge or self-evaluation skills to disentangle them. Real feedback on their thinking is immensely valuable to them. Unfortunately, this is one of the first things to be cut when financial or personnel constraints get tight and classes get large. Homework grading disappears in favor of only assigning problems for which the answer is given in the book. Detailed solutions can help somewhat, but since there are often many ways to approach a particular problem, they do not always help students debug their own thinking.

> *Redish's third teaching commandment:* One of the most useful aids you can give your students is detailed feedback on their thinking—in an environment in which they will take note of and make use of it.

HOMEWORK

Sagredo says that he always assigns homework in his physics classes, since it's in doing the problems that the students "really learn the physics." Here's a place where he and I strongly agree. In my experience, far too many students have the expectation that physics is a set of simple "facts" and a list of equations, which, if memorized, should permit them to get an A in my class. Far too few are able to solve the complex kinds of problems that require them to think deeply about the physics and to build a strong conceptual understanding and an appropriate and well-organized knowledge structure. Homework can be of immense value in helping students learn physics, but the wrong kind of homework can also send the wrong message, confirming their view that "facts and equations are enough."

Unfortunately, sometimes despite our best efforts, the wrong message still gets through. End-of-chapter problems given in most introductory texts come in a variety of formats—questions (often conceptual), exercises (straightforward plug-and-chug), and problems (more complex reasoning required). Sagredo and I have both assigned a mix of problems for many years. But recently, I have come to suspect that for many of our students this is still doing them a disservice. A significant fraction of my students appear to think (or hope) that what my course is really about is straightforward plug-and-chug. This is a pretty standard wishful thinking, since these are the problems they can comfortably do. If they write something arbitrary on the questions, get all the exercises right, and write down enough to get partial credit on the problems, they are satisfied that they have gotten all that is necessary out of the assignment. In fact, these students largely miss the main point. The questions are supposed to get them to think deeply about the concepts, and the problems are supposed to help them build a good understanding of how to work with physics. Given the "escape hatch" of the exercises, some students make little effort to get any farther.

As a somewhat draconian attempt to close this loophole, I have stopped giving students exercises as part of their homework. Instead, I reduce the total number of problems and give a smaller number of harder problems, including essay questions and context-rich problems. (See the explanation of these problems below.) One of my better students complained about this. He said he needed the exercises to gain familiarity with the material. I responded that he was perfectly welcome to do them and that many had answers in the back of the book so he could check his work—but that success in this class means much more than plug-and-chug problems and that he would be evaluated on what I wanted him to learn in the end.

Even when we give complex problems, we tend to "degrade" them in response to student pressure. As described in chapter 2 in the section on multiple representations, a major difficulty in solving a realistic physics problem is the extraction of a solvable physics problem from a complex real-world situation. When we assign variable names to the quantities in the problem situation, when we only mention quantities that are relevant and give no extraneous information, when we set up the problem so that all choices of what's important and what's not are explicitly indicated, we steal from the student the opportunity to learn a whole range of essential problem-solving skills.

When students find some tasks difficult, we have a tendency to give in to their pressure and modify the problems so that there is no need for them to develop those skills. A crucial example is the ability to manipulate and interpret equations containing many symbols. Sagredo is aware that students have this difficulty and is disturbed by it. "Why should it matter? A symbol just stands for a number anyway. They should be able to use equations either way." I agree, Sagredo, but recall that experience and familiarity with representations and manipulations make a big difference in the ability to recognize appropriate actions.[2] I once looked carefully at a couple of standard introductory calculus texts. To my surprise I found that very few equations anywhere in the text contained more than one symbol. Those that did followed a strict convention: x, y, z, and t were variables, while a, b, and c were constants. If we want students to learn the skill of working with multisymbol equations, we need to provide examples that will make them use it. Problems in which numbers are given right away allow the students to avoid having to learn the skill of interpreting equations with multiple symbols.

I find that weekly homework using a variety of challenging problems (especially if some are marked "from an hour exam" or "from a final exam") can be effective in helping to reorient students' ideas of what they need to do to learn physics.

In addition to providing a rich venue for student activity, homework also plays a critical role in providing students with formative feedback. In many large classes, we cannot give quizzes and examinations frequently enough to provide students enough feedback to modify a learning approach they may have developed in high school or in other science classes. As I discussed in chapter 3, student expectations play a major role in what they do. Even if you give them good feedback, there is no guarantee that they will use it. If homework problems are graded with comments but handed back weeks after they were done, the students might

---

[2] Recall, for example, the Wason card (K2A7) example discussed in chapter 2.

not be able to reconstruct their state of mind when they did the problems. Any difficulty they have interpreting written feedback is likely to result in their ignoring it.

GETTING FEEDBACK FROM YOUR STUDENTS

The Ausubel quote in the epigraph, our cognitive model, and my experience all agree: Feedback needs to work both ways. You can improve your teaching substantially by getting good and regular feedback about where your students are, what they are thinking, and how they are interpreting the information provided in your course.

> *Redish's fourth teaching commandment:* Find out as much as you can about what your students are thinking.

Four plausible and frequently used approaches to getting feedback from your students and evaluating their learning are:

1. Observe student behavior in class and in office hours.
2. Measure student satisfaction with a survey or questionnaire.
3. Measure student learning using a closed-ended question (multiple-choice or short-answer), designed using the results of physics education research on commonly found errors to specify attractive distractors.
4. Measure student learning using open-ended (long-answer or essay) exam questions—problems or open-expression questions in which students explain and discuss their answers.

The first approach is an essential part of understanding our students and seeing how they are responding to our instruction. Here, we have to be particularly careful not to go into wishful-thinking mode and assume that a question that a student asks in office hours simply needs to be answered in a direct and straightforward manner. This may sometimes be the case, but more often, I have found that asking a few well-placed questions such as "Well, what do *you* think is going on?" or "Why do you ask?" or even "Could you explain why you're stuck?" often produces a lot of information—and indicates that my "gut-response" answer would have been useless.

> *Redish's fifth teaching commandment:* When students ask you a question or for help, don't answer right away. Ask them questions first, in order to see whether your assumptions about their question are correct.

The second approach, an attitude survey or questionnaire, is the simplest and most commonly used, but although student satisfaction is important in motivating student work, and presumably therefore in producing student success, the link between satisfaction and learning is highly indirect. Indeed, students whose primary goal is a good grade may find higher satisfaction in a course that produces a good grade without improved learning, since improved learning often requires time and painful effort.

The third method, multiple-choice questions with research-based distractors or short-answer questions with research-motivated contexts, is easy to deliver, but, as explained in chapter 5, requires a substantial effort to develop effectively. I discuss four kinds of such questions below: multiple-choice with research-based distractors, multiple-choice multiple-response, representation translation, and ranking tasks. Other formats for short-answer questions have been developed and can be quite effective [Peterson 1989].

The fourth approach, long-answer or open-ended questions, is easy to deliver, but the grading and analysis of long-answer questions and problems can be time consuming. Student answers must be read in detail and evaluated by the understanding displayed. Grading here can be quite subtle. There is a tension between grading that is too draconian and only gives credit for detailed and careful reasoning, and grading that is too casual and gives points for remembered equations or principles that the student has no idea how to use. The former tends to produce grades that are too low and the latter to send the message that understanding is not important. Later in this chapter, I discuss four kinds of long-answer questions: open-ended or context-rich reasoning problems, estimation questions, qualitative questions, and essay questions.

TESTING

Often the standard testing we carry out to assess our students is limited, especially in large classes. In order to facilitate grading, we might choose to resort to machine-graded closed-answer questions—multiple-choice end-of-chapter problems that rely primarily on recognizing the correct equation and performing some straightforward algebraic manipulations to get a number. Decisions to test in this way have a powerful impact on our instruction. Not only do we severely limit the kind of feedback we receive about how our students are doing, but we send them a powerful message about what we believe is important for them to learn.

College students are as busy as their instructors, having many courses to consider, not just physics. In addition, many of them are young adults seeking partners and engaging in extramural sports or clubs, and some are working long hours to pay for their tuition. Some of these extramural activities will have more impact on their lives and careers than their physics class, hard as this might be to believe! As a result, even if I provide my students with good advice—such as that working something out in multiple ways is important for their learning, or that once they have finished a problem they should think about it to see that it makes sense—they are likely to ignore it if they can't see how doing it will pay off directly in points and grades. Most students regularly practice "time triage," doing only those activities that seem absolutely necessary for getting through the class successfully.[3]

In order to get students to engage in the kind of intellectual activity we want them to engage in, we have to let them know by testing them for it directly. In particular, in choosing exam questions, we have to be aware of the mixed messages we may be sending. This can be quite difficult, especially in these days where a class's success is measured by "student happiness" as evaluated with end-of-class student anonymous comments, as opposed to being measured by some assessment of student learning. Many students are satisfied with getting

---

[3] And of course "successfully" means different things to different students. For a few it means showing their parents that they were not cut out to be doctors or engineers or whatever goals their parents may have imposed on them.

through a class with a decent grade without having to do too much work—even if they learn little or nothing—and on end-of-semester surveys they may reward teachers who help them achieve this goal and punish those who make them work too hard.[4]

Designing exams

The kinds of questions we choose for an exam can make a big impact on what students choose to do in our classes. If, for example, we construct our exams from multiple-choice questions that can be answered by recognition (that is, there are no "tricky" or tempting alternatives), our students are likely to be satisfied with reading their text and lecture notes "lightly" a number of times. If we construct our exams from homework problems they have been assigned, our students are likely to seek out correct solutions from friends or solution books and to memorize them without making the effort to understand or make sense of them. If all our exam problems can be solved by coming up with the "right equation" and turning the crank, students will memorize lists of equations and pattern match. If we allow our students to create a card of equations and bring it into the exam and then only test the application of equations, students are likely to forego understanding altogether and only practice equation manipulation. These kinds of "studying" have minimal impact on real learning and are often highly ephemeral.

If we really want students to study physics deeply and effectively in a way that will produce long-term learning, the activities we provide are not enough; the learning they foster has to be validated via testing in their examinations.

> *Redish's sixth teaching commandment:* If you want your students to learn something, you have to test them on it. This is particularly true for items in the "hidden curriculum" (cf. chapter 3).

Exams as formative feedback

Exams and quizzes are not only means of carrying out summative assessments of what students have learned. They can also provide formative feedback to students as to what they have learned and what they need to work on. Unfortunately, it has been my experience that most students have the expectation that exams only serve as summative assessments. If they do poorly, they respond, "Well I sure messed that up. I need to do better on the next topic." Of course, since physics is highly cumulative, both in content and in skill building, having a significant "hole" in one's knowledge can lead into a downward spiral.

There are ways of designing the presentation and delivery of exams and quizzes to help encourage students to pay attention to the mistakes they make on exams. I've developed a pattern of exam delivery that seems to work reasonably well to encourage at least some students to debug their thinking from the feedback they get on exams.

- The exam is given in class during the last class of the week.
- The exam is graded immediately (over the weekend) and returned in the first class of the next week.

---

[4] Be careful! The interpretation that one's bad responses from students is because one is making them work too hard is a classic "wishful thinking."

- I return the exams to the students at the beginning of the class and go over the exam in class, explicitly showing the partial credit grading patterns, if any.
- I encourage the students to look for grading errors or possible interpretations of the problem's wording that would improve their grade. I tell them to submit a written request for a grade change with a detailed explanation of why they think they should have received more points. (Just saying "please look at problem 4" doesn't count. They have to explain why they think they were right and the grader was wrong.)
- If they are dissatisfied with their grade on the exam, they may take an out-of-class makeup test on the same material (but not the same test). However, the grade on the makeup does not replace the first test: they get the average of the two grades.
- I tell them that my experience is that a student who got a low grade who simply goes back and studies again in the same way (or does not study at all) is as likely to go down on the makeup and lose points as a result. On the other hand, students who specifically study their first exam in order to understand the mistakes they made and why have a high probability of raising their grade substantially.
- The class averages on each question are reported and they are told that the final exam will include at least one question based on the questions the class performed the most poorly on.

Since this procedure takes up two lecture periods for every exam and requires me to write two exams for every exam, I tend to reduce the number of hour exams I give in a semester: two instead of three in a course with three 50-minute lectures per week; one and two shorter quizzes in a course with two 75-minute lectures per week. In a lecture class of 100 to 200, I usually get about 25% of the students taking the makeup. This has been effective in making the students more comfortable with my unfamiliar exam style. (I always include essay questions, estimation questions, and representation translation questions, along with a more traditional problem or two.)

The last point in the bulleted list sends the message that I expect them to know everything I am testing them on. If the class as a whole performs badly, I review the physics behind the problem carefully when I go over it in class and tell them that they all need to go back and look at that material again—and that they are certain to have a question concerning that topic on the final exam.

EIGHT TYPES OF EXAM AND HOMEWORK QUESTIONS

A wide variety of structures are available for presenting physics problems to our students. Different types of structures tend to activate different kinds of associations and resources. A critical element in developing a course that gets students thinking more deeply about physics is choosing from a broad palette of question types. In the rest of this chapter, I discuss eight kinds of problems, briefly describe their value, and give an example or two of each. In addition to the kinds of problems discussed here, see also the discussion of Peer Instruction and JiTT in chapter 7 and of Cooperative Problem Solving in chapter 8 where other problem structures are used.

The Physics Suite has a large collection of problems, being the union of the set developed over the years for *Fundamentals of Physics* [HRW6 2001], the set of *Workshop Physics* problems, and the set I have developed over the years for the Activity-Based Physics project. These problems are distributed in the text, *Understanding Physics*, in the problem volume, and some are available on the Resource CD in the back of this book. Additional sources of problems and approaches to testing may be found in [Arons 1994], [Tobias 1997], and the books listed in the file "Other Resources" on the Resource CD associated with this volume.

Multiple-choice and short-answer questions

Multiple-choice and short-answer questions are tempting to use because they are easy to grade. The results can be highly suggestive, but multiple-choice tests can be difficult to interpret. They tend to overestimate the student's learning since they can sometimes be answered correctly by means of incorrect reasoning[5] or by cued responses that fail to represent functional understanding. On the other hand, the use of common misconceptions or facets as distractors produces "attractive nuisances" that challenge the students' understanding. Students who get the correct answer despite these challenges are likely to have a reasonably good understanding. A well-constructed multiple-choice question based on research into common naïve conceptions can therefore give some indication of the robustness of a student's correct answer. Note, however, that standard multiple-choice questions developed by instructors who have not studied the relevant research often have distractors that are too trivial to provide a real test of students' understanding. Instructors unaware of research results in physics education sometimes find it difficult to imagine the kinds of errors that students will commonly make.

An example of a good multiple-choice question with research-based distractors is shown in Figure 4.1. The distractors have to correspond to students' naïve conceptions, not to what faculty think. Since most physics instructors know very well that imbedding an object in a fluid produces a buoyant force, they may find the answers to the problem in Figure 4.1 peculiar. But

A book is at rest on a table top. Which of the following force(s) is (are) acting on the book?

1. A downward force due to gravity.
2. The upward force by the table.
3. A net downward force due to air pressure.
4. A net upward force due to air pressure.

(A) 1 only
(B) 1 and 2
(C) 1, 2, and 3
(D) 1, 2, and 4
(E) None of these. Since the book is at rest, there are no forces acting on it.

Figure 4.1 A multiple-choice question from the FCI [Hestenes 1992].

---

[5] See, for example, [Sandin 1985].

it is well documented that many high school students and some college students think that air pressure is responsible for gravity, pushing down on an object with the weight of the air above it, so item (C) is an attractive distractor for those students.[6] A large collection of multiple-choice questions with research-based distractors is provided in the conceptual surveys of various topics in physics given on the Resource CD associated with this volume.

Short-answer questions can also explore students' functional understanding of physics concepts effectively. The key is not to make the question a simple recognition test but to require some reasoning, perhaps using ideas or principles that are not directly cued by the problem. In the sample shown in Figure 4.2, the problems naturally cue up buoyant forces, but they require a fairly sophisticated application of free-body diagrams. (The answer to 2.2, for example, is obviously "=" since both net forces are equal to 0 by Newton's first law. Students who do not comfortably distinguish individual and net forces have trouble with this.)

Multiple-choice multiple-response questions

A substantial amount of thinking and reasoning can be required from a student in a multiple-choice multiple-response test. A sample is shown in Figure 4.3. To solve this problem, a

For each of the following partial sentences, indicate whether they are correctly completed by the symbol corresponding to the phrase *greater than* (>), *less than* (<), or *the same as* (=).

(2.1) A chunk of iron is sitting on a table. It is then moved from the table into a bucket of water sitting on the table. The iron now rests on the bottom of the bucket. The force the bucket exerts on the block when the block is sitting on the bottom of the bucket is _______ the force that the table exerted on the block when the block was sitting on the table.

(2.2) A chunk of iron is sitting on a table. It is then moved from the table into a bucket of water sitting on the table. The iron now rests on the bottom of the bucket. The total force on the block when it is sitting on the bottom of the bucket is _______ it was on the table.

(2.3) A chunk of iron is sitting on a table. It is then covered by a bell jar which has a nozzle connected to a vacuum pump. The air is extracted from the bell jar. The force the table exerts on the block when the block is sitting in a vacuum is _______ the force that the table exerted on the block when the block was sitting in the air.

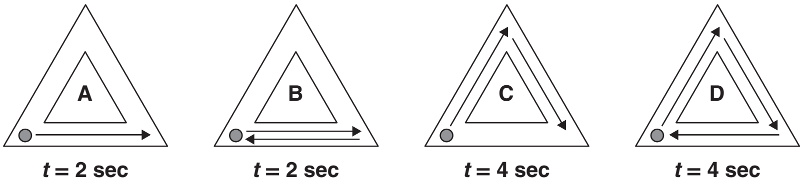
(2.4) A chunk of iron is sitting on a scale. The iron and the scale are then both immersed in a large vat of water. After being immersed in the water, the scale reading will be _______ the scale reading when they were simply sitting in the air. (Assume the scale would read zero if nothing were sitting on it, even when it is under water.)

Figure 4.2 Sample short-answer question combining the use of different physics principles.

---

[6] Then, Sagredo asks, where does the weight of the air come from? A good question, but one not often asked by these students.

Four different mice (labeled A, B, C, and D) ran the triangular maze shown below. They started in the lower left corner and followed the paths of the arrows. The times they took are shown below each figure.



For each item below, on your answer sheet write the letters of all of the mice that fit the description.

(a) This mouse had the greatest average speed.
(b) This mouse had the greatest total displacement.
(c) This mouse had an average velocity that points in this direction (⇒).
(d) This mouse had the greatest average velocity.

Figure 4.3 A sample multiple-choice multiple-response question.

student has to have good control of the concepts of vector displacement and average vector velocity, and to be able to clearly distinguish velocity and speed.

This type of question requires a student to evaluate each statement and make a decision about it. It is a particularly useful type of question in cases where students tend to have mixed context-dependent models of a physical situation.

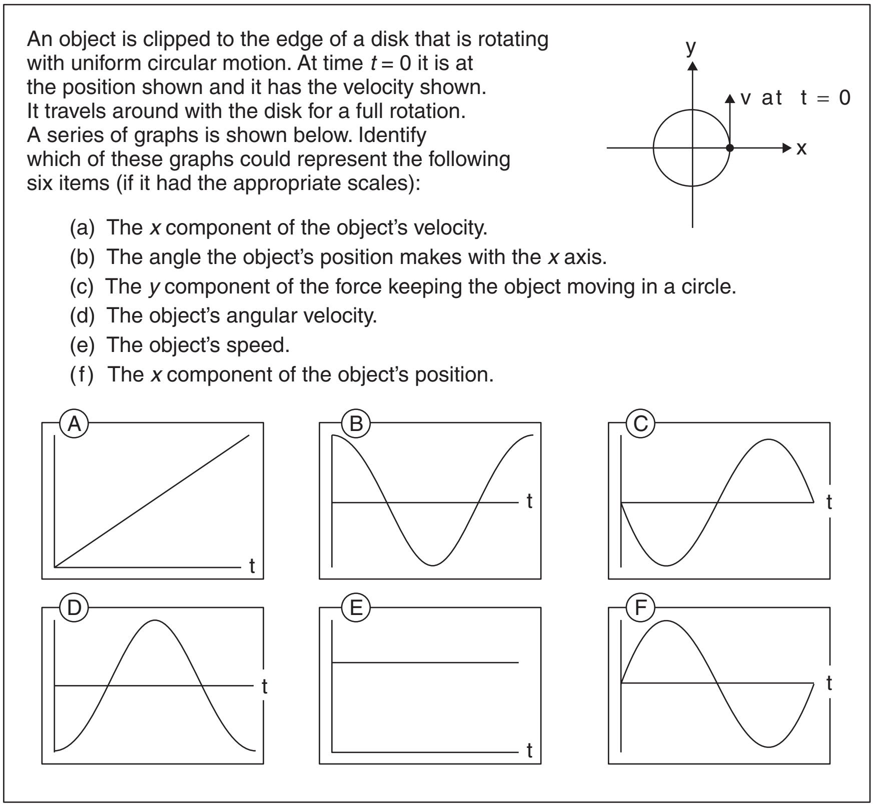
Representation-translation questions

As discussed in chapter 2, learning to handle the variety of representations we use can be quite a challenge for introductory students, but it can be one of the most valuable general skills they develop from studying physics. The presentation of a single situation in different ways facilitates understanding and sense-making of different facets of the situation. Furthermore, one of the primary difficulties novice students have with problem solving is their failure to be able to visualize the dynamic situation and map a physics description onto that visualization.

Thornton and Sokoloff [Thornton 1998] pioneered the use of problems in which students are required to match a statement in words describing a physical situation with graphs or pictures that describe those situations. You will find many such problems in their concept surveys contained in the Action Research Kit on the Resource CD: the FMCE, the ECCE, the VET, and the MMCE.

I have had success getting students to think about the meaning of physics variables with problems in which they are shown a situation and a number of graphs. The graphs have their abscissa labeled as time, but the ordinates are unmarked. The students then have to match a list of physical variables to the graphs. An example is shown in Figure 4.4.

I had an interesting experience with a student concerning the representation translation shown in Figure 4.5. After the exam in which this problem was given, the student (an



An object is clipped to the edge of a disk that is rotating with uniform circular motion. At time $t = 0$ it is at the position shown and it has the velocity shown. It travels around with the disk for a full rotation. A series of graphs is shown below. Identify which of these graphs could represent the following six items (if it had the appropriate scales):

(a) The x component of the object's velocity.
(b) The angle the object's position makes with the x axis.
(c) The y component of the force keeping the object moving in a circle.
(d) The object's angular velocity.
(e) The object's speed.
(f) The x component of the object's position.

Figure 4.4 A sample representation-translation question.

engineer) came in to complain. "I don't understand how to do these problems," he grumped. I asked what answer he chose for part (a). He reported that he had selected option (e). "And why did you choose that one?" I asked. "Well," he said, "it's a wave and waves are supposed to be wiggly. That was the wiggliest."

"Okay," I responded. "Now tell me what's happening."

"What do you mean?" he asked.

I said, "Describe the string, tell me where the dot is. Then tell me what happens to the string and the dot as the pulse moves down the string."

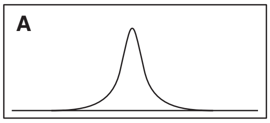
"Hmmm," he said. "Well, the pulse moves to the right, when it gets to the dot, the dot moves up and then down . . . Oh, [expletive deleted]!" Once he had worked through visualizing what happened, he was almost trivially able to solve the rest of the problem.

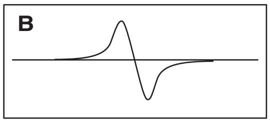
Although I have not carried out explicit research on the topic, in my experience, the students in both algebra- and calculus-based physics struggle with these problems and the struggle is extremely productive.

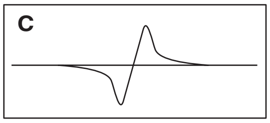
Consider the motion of a pulse on a long taut string. We will choose our coordinate system so that when the string is at rest, the string lies along the *x* axis of the coordinate system. We will take the positive direction of the *x* axis to be to the right on this page and the positive direction of the *y* axis to be up. Ignore gravity. A pulse is started on the string moving to the right. At a time t_0 a photograph of the string would look like Figure A below. A point on the string to the right of the pulse is marked by a spot of paint.

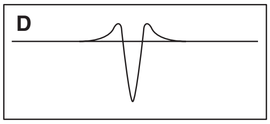
For each of the items below, identify which figure below would look most like the graph of the indicated quantity. (Take the positive axis as up.) If none of the figures look like you expect the graph to look, write *N*.

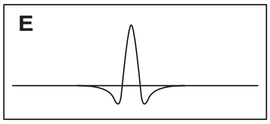
(a) The graph of the *y* displacement of the spot of paint as a function of time.
(b) The graph of the *x* velocity of the spot of paint as a function of time.
(c) The graph of the *y* velocity of the spot of paint as a function of time.
(d) The graph of the *y* component of the force on the piece of string marked by the paint as a function of time.


A


B


C


D


E

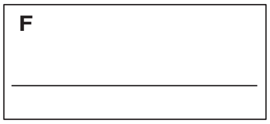

F

Figure 4.5 A sample representation-translation problem.

Ranking tasks

Another class of easy-to-grade but effective problems are *ranking tasks*—problems in which the student must order a series. These have been used by a number of researchers and curriculum developers effectively.

Ranking tasks are effective because they easily trigger reasoning primitives[7] such as "more of a cause produces (proportionately) more of an effect" [diSessa 1993]. An example of a ranking task from the UWPEG is shown in Figure 4.6.

David Maloney has been a long-time user of ranking tasks in his research. Recently, he and his collaborators, Tom O'Kuma and Curt Heiggelke, published a collection of these problems [O'Kuma 1999]. A sample is given in Figure 4.7. Many of the Reading Exercises in *Understanding Physics* are ranking tasks.

---

[7] Primitives are discussed in detail in chapter 2.

In the picture shown below are 5 blocks. The blocks have equal volumes but different masses. The blocks are placed in an aquarium tank filled with water, and blocks 2 and 5 come to rest as shown in the lower figure. Sketch on the figure where you would expect the blocks 1, 3, and 4 to come to rest. (The differences in mass between successive blocks is significant—not just a tiny amount.)

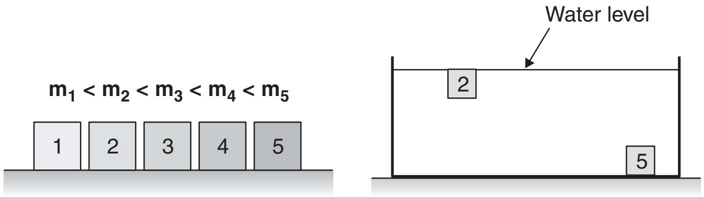


Figure 4.6 A ranking task [Loverude 1999].

Given below are seven arrangements of two electric charges. In each figure, a point labeled *P* is also identified. All of the charges are the same size, *Q*, but they can be either positive or negative. The charges and point *P* all lie on a straight line. The distances between adjacement items, either between two charges or between a charge and the point *P*, are all *x* cm. There is no charge at point *P*, nor are there any other charges in this region.

Rank these arrangements from greatest to least on the basis of the strength of the electric field at point *P*. That is, put first the arrangement that produces the strongest field at point *P*, and put last the arrangement that produces the weakest field at point *P*.

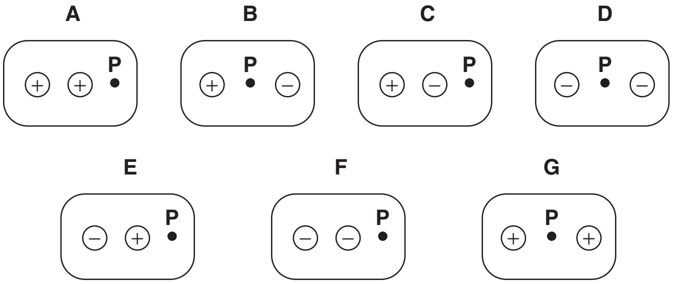


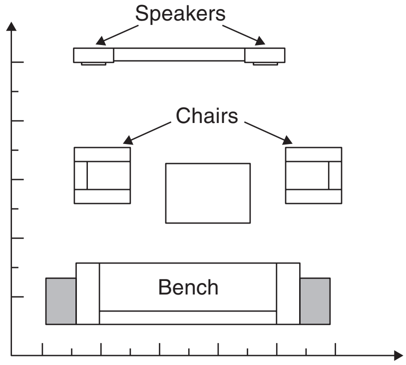
Figure 4.7 A ranking task from [O'Kuma 1999].

Context-based reasoning problems

The problems I find most valuable, both on homework and on exams, are those I call *context-based reasoning problems*.[8] In these problems, students are given a reasonably realistic situation and have to use physics principles—often in ways or circumstances in that they have not been previously seen—to come to a conclusion. The crucial fact is that the answer to the

I have set up my two stereo speakers on my back patio as shown in the top view diagram in the figure at the right. I am worried that at certain positions I will lose frequencies as a result of interference. The coordinate grid on the edge of the picture has its large tick marks separated by 1 meter. For ease of calculation, make the following assumptions:

- Assume that the relevant objects lie on integer or half-integer grid points of the coordinate system.
- Take the speed of sound to be 343 m/s.
- Ignore the reflection of sound from the house, trees, etc.
- The speakers are in phase.



(a) What will happen if I am sitting in the middle of the bench?

(b) If I am sitting in the lawn chair on the left, what will be the lowest frequency I will lose to destructive interference? (If you do not have a calculator, leave the result as an expression with numbers that could be simply evaluated and estimate the result to one significant figure.)

(c) Can I restore the frequency lost in part (a) by switching the leads to one of the speakers, thereby reversing the phase of that source?

(d) With the leads reversed, what will happen to the sound for a person sitting at the center of the bench?

Figure 4.8 A context-based reasoning problem.

---

[8] These are similar in spirit to *Context-Rich Problems* used by the Minnesota group. See the discussion of Cooperative Problem Solving in chapter 8.

problem should be of some reasonable real-world interest. Unfortunately, too many end-of-chapter physics problems are poorly motivated. They seem simply an exercise in carrying out some obscure physics calculations for no obvious purpose. A problem such as “How much work is done by a weight-lifter in raising a 200-kg barbell a distance of 2 meters” is of this type. Why is the “work” something I should care about? What makes this calculation relevant? A problem such as “Estimate the number of calories a marathon runner burns in running 26 miles. Does he need to stock up on carbohydrates before beginning?” is better motivated.

An example that I like is shown in Figure 4.8. This is particularly nice since it relies on the fundamental idea of interference—figure out the path difference and see how many wavelengths fit in. The geometry does not permit the use of the small-angle approximation that leads to the standard interference formulas. Students have to calculate distances using the Pythagorean Theorem. It’s not too realistic because with sound, reflections from nearby walls and objects are of primary importance. But this calculation is one part of understanding what is going on.

Everyday examples in newspapers, advertising, television, and movies in which the physics is wrong make nice problems of this type. An example is given in Figure 4.9.

In the movie *Jurassic Park,* there is a scene in which some members of the visiting group are trapped in the kitchen with dinosaurs on the other side of the door. The paleontologist is pressing his shoulder near the center of the door, trying to keep out the dinosaurs. The botanist throws her back against the door at the edge next to the hinge. She cannot reach a gun on the floor because she is trying to help hold the door closed. Would they improve or worsen their situation if he moved to the door’s outer edge and she went for the gun? Estimate the change in the torque they would exert on the door if they changed their positions as suggested.

Figure 4.9 A context-based reasoning problem.

> Estimate the number of blades of grass a typical suburban house's lawn has in the summer.

Figure 4.10 A typical Fermi question.

Estimation problems

Estimation problems were made famous by Enrico Fermi, who was a master at them. His classic question was "How many barbers are there in Chicago?" Questions of this type can have considerable value because students

- get to practice and apply proportional reasoning
- learn to work with large numbers
- learn to think about significant figures (I always deduct points for too many significant figures.)
- learn to quantify their real-world experience

An example of a typical Fermi question is given in Figure 4.10. I use this question to explain to my class what it is that I want them to do when they solve problems of this type.[9] To get a number, students must picture a lawn, estimate its area, and estimate the number of blades of grass in a given area of lawn and then scale up. I always require that they start from a number they can plausibly know. Thus, if they said "let's assume that a square meter of lawn has a million blades of grass," I would give no credit for that estimate. If, on the other hand, they said,

> Consider a square centimeter of grass. There are about 10 blades of grass counting along each side. So there will be 100 in a square centimeter and 100 × 100 times that in a square meter. So assume one million blades of grass per square meter.

that answer would receive full credit. I always grade (or have graded) my estimation problems so that points are given for each part of the reasoning required. I explain carefully to my students that they are going to be evaluated on how they come up with the answer, not just on the answer.

An instructor needs to be both persistent and patient to include estimation problems. At first, students may not believe that you are serious about asking them to do problems of this type. For this reason, I give an estimation problem on every homework and exam, and I identify some homework estimations as having come from previous exams so that students

---

[9] This question is appropriate for Maryland, where the lawns in the summer still have grass. It might be too easy for a student at the University of New Mexico in Albuquerque.

will know I'm serious. Second, students may not believe that this is the sort of thing they are supposed to be learning in physics class and resent doing them at first. One student complained on his post-class anonymous survey that "exam points were given for being a good guesser." In my experience, this difficulty passes as they gain skill and confidence. Near the end of one of my algebra-based physics classes, one of my students told me with great glee, "Prof, you know those estimation problems? Well I'm taking a business class and we're doing, like, making a business plan and, you know, it's just like estimations? I was the only one in the class who knew how to do it!"

As the class moves on through the year, I begin to blend more physics into estimation problems so that they become design problems. These seem to play a big role in helping students understand the long-term value of physics for their professional future. An example of an "estimation/design" problem is given in Figure 4.11.

Qualitative questions

Qualitative questions can be quite effective in getting students to learn to think about concepts—and in helping instructors realize that students may not have learned what they thought they had from observing good performance on quantitative problem solving.[10] The University of Washington group has been particularly effective in using this kind of question in curriculum design.[11]

A sample qualitative problem that is an extension of a ranking task is given in Figure 4.12. I first encountered questions of this type when visiting the University of Washington on sabbatical in 1992. My inclination was to label each resistor, *R*, label each battery, *V*, write down all the equations using Kirchhoff's laws, and solve them for the relevant unknowns. This method certainly works! But one of the facilitators[12] asked me whether I could solve it without equations. "Why should I?" I responded. "Because," he said, "perhaps your students are not as facile with equations as you are." I proceeded to give it a try and was quite surprised at how difficult it was. I realized that I was using the physics equations as a scaffold for organizing my conceptual knowledge. For students who have not developed this knowledge structure and who might be hazy about what the concepts mean, my approach would be largely inaccessible. The approach of reasoning through the problems conceptually, on the other hand, helps reveal students' naïve conceptions (such as the idea that a battery is a source of constant current). As they work through these problems on homework and struggle with them on exams, they seem to build the concepts more firmly than they would if they could get away with a purely mathematical approach. Note that an essential part of questions of this type is the phrase "explain your reasoning." Many students don't know what reasoning or building a case means [Kuhn 1989]. Having them explain their reasoning, discussing what you mean by reasons, and giving them feedback on the sort of things we mean by reasoning can be an important part of what they learn in a physics class.

---

[10] See the discussion of Mazur's experience in chapter 1.

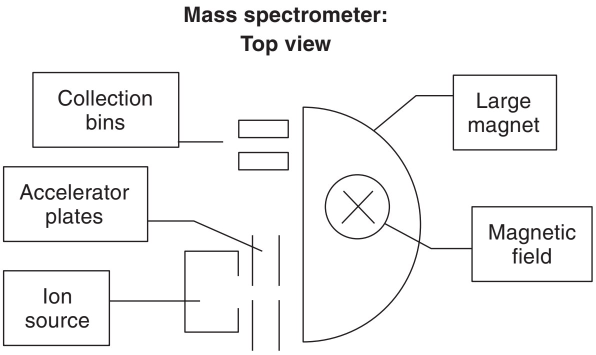
[11] See the discussion of Tutorials in Introductory Physics in chapter 8 and Physics by Inquiry in chapter 9.

[12] Richard Steinberg, with whom I later had the privilege of collaborating extensively.

You are assigned the task of deciding whether it is possible to purchase a desk-top-sized magnetic spectrometer in order to measure the ratio of C^{12} to C^{14} atoms in a sample in order to determine its age.

For this probem, let's concentrate on the magnet that will perform the separation of masses. Suppose that you have burned and vaporized the sample so that the carbon atoms are in a gas. You now pass this gas through an "ionizer" that on the average strips one electron from each atom. You then accelerate the ions by putting them through an electrostatic accelerator—two capacitor plates with small holes that permit the ions to enter and leave.

The two plates are charged so that they are at a voltage difference of ΔV Volts. The electric field produced by charges on the capacitor plates accelerate the ions to an energy of $q\Delta V$. These fast ions are then introduced into a nearly constant, vertical magnetic field. (See the figure below.) If we ignore gravity, the magnetic field will cause the charged particles to follow a circular path in a horizontal plane. The radius of the circle will depend on the atom's mass. (Assume the whole device will be placed inside a vacuum chamber.)



Answer three questions about how the device works.
(a) We would like not to use too high a voltage. If ΔV is 1000 Volts, how big a magnetic field would we require to have a plausible "table-top-sized" intrument? Is this a reasonable magnetic field to have with a table-top sized magnet?

(b) Do the C^{12} and C^{14} atoms hit the collection plate far enough apart? (If they are not separated by at least a few millimeters at the end of their path we will have trouble collecting the atoms in separate bins.)

(c) Can we get away with ignoring gravity? (*Hint:* Calculate the time it would take the atom to travel its semicircle and calculate how far it would fall in that time.)

Figure 4.11 An estimation/design problem.

In the circuits below, all the bulbs are identical, and the batteries are identical and nearly ideal. Rank the brightness of amount of (current through) the bulbs. Explain your reasoning.

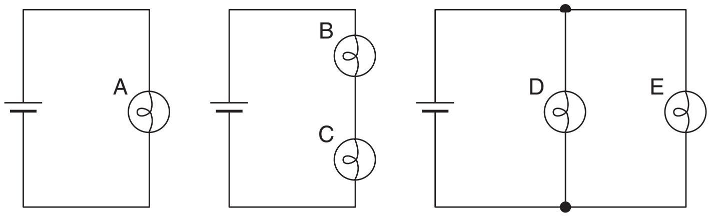


Figure 4.12 Sample qualitative question on direct current circuits [McDermott 1992].

Of course, in the end I want my students to develop a well-organized knowledge structure and to have equations mapped firmly into that structure. But it is important not to demand too much too early.

Qualitative questions that require identifying relevant physics principles and concepts, qualitative reasoning, and writing an explanation can be effective in helping students make the connection between their real-world personal experiences and the physics they are learning. An example of a question of this type is given in Figure 4.13.

Essay questions

Essay questions can be the most revealing of students' difficulties and naïve conceptions of any form. In the example given in Figure 4.14, the students are not asked to recall the law—it is given to them. But they are asked to discuss its validity. In this case, we had completed an ABP Tutorial (see chapter 8) exploring Newton's third law (N3) with two force probes. I had wanted them to refer to experimental evidence of some sort in justifying their belief in N3. Interestingly enough, a significant fraction of students did not expect N3 to hold in all

In public restrooms there are often paper towel dispensers that require you to pull downward on the towel to extract it. If your hands are wet and you are pulling with one hand, the towel often rips. When you pull with both hands, the towel can be extracted without tearing. Explain why.

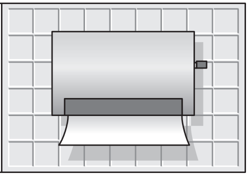


Figure 4.13 A sample qualitative reasoning problem that makes a connection to the student's personal experience.

Newton's 3rd law says that objects that touch each other exert forces on each other.

> *If object A exerts a force on object B, then object B exerts a force back on object A and the two forces are equal and opposite.*

Consider the following three situations concerning two *identical* cars and a much heavier truck.

(a) One car is parked and the other car crashes into it.
(b) One car is parked and the truck crashes into it.
(c) The truck is pushing the car because the car's engine cannot start. The two are touching and the truck is speeding up.

In which of these situations do you think Newton's 3rd law holds or does not hold? Explain your reasons for saying so. (Your answers are worth 5 points, your reasons 10.)

Figure 4.14 A sample essay question on Newton's 3rd law.

these cases. The results provided me guidance for follow-up discussions in lecture and later problems on exams. In the example given in Figure 4.15, I learned in a dramatic fashion that a large majority of my students had significant difficulties with the concept of "field" despite my careful attempts to be precise in lecture.

Note that giving essay questions in an exam situation has to be done with some care. There are three points to consider.

1. In some university environments (such as mine), introductory students often have all their classes as large lectures and they may have little or no experience with actually having to write during an examination.

In this semester we have studied two fields: electric and magnetic. Explain why we introduce the idea of field and compare and contrast the electric and magnetic fields. In your comparison, be certain to discuss at least one similarity and one difference.

Figure 4.15 A sample essay question on the field concept.

2. Remember that in exams students are under tremendous emotional pressure. If the exam is long so that they are also under time pressure, they may not have the opportunity to think about what they are writing.
3. The person who learns the most from essay questions on exams is the person grading them.

As a result of these points, I always try to keep exams short enough that the students are not overly rushed and I always grade my own essay questions, even if I have a large class and have TAs available for grading the other parts of the exam.