National Study of Education in Undergraduate Science
www.nseus.org

The Office of Research on Teaching in the Disciplines
The University of Alabama

Challenges and Possibilities of Meaningful Assessment in Large Lecture Introductory Physics

David Hammer

Tufts University

A Bit of Background in Research on Learning

My main purpose in this paper is to describe my own attempts as a practitioner to enact and promote meaningful practices of assessment — of students in my course and, more important, *by* students in my course. Before I do that, it will be helpful to back up and review the research behind those attempts. I'll begin with some discussion about the idea of *assessment* in science and science courses.

How Students Assess the Quality of Their Ideas and Understanding

I'd like to open with an example of data from my dissertation (Hammer, 1991, 1994), a study of students taking calculus-based introductory physics at Berkeley.

"Roger" and "Tony" were both students with outstanding scores and records in mathematics. I was interviewing them as part of my research on intuitive epistemologies: What, if any, beliefs did they show about knowledge and learning in the context of their course? I didn't ask them about their beliefs; I watched and listened to what they said and did over the semester. Roger and Tony were two of the six cases, and by coincidence it happened that they each made the same mistake in solving one of the problems in the set they had assigned that week.

The problem was to find the acceleration of the blocks, set up as shown, with the masses and angle of incline specified. Roger and Tony each correctly found an expression for the total force acting on the pair of blocks (treating it as a one-dimensional system), but they each applied that total to the blocks separately. This gave them different values of acceleration for the two blocks.

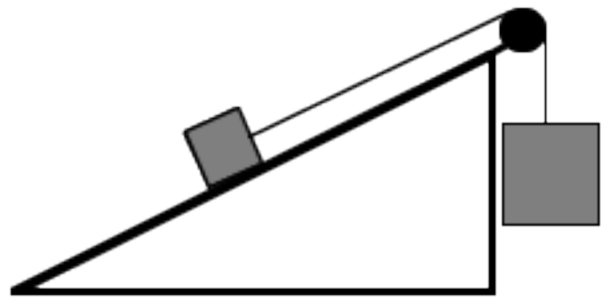


Fig. 1: Roger's and Tony's problem

What happened next showed a striking contrast in Roger's and Tony's epistemologies. Roger, seeing that his answer meant the two blocks would move at different speeds, said "Oh geez, how could one be accelerating faster than the other," and checked his algebra for mistakes. When he didn't find any, he decided he was probably right. His next step, told me, would be to "check and see if I got the right answer," using the answers in the back of the textbook, but he was "90% sure."

Tony, in contrast, immediately concluded that he'd "done something basically wrong," confident that "they have to be moving at the same speed." He reconsidered the expressions he'd written and decided that in writing them he had "said that this force was going to be [on the blocks separately]. "Now," he told me, "I'm saying that's not true." He wrote new expressions of this different conceptualization of the forces and accelerations and arrived at a correct solution.

My argument in my thesis was that this behavior fit the patterns I saw across their interviews. Roger, in doing his work for the course and in talking about what he was experiencing, showed a pattern that suggested he believed knowledge in physics is comprised of pieces of information, mainly located literally in the formulas, and that learning physics is a matter of receiving and retaining this information as provided by the instructor and textbook. Tony, in contrast, showed a pattern that suggested he believed knowledge in physics to be a coherent system of ideas, the ideas and coherence expressed in those formulas — "It's just a matter of putting common sense into equations" — and learning as centrally involving personal sense-making — "I usually tend to modify my common sense."

Years later, and specifically thanks to Janet Coffey (Coffey, 2003, 2005; Coffey, Hammer, Levin, & Grant, 2011), my former colleague at the University of Maryland, I came to see that much of what I was studying was, fundamentally, about *assessment*. Roger had seen that his answer implied the two masses would move at different speeds, but to find out if he had the right answer he would consult the authority of the textbook. Tony, seeing exactly the same thing, decided immediately he had done something wrong, based on a mismatch with what he took as solid conceptual ground, that the two masses had to move at the same speed. In this way, Tony was doing something recognizable as part of the discipline: in science, an answer is "true" if it coheres with established theory and evidence. Roger, on the other hand, checking the back of the book, was doing something recognizable only as part of school: an answer is correct if it matches the answer key. Thus the students' were enacting quite different practices of assessment.

There is little question that either student was capable of doing what the other was doing, given the right circumstances. At the time, I was focused on understanding students' epistemologies specifically in the context of their introductory physics course; in more recent work I and my colleagues have studied the contextual variations and dynamics of student epistemologies (Hammer & Elby, 2002; Hammer, Elby, Scherr, & Redish, 2005; Hutchison & Hammer, 2010; Scherr & Hammer, 2009). People can and do participate in a wide assortment of epistemic activities, ranging from some in which the right answer is determined by authority ("guess my middle name!") and others in which it depends on logical coherence ("is the set of prime numbers infinite?") or empirical ("see if cake's done"). Much depends on how any of us understands the situation we're in — how we "frame" it (Goffman, 1986; Redish, 2004; Tannen, 1993).

Unfortunately, a great deal of what students experience over the years in schools suggests they should be assessing the quality of their thinking in the ways Roger did, for alignment with the authoritative account rather than for the sense it makes given what they know and have seen. This includes undergraduate instruction, where there is evidence that students' expectations regarding knowledge and learning, including with respect to assessment, tend "to deteriorate rather than improve as a result of a semester of introductory physics" (Redish, Steinberg, & Saul, 1998).

How Educators Assess the Quality of Students' Ideas and Understanding

Much of what students experience as assessment, of course, is in the form of standardized tests, much more so in recent years than when most current college instructors were students. These, of course, enact assessment-by-answer-key in its purest form, and unfortunately the construction of these tests typically centers of know-it-or-don't terminology. In this respect, the tests are at odds with scientific rationality. Certainly, there are compelling reasons for standardized tests, and no doubt they vary in quality, but it is important to consider not only the information they provide but also the practices of assessment they model for students.

Something similar happens within class, including in the name of "formative assessment." Assessment practices that may be "effective" in serving institutional structures of schools can be at odds, even antithetical, to practices of assessment within science. Coffey *et al* (2011) review prominent examples from the literature on formative assessment to show their emphases on terminology over meaning (e.g. the term "mass" is correct; the term "weight" is not), their evaluation of correctness by authority, and their general lack of attention to the substance of students' reasoning within the students' own inquiries. Over the years, students' amass a great deal of experience of assessment practices based on alignment with authority rather than sense-making. It is not surprising they take up these practices themselves, that they become part of how they frame what takes place in science class.

For individual instructors, much of the challenge is the salience of "correctness" as judged against the familiar, established canon (Hammer, 1995). For a nice example, Scherr & Redish (2005) described a student, Melanie, who explained that a swimmer gets moving by "pushing off of the wall of the pool," reasoning that the force by the swimmer on the wall causes her acceleration. Melanie was clearly incorrect: From a Newtonian perspective, the swimmer's acceleration is related to the sum of the forces by other objects, including the wall, acting on her.

At the same time, the student's reasoning reflects aspects of disciplinary assessment that is important to encourage, of checking that ideas make sense — that they fit with other ideas and known phenomena — the sort of thing Tony was doing but Roger was not. To appreciate her thinking at that level is more difficult. Indeed, part of this challenge for the instructor is precisely the challenge instructors routinely expect students to meet, to understand a line of reasoning and consider its merits.

Someone might object, "But the Newtonian account makes sense, and hers does not." That, of course, is precisely what students experience, although in the opposite direction. It would be helpful for the instructor to articulate precisely what does not make sense in her reasoning, much as it would be helpful for the student to do the same with respect to the instructor's explanations. In this instance, genuinely to enter the student's thinking would be to appreciate its internal consistency.[1]

Challenges and Importance of Meaningful Assessment

There are, however, a number of challenges to meaningful assessment in university physics instruction. I've highlighted two:

- The practices of assessment students have experienced are largely at odds with scientific rationality, but they are what students have come to expect in science classes. Instructors who work to do something different can encounter resistance to change, especially among students who have learned to be "successful" in the system.

- It is much easier for instructors to recognize the consistency of an idea with respect to *our* understanding than it is to recognize it with respect to the students'. It is obvious, for example, when students reach conclusions we know to be correct; it is less obvious to follow the lines of reasoning by which they reached those conclusions.

The "Real World" of the Lecture Course

With that as introduction, the remainder of this paper focuses on practices of assessment in the context of large-lecture introductory physics. The particular example I'll use is the course I just finished teaching (as of this writing): Physics 11 at Tufts University, the calculus-based first

[1] In fact, Melanie's logic represents a perfectly viable alternative formulation of Newton's Second Law, were she to apply it systematically: "The sum of the forces exerted by a body is equal to its mass times the *negative* of its acceleration": $\mathbf{F} = -m\mathbf{a}$ (Scherr & Redish, 2005).

semester, which ended with 105 students. Here, I speak as an instructor informed by research on learning and teaching, in particular by arguments and evidence for attending to epistemologies, for understanding students as richly endowed with productive resources for learning, and for the importance of responsiveness in teaching.

The research suggests two principles for designing how we — the teaching assistants and I — assess students' work:

1) What we do should model and encourage scientific practices of assessment.

2) What we do should give us information about students' taking up those practices, as well as about their understanding of target concepts.

That is to say, assessment is a fundamental part of our objectives for students, part of their learning how to learn.

What makes things especially interesting is that, while these objectives are mutually dependent, in particular moments they are often in tension (Hammer, 1995). Indeed, this was the core import of misconceptions research, to help instructors recognize and appreciate the rationality of students' reasoning (Strike & Posner, 1992), even when it conflicts with established ideas; it is why so much of that work was based on comparison to historical views (McCloskey & Kargon, 1988). Students who have theoretical and empirical arguments to support their ideas are assessing in ways more like scientists than are students who check answers by authority.[2]

Regarding the first principle, it is essential to recognize that the assessment of ideas and understandings is something for course instructors as well as for students. Students need to do this assessment when they hear an idea, from me for from a fellow student: They need to assess "Have I understood what this person is saying?" as well as "How does it fit with other things I know?" The TAs and I have to do the same, when we hear students express their ideas, and Principle 1 says *we'd better model good behavior!*

Regarding the second principle, as course instructors we should be attentive not only to the "content" of what students say, in the traditional sense, but also to the ways students are taking up the pursuit of scientific understanding. How are they seeking coherence with respect to other knowledge and evidence they have of phenomena in the physical world (Sikorski, 2012)? Are they working toward mechanistic understanding (Russ, 2006; Russ, Coffey, Hammer, & Hutchison, 2009)?

For this paper, I'll focus mainly on lectures and exams, with some examples from this semester. See Redish and Hammer (2009) for more about problem sets and labs.

[2] It is unfortunate that in popular treatment, this tension has almost entirely resolved toward canonical correctness, and so "misconceptions" are widely depicted and experienced as impediments to learning science rather than as positive signs of sense-making.
[2] We use the "iclicker" system, but there are many choices.

Lectures

We use "clickers,"[3] and I organize lectures around "clicker questions." I do something similar to *Peer Instruction* (Crouch & Mazur, 2001; Mazur, 1997; Turpen & Finkelstein, 2009), but with more spontaneous interaction and improvisation. I prepare conceptual questions in advance, pose them to the students and poll the class, such as in Figure 2. Often students answer independently at first and what happens next can depend on the results. If there's broad agreement around the correct answer, I may just collect a few lines of argument for that answer and then agree with it. If there is significant disagreement, I may either have students call out arguments "someone might give" for different answers — "someone might give" to allow students rhetorical distance from the reasoning they articulate (Conlin, 2011) — or I might have students talk to people sitting around them.

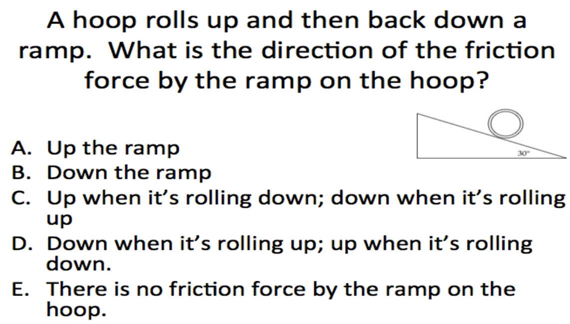


Figure 2: A clicker question
(including typo!)

I emphasize and try to model several things in these discussions. First is the importance of tangible clarity: I ask that students explain ideas in plain, simple language. *Can you explain the basic idea in terms an eighth-grader could understand?* This is a theme throughout the course, part of my advice for how they should study for exams, to be sure they can give that explanation and, more, be able to anticipate and respond to questions the eighth-grader might ask. It is also explicitly advice on assessment: If you can't explain the basic idea in simple language, you don't understand it.

As part of the discussions, I emphasize the importance of responding to counter-arguments. It is essential for students to *understand* the reasoning behind answers that are different from theirs, and it is essential to *respond* to that reasoning, to find the specific way that reasoning breaks down. This is also assessment: ideas should make sense to the point that we can reconcile apparent inconsistencies with other ideas. "The whole of science is nothing more than a refinement of everyday thinking" (Einstein, 1936), I remind them often, and the refinement is toward principled coherence.

I then typically re-poll, to find out how the discussions are influencing students' thinking, and from there I may follow with more discussion, or I may explain the "right answer."

Sometimes I withhold my take on the question, to keep the topic open for later, or because I believe students have attained their own clarity. It is an ongoing struggle for me to handle many students' need for the authoritative ruling, especially when the discussion or in some cases the demonstration has arrived at what I believe they could recognize as a clear answer.

What I am aiming for is something more like a conversation than a lecture punctuated with questions. I do offer some explanations of ideas, but mostly in response to student thinking I have heard, and I don't prepare any slides for presentation. The only slides I prepare in advance are questions, and lectures almost always open with one of them—such as the question in Figure 2, which was a follow-up to the problem set they were just handing in. It produced Figure 3, and the discussion that ensued lasted for 20 minutes, with multiple divergent arguments.

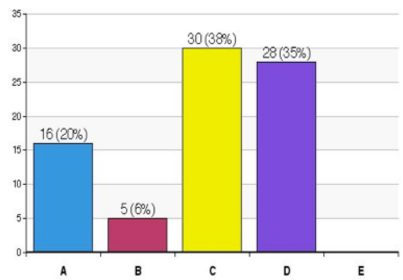


Figure 3: Polling results

Moreover, questions emerge during lecture, sometimes because the student thinking I hear makes me think of something new to ask, and often because a student's question seems like something to reflect back to the class. Question posing, including to oneself, is an important practice of assessment for them to learn, and I often solicit their participation in generating clicker questions, such as in Figure 4. In that instance, I had just posed a question asking them to compare the force by the horse on the buggy to the force by the buggy on the horse, and this was the next slide.

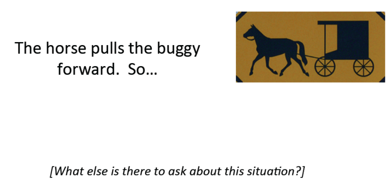


Figure 4: "What else is there to ask?"

By the end of the semester, this year and generally, the emergent questions "take over" many of the lectures, although I have not found a good way to distribute this work across students.

For the past two years I've been using *smartPhysics* (Gladding, Selen, & Steltzer, 2012), specifically the prelectures and checkpoint questions. These have been helpful in two ways. The "prelectures" are ~15-20 minute videos that are, essentially, traditional lectures, and in this way they serve the role of textbook reading. One benefit is that they have greatly reduced the number of complaints that I don't lecture! The checkpoints are more conceptual questions about that prelecture's ideas. The most powerful advantage to me is more assessment: I take an hour the morning before my lecture to review students' answers and comments on the prelectures and checkpoints, so that when I walk into the hall I already have information about what and how students are thinking, and I respond to it as part of what happens. A student's question on a *smartPhysics* checkpoint can easily become a clicker question in lecture that day. For me, *smartPhysics* is a compromise — I don't think the curriculum in itself promotes scientific assessment practices — but I like the way it helps the students and me interact. I don't use its homework assignments.

I also generally respond to 15-20 students by e-mail, to let them know I'm reading and paying attention to their work online, as well as to nudge them into better practices. The most common advice I give is that they work to articulate their confusion: Say precisely what it is that is confusing. It is an answer nobody can provide but the student.

In all these various ways, I design the lecture to model assessment practices reflecting disciplinary reasoning. Between clicker data and the arguments students express, I get a significant amount of information to help me assess how the class as a whole is doing, and the results inform what happens in that lecture as well as on problem sets and exams. The clicker points do collect information regarding individuals, but in practice I have made little use of that during the semester. That is something I would like to improve.

Exams

I don't think my exams are very innovative, but the following is how I describe them in the syllabus and advise students to study, much of which I discuss in lecture as well.

I try to write exams so that memorization without understanding doesn't succeed. The best way to be ready for my exams is to keep up with the course, staying on top of the ideas, asking questions to make sense of them, all along the way. Cramming to memorize equations the night before won't work.

So, here's what I advise students to do as they study, but really this is for all along the way in the course.

> "Use problems in the weekly assignments to help you discover gaps and confusions in your understanding—that's what they're for! Don't shy away from confusion—look for it, pin it down, and work it out. Don't just find a way to solve the problem; figure it out until you own it. Get so that you can explain what's going on in simple understandable language—I mean language that would be accessible to an intelligent 8th grader."
>
> "Be able to explain why other ways of solving the problem that lead to different answers don't work. It often happens in physics problems that one line of reasoning takes you in one direction and another takes you in a different direction. It's not enough to know which direction is right; you need to be able to explain why the other direction is wrong. For this in particular, it's important to work with other people, because they'll come up with ways of thinking you didn't."
>
> "Finally, you should be able to solve variations of that problem. Maybe in some of those variations one of those other lines of reasoning would be the way to go! So, pose yourself new questions—what if there were friction, what if the rock was moving up, what if the two cars had equal mass, whatever. That's a lot of how I come up with questions for exams: I look at problems we've solved, and I think of variations. Not, I should say, the same problem with new numerical values, but a variation of the problem that needs a variation of reasoning."

The first exam from this year is in the appendix, a mix of multiple choice questions like clickers in lecture and short-answer questions like the problem sets, although I take care to be sure they do not need calculators.

None of the course has students rehearsing "problem-solving" algorithms, and as I tell them I try not to have any question students could answer (or later think they could have answered) by memory. I explain in lecture that the world no longer needs people who are trained to perform routine calculations, and I hold up my smartphone to show them why not. Rather, we need people who can reason and invent, people who can do things these pervasive devices cannot.

I write the exams both to find out what they've understood and how they can reason, about material we've discussed at length and about new questions we have not, and I write them again with attention to what they "tell" students they should be doing to assess their understanding. The latter is especially important on the first exam — and this year as often the score was humbling for them and for me.

A 2 kg block sits on on a cart of mass 8 kg. The cart is accelerating to the right at 2 m/s^2, and the block accelerates with it, without slipping. The coefficient of static friction between the block and cart is μ_S=0.3.

What is the size and direction of the friction force on the block by the cart?

1. 6 N to the left
2. 6 N to the right
3. 4 N to the left
4. 4 N to the right
5. I don't know.

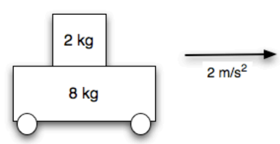


Figure 5: A clicker question that generated lots of discussion

This year, for example, we had spent a fair amount of time in lecture discussing the clicker question shown in Figure 5. During lecture, we modified this question to ask several others, including "what is the maximum force you could exert on the 8 kg cart without the 2 kg cart slipping off?"

Question 9 on the exam was another modification of this problem: "Suppose instead of pushing on the 8 kg cart you push on the 2 kg block. What would be the maximum force you in that case…?"

That turned out to be quite difficult: Almost all students answered 6N. That gave me information about how well they had understood the solution to the problem we had discussed (not very!).

To grade these questions, the first step is to examine the range of student answers. The TAs and I go through a sampling, typically 10 answers, writing comments and tentative scores on sticky notes. We then meet to go over the results, compare our scorings, and in this way we examine and discuss the data, to consider what it suggests about how things are going for the course as a whole and for individual students. Typically, we agree within a few points on almost all scores; divergences are moments to discuss differences in interpretation of what students did and/or differences in point values for particular accomplishments or errors.

I hope for an average of 65 on my exams, and if I achieve it, the grade cuts are 80 for an A, 70 for a B, and so on. I do this to give me the freedom to ask challenging questions as well as to limit the damage possible by a careless mind-slip or a student leaving a question blank. I never adjust the grade cuts up, but if I don't meet my target, I lower them. In this case, the average was 63, 2 points below my target.

On one question in every exam, question 11, I include a problem I think will be very difficult for students given what we've discussed so far. What I am mainly after in question 11 is whether students *see the problem*, something I emphasize in lectures and on problem sets I am also hoping to continue the emphasis from lecture that it is essential to find the flaw in a conflicting line of reasoning. These questions always include three parts: (a0 Explain what they think is the correct answer; (b) explain another compelling line of reasoning; and (c) respond to the reasoning in part (b).

In this case, the problem asked for the tension in the rope, in the situation shown in Figure 6. The person on the bench has a mass of 70 kg, the bench 10 kg, and the rope's mass is negligible. The question asked for the tension in the rope.


Fig 6: Exam question

For this problem, no student thought the answer was obviously 400 Newtons, the correct answer; a number of students thought it was obviously 800 Newtons. The worst answers give one obvious explanation and then a contorted alternative that is easy to dismiss in part C. My favorite answers are by students who change their minds: They write their first reasoning, genuinely consider another possibility, and end up choosing that one as the one they believe.

I'm not confident in any exam as a definitive measure of student learning, in part because over the years I've seen a number of cases in which it was clear that performance anxiety interfered with students I had seen often and gotten to know well. Much for that reason, I limit the exams to count for 50% of the grade. The remaining 50% comes from problem sets, scored primarily for "honest effort" (including sensible reasoning we can follow), labs (scored for clarity, insight, and innovation — see Redish & Hammer, 2009), and "participation points" from clickers, bonuses for catching my mistakes, and now, for *smartPhysics* prelectures and checkpoints.

Closing Note

One of the challenges of meaningful assessment, I always find, is that the results can be quite humbling to me as an instructor. Hardly any students were able to get Question 9 on my first exam, despite the time we spent on a different version of it. I find I have to steel myself to face the data without flinching, see what it says, and respond. Here, it meant going back to the basic ideas of Newton's laws and friction, as well as again to students' practices of assessing their understanding. (In fact, this result led the TAs and me to redesign the next laboratory to allow students opportunities to design experiments around aspects of friction that continued to puzzle them, including its role in rolling as well as the exam question.)

And, no question, teaching is humbling to me as a researcher. There's just no escaping that the existing research is not sufficient and will not be for a good while. To be sure, that research points to an account of framing in terms of a complex system, with conceptual, social, material, institutional, affective, epistemological aspects in non-linear interactions. Much as it can difficult to predict the weather with precision, it can be difficult to predict particular classroom dynamics, which is part of the argument for the importance of adaptive, responsive teaching. Still, it is striking to me teaching introductory physics, or anything else for that matter, how much more work there is to do in understanding learning and instruction.

But this is not remotely an argument against education research! It is an argument for it, and for recognizing it as a long-term endeavor. Nobody doubts the importance of research on mechanisms of the immune system, for example, despite the clear fact that what it has produced thus far is not remotely sufficient to prescribe optimal practices for health; society expects progress over decades. It has been difficult, in contrast, to achieve commitment to research on learning and instruction as a pursuit for the long term. Of course we should do what we can to understand education in ways that we can put to immediate use, but at the same time we should appreciate the depth and complexity of the phenomena.

Author Biography

David Hammer's research has focused on the learning and teaching of science (mainly physics) from elementary school through university, with particular emphases on students' intuitive epistemologies, how instructors interpret and respond to student thinking, and resource-based models of knowledge and reasoning. From 1998-2010, he held a joint position in Physics and Curriculum and Instruction at the University of Maryland, College Park; he is now Professor of Education and Physics, and Co-Director of the Center for Engineering Education and Outreach, at Tufts University.

The dissemination of this material was partially supported by the National Science Foundation under Grant No. 0554594. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Coffey, J. E. (2003). Involving students in assessment. In J. M. Atkin & J. E. Coffey (Eds.), *Everyday Assessment in the Science Classroom* (pp. 75-88). Arlington, Va.: NSTApress.

Coffey, J. E. (2005). Classroom Assessment in the Service of Student Learning. *Teaching Science in the 21st Century*, *39.*

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching*, 48(10), 1109-1136.

Conlin, L. D. (2011). *Builing shared understandings in introductory physics tutorials through risk, repair, conflict & comedy.* Doctoral dissertation, University of Maryland, College Park.

Crouch, C. H., & Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics, 69*, 970.

Einstein, A. (1936). Physics and reality. *Journal of the Franklin Institute*, 221(3), 349-382.

Goffman, E. (1986). *Frame analysis : an essay on the organization of experience*. Boston: Northeastern University Press.

Hammer, D. (1991). *Defying common sense: Epistemological beliefs in an introductory physics course*. University of California at Berkeley.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction, 12*(2), 151-183.

Hammer, D. (1995). Student inquiry in a physics class discussion. *Cognition and Instruction, 13*(3), 401-430.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal Epistemology: The Psychology of Beliefs about Knowledge and Knowing* (pp. 169-190). Mahwah, N.J.: Lawrence Erlbaum.

Hammer, D., Elby, A., Scherr, R. E., & Redish, E. F. (2005). Resources, framing, and transfer. In J. Mestre (Ed.), *Transfer of Learning from a Modern Multidisciplinary Perspective* (pp. 89-119). Greenwich, CT: Information Age Publishing.

Hutchison, P., & Hammer, D. (2010). Attending to Student Epistemological Framing in a Science Classroom. *Science Education, 94*(3), 506-524.

Mazur, E. (1997). *Peer Instruction*. Upper Saddle River, NJ: Prentice Hall.

McCloskey, M., & Kargon, R. (1988). The meaning and use of historical models in the study of intuitive physics. In S. Strauss (Ed.), *Ontogeny, Phylogeny, and Historical Development* (pp. 49-67). Norwood, NJ: Ablex.

Redish, E. F. (2004). A theoretical framework for physics education research: Modeling student thinking. In E. Redish & M. Vicentini (Eds.), *Proceedings of the Enrico Fermi Summer School, Course CLVI* (pp. 1-63). Bologna, Itally: Italian Physical Society.

Redish, E. F., & Hammer, D. (2009). Reinventing college physics for biologists: Explicating an epistemological curriculum. *American Journal of Physics, 77*(7), 629-642.

Redish, E. F., Steinberg, R. N., & Saul, J. M. (1998). Student expectations in introductory physics. *American Journal of Physics, 66*(3), 212-224.

Russ, R. S. (2006). *A framework for recognizing mechanistic reasoning in student scientific inquiry.* Doctoral dissertation, University of Maryland, College Park.

Russ, R. S., Coffey, J. E., Hammer, D., & Hutchison, P. (2009). Making Classroom Assessment More Accountable to Scientific Reasoning: A Case for Attending to Mechanistic Thinking. *Science Education, 93*(5), 875-891.

Scherr, R. E., & Hammer, D. (2009). Student Behavior and Epistemological Framing: Examples from Collaborative Active-Learning Activities in Physics. *Cognition and Instruction, 27*(2), 147-174.

Scherr, R. E., & Redish, E. F. (2005). Newton's Zeroth Law: Learning from Listening to Our Students. *The Physics Teacher, 43*, 41.

Sikorski, T.-R. (2012). *Developing an alternative perspective on coherence seeking in science classrooms.* Doctoral dissertation, University of Maryland, College Park.

Strike, K. A., & Posner, G. J. (1992). A revisionist theory of conceptual change. In R. A. Duschl & R. J. Hamilton (Eds.), *Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice* (pp. 147-176). Albany: State University of New York Press.

Tannen, D. (1993). *Framing in discourse*. New York: Oxford University Press.

Turpen, C., & Finkelstein, N. D. (2009). Not all interactive engagement is the same: Variations in physics professors' implementation of Peer Instruction. *Physical Review Special Topics-Physics Education Research, 5*(2), 020101.

PHYSICS 11 FIRST EXAM

Spring 2012, David Hammer

Answer all questions on these sheets. Please write clearly and neatly; we can only give you credit for we can read. **We need your name and recitation section number on every page, because we will separate the pages for grading.**

The first set of questions are multiple choice. There's no partial credit for these – just choose the best answer and indicate it clearly.

The second set are "short answer," and they all require explanations, whether the word "explain" appears in the question or not! Your explanation may be in words, mathematics, and/or diagrams, but we need to be able to follow it. Full credit is for a correct answer with a clear explanation. You'll get partial credit for sensible reasoning, even if the answer is incorrect. **<u>But you'll get no credit, even if your answer is correct, if we can't follow your reasoning.</u>**

THIS IS A CLOSED BOOK, CLOSED NOTES EXAM. NO CALCULATORS OR ANYTHING WITH A SCREEN ALLOWED!

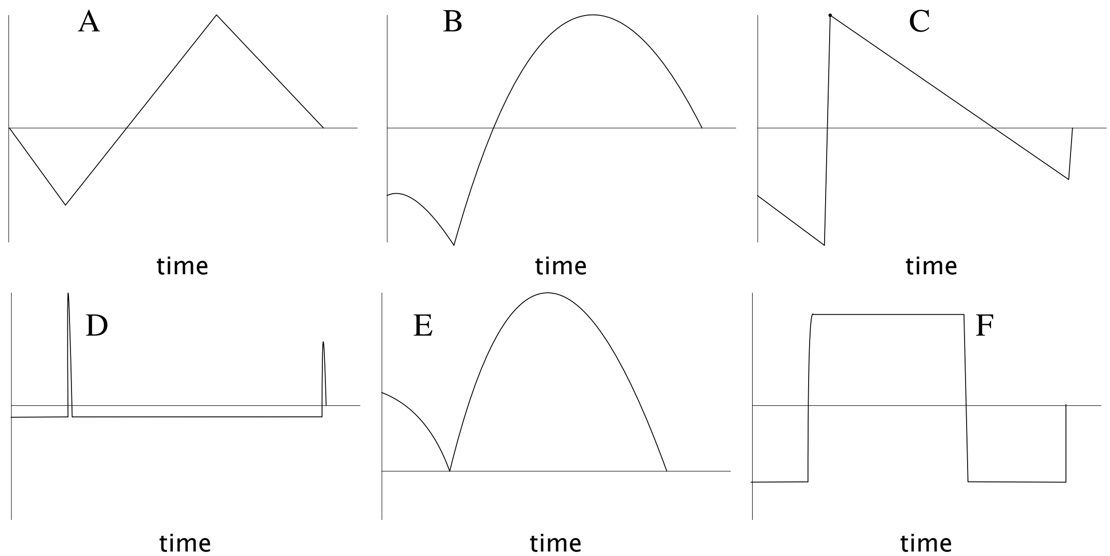
Finally, please: Think about the physics, not about the psychology of how I write exams!

____

Name (printed)

Multiple choice questions (5 each, 40 total)

Just the answer counts for these; no partial credit.



1) You throw a ball at the floor, let it bounce up over your head, and then catch it as it comes down, at about the same height from which you let it go. Which of the graphs above would best depict the vertical component of the velocity of the ball (dh/dt), from when it leaves your hand to when you catch it?

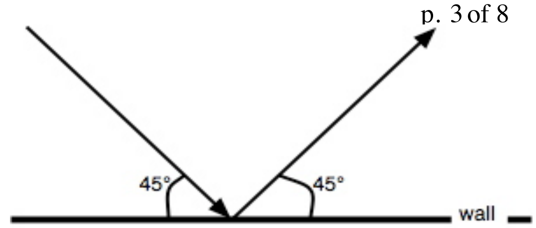
2) Same situation, but now pick the graph that would best depict the sum of all the forces acting on the ball, from when it leaves your hand to when you catch it?

3. You're in a glass-walled elevator, and it is moving *down* at a constant speed of 2 m/s. To pass the time, you bounce a ball off the floor: You throw the ball down, at a speed of 4 m/s, from a height of 1 meter, it bounces off the floor, and you catch it as it comes back up. From your perspective, the ball moves 1 meter down, and it takes ¼ second to hit the floor.

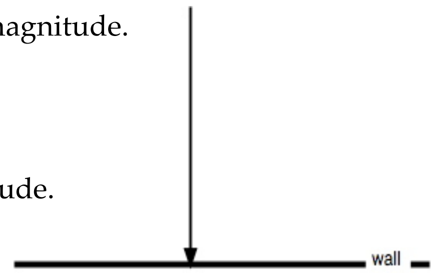
   To someone watching from outside, how far does the ball move down, and how long does it take to hit the elevator floor?

   A. 1 meter, 0.45 seconds
   B. 0.67 meters, 0.1 seconds.
   C. 0.67 meters, 0.25 seconds.
   D. 1.5 meters, 0.45 seconds.
   E. 1.5 meters, 0.25 seconds.

4) I have a rubber ball and a lump of clay, and I throw them each at a wall with the same speed, but at different angles as shown. The ball bounces off the wall at an angle of 45°; the clay I throw straight at the wall where it just sticks. How does the change in velocity of the ball $\Delta\mathbf{v}_b$ compare with the change in velocity of the lump of clay $\Delta\mathbf{v}_c$?



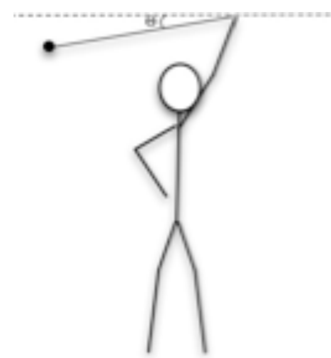
The ball bounces at a 45° angle



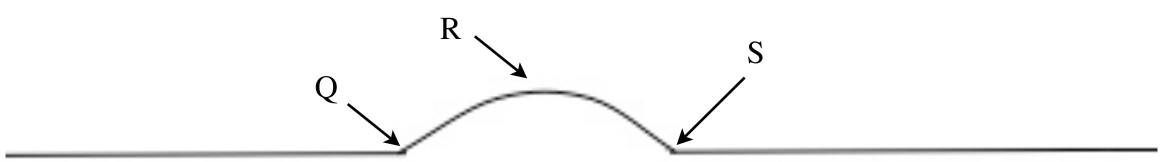
The clay just sticks

A. They are in different directions, but they have the same magnitude.
B. They are in different directions, and $\Delta\mathbf{v}_b$ is larger.
C. They are in different directions, but $\Delta\mathbf{v}_b$ is smaller.
D. They are in the same direction and have the same magnitude.
E. They are in the same direction, but $\Delta\mathbf{v}_b$ is larger.
F. They are in the same direction, but $\Delta\mathbf{v}_b$ is smaller.

5) I swing a ball from a string around my head. What forces act on the ball? There's the gravitational pull by the earth, and there's also the



A. force inward by the string.
B. force outward by the ball's motion.
C. force inward by the ball's motion.
D. force of the ball's motion, in the direction it is moving.
E. More than one of a - d.



6) You're driving along when you come to a little rise in the road — a small hill (makes more sense than a "hump"). Suppose the car moves left to right, in this picture. So point Q is at the very start of the hill; point R is at the top of the hill, and point S is at the end of the hill. F_s is the magnitude of the force by the car seat on you; and mg is your weight. At which points, if any, is $F_s > mg$?

A. At point Q only.
B. At points Q and R.
C. At points Q and S.
D. At point R only.
E. At point S only.
F. At no points Q, R, or S.

7) Meet Xena and Yara. Xena's mass is half of Yara's mass ($m_X = ½\ m_Y$). They are standing on smooth ice. Xena holds a rope that's tied around Yara's waist. For this problem, we'll assume the rope has negligible mass, and the friction between either of them and the ice is negligible too.

They're standing still until Xena gives the rope a tug. As a result…

A. Yara starts moving toward Xena; Xena doesn't move.

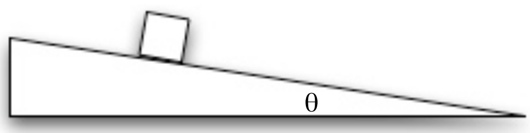
B. Xena starts moving toward Yara; Yara doesn't move.

C. They start moving toward each other; Yara moves twice as fast as Xena: $v_Y = -2\ v_X$.

D. They start moving toward each other at the same speed: $v_Y = -\ v_X$.

E. They start moving toward each other; Yara moves half as fast as Xena: and $v_Y = -\ ½\ v_X$.

8) A block of mass m sits on a ramp, inclined at angle θ. Which of the following would you expect to be true, of the magnitude of the friction force F_f?



A. $F_f = mg\sin\theta$

B. $F_f = \mu_k mg\cos\theta$

C. $F_f = 0$

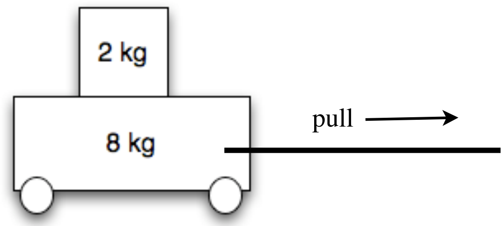
D. Both A and B

E. None of these

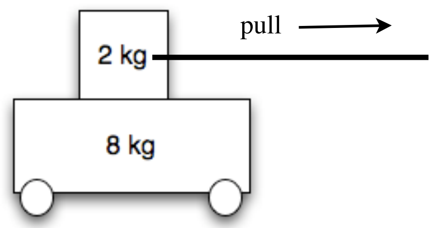
Short answer questions

For these you **must** explain; no credit without explanation!
(And the clarity of your explanation counts.)

9) Back to this situation: A 2 kg block sits on on a cart of mass 8 kg. The coefficient of static friction between the block and cart is μ_S=0.3. In lecture we found that the maximum force you could exert on the 8 kg cart without the 2 kg block slipping was 30 N. For example, you could pull it with a rope, as I've shown.



Suppose instead of pushing on the 8 kg cart, you push on the 2 kg block. For example, you could pull it with a rope, as I've shown.



What would be the maximum force in that case, that you could exert on the 2 kg block without it slipping? As before, assume there's no friction between the 8 kg cart and the surface it's on.

10) We've derived a bunch of expressions for constant acceleration, including $x(t) = x_0 + v_0 t + \frac{1}{2}at^2$ and $v_f = v_i + at$. So we know those formulas work when a is constant.

If a isn't constant, can we still use them, as long as we substitute the *average acceleration* in for a? That is, suppose we know the average acceleration $\overline{a}(t)$ of, say, a car, from 0 to t. If we use that value in the equation $x(t) = x_0 + v_0 t + \frac{1}{2}at^2$, would we get the right answer? How about if we used it in $v_f = v_i + at$? Give a convincing argument you can or you can't.

11) The picture shows Hanzo sitting on a bench that's attached to a rope. The rope goes through a pulley; the pulley is attached to the ceiling. (In case it's helpful, the second picture is a less artistic diagram of the same situation.)

The bench's mass is 10 kg; Hanzo's mass is 70 kg, and the rope's mass is very small so we'll ignore it. And assume there's negligible friction in the pulley. If Hanzo is stationary, what is the magnitude of the force between him and the rope?

a) Give a sensible argument for the answer you believe.

b) Give a sensible argument for an answer you think someone else might believe.

c) Try to explain the flaw in the reasoning for part b.