2005 National AAPT Meeting – Winter
Notes from January 10-12
Albuquerque, NM

The following notes were taken by Dr. Van Domelen during the regular sessions of the 130th National Meeting of the American Association of Physics Teachers. One of the themes of this meeting was Einstein's "Annus Mirabilis" and the International Year of Physics.

Abbreviations: There are a few concepts that came up frequently enough in the talks I attended that I gave them their own abbreviations. (Not all will appear this time; I just pasted in the 2003 glossary and added to it as necessary.)

CSEM – Conceptual Survey of Electricity and Magnetism
ECR – Elicit, Confront, Resolve. The UWash method.
FCI – Force Concepts Inventory, a common mechanics test.
FMCE – Force and Motion Concept Exam, a common mechanics test. I may sometimes type it as FCME.
GG – Gender Gap, a disparity between male and female results, usually to the detriment of the women.
IE – Interactive Engagement, one of many strategies that get students more involved in their own learning.
ISLE – Investigative Science Learning Environment
MPEX – Maryland Physics Expectations Survey, a test of student expectations about physics and physics courses. MPEX2 is a recent revision.
N1L, N2L, N3L – Newton's First Law, Second, Third.
PRS – Personal Response System, a means of letting students answer multiple choice questions in class using a remote control.
PS – Problem-Solving
STEM – Science, Technology, Engineering and Math. Majors typically targeted by calculus-based physics courses.

So that I don't need to say this every time it happens, I apologize in advance if I misrepresented anyone's arguments. A lot of talks went too quickly to take accurate notes, averaging a slide every 30 seconds or so. This was particularly hard on me with the invited talks, because I didn't get a break to rest my writing hand as often. Heh.

Seriously, if you're giving a talk at one of these meetings, I'd strongly suggest practicing with someone trying to take notes. If they can't keep up, then you're going too fast. We all know the limited effectiveness of the lecture format, and it's even less effective if material is presented too quickly for anyone to process it. This message has been brought to you by Futura Wrist Braces.

Monday, January 10, 2005

BE – Student Conceptions and Einstein's Annus Mirabilis (Invited)

BE01: Einsteinian Thinking Among Ordinary Students: Examples from Special Relativity – Rachel Scherr, University of Maryland.

Rather than looking to see if students have become "Newtonian thinkers," the goal here was to consider whether any of them were on the road to becoming "Einsteinian thinkers".

A Newtonian thinker:

- Believes in and uses Newton's 0th through 3rd Laws (The 0th Law is a recent proposal that states, essentially, that an object only reacts to the forces currently acting on it, and does not "remember" previous forces).
- Believes in absolute spacetime "provided by the universe".
- Believes in a mechanistic universe.
- Believes "the truth is out there" and can be deduced from observations.

Meanwhile, an Einsteinian conceptual framework involves non-absolute quantities defined by us (rather than provided by the universe), and includes:

- Frame equivalence
- Isotropy
- Constancy of the speed of light
- Relativity of space and time.

A correct conceptual framework should practice logical economy, being as simple as possible but no simpler.

Common student non-Einsteinian beliefs include:

- Simultaneity is absolute.
- If signals reach the observer at the same time, the events are absolutely simultaneous.
- Each observer is a unique frame, rather than being in the same frame as any other element of the system or other observer.

The first topic investigated was the "Bubbles" question. Given two observers, Alan and Beth, with Beth moving away from Alan, an expanding wavefront is centered on Alan. Students were asked to describe things in Alan's frame, such as the shape of the expanding wavefront. Then the situation is shifted to Beth's frame, with Alan moving away and still generating waves, and the questions are asked again.

Students did display some good Einsteinian thinking habits:

- They value symmetry, and use it almost as second nature.
- They treat conceptual oddities raised by the situation (i.e. is there just one bubble or two different ones?) as curiosities, not obstacles.
- Isotropy is assumed (bubbles consistently drawn as circular, not lumpy).

- They easily apply the invariance of c.

However, the questions themselves were easy (for relativity, anyway), and may not demonstrate a robust student understanding. They may be useful for scaffolding, however.

The abstraction of the setup (stick figures and circles) did remove excess information that might have hurt learning, notably the lack of eyes on the figures removing a cue for reception of light signals. By distilling it down, the situation was made easier to comprehend.

However, despite doing well on this item, students still did poorly on exam questions and other standard tasks. It is believed that this is because the basic concepts are both hard to activate without prodding, and because they conflict with Newtonian (or Aristotelian) views.

The second topic was the "Beeper" question. Alan is some distance away from a beeping box. He has a ruler, a timer, and helpers to take measurements for him. The goal is to figure out a way to determine at what time the box beeps, both using knowledge of the speed of sound and not using that information.

Students got hung up on the obviousness of the problem and spent a lot of time looking for the "trick".

Students were given a historical equivalent of the problem, the issue of train scheduling in Europe about 1900 and the choice of a "master clock". Each town had its own master clock, and the time at an outlying location would be the moment when the signal (bell ringing) from the master clock reached it. While this small variance in times wasn't relevant to train schedules, it did get some scientists thinking. A similar modern situation would be watches that synchronize with the atomic clock's radio signal out of Colorado, and are actually "out of synch" by the fraction of a second it takes the signal to travel from Colorado to one's location.

In solving the Beeper problem, students correct for signal travel time and use Einsteinian time (all locations assigned the same time), which is another building block of Einsteinian thinking. However, context and question priming may affect these results.

Conclusion 1: Students can think in an Einsteinian sense, at least in response to certain questions. This is very fragile, however.

Conclusion 2: We have much to learn from attending to what students do **correctly**, as well as where they have difficulty.

BE02: The Development of Student Understanding of the Photoelectric Effect – Graham Oberem, CSU San Marcos (with Charles DeLeone).

The photoelectric effect is interesting because it forces consideration of the particle model of light, requiring students simultaneously hold two or more models for the same phenomenon at once. (The "or more" comes from the existence of semi-classical models using purely wave light coming from a quantized atom.)

The class investigated in this study was a low-level modern physics course taken by Physics majors right after they had completed the introductory year of calculus-based physics. 2 labs and 2 lectures covered the photoelectric effect, using a format of "pretest

– small lecture – tutorial – small lecture – lab – posttest". Students were taped during labs, and the results transcribed. The tutorial was online, and the computer system saved records of all responses.

The pretest covered the classical model, the actual circuitry of the apparatus (using free electrons without involving how they were generated) and then the photon model after the first parts were handed in. Scores were very low, even on the classical model.

The computer tutorial used simplified circuit diagrams and asked students to graph Voltage versus Current. Students were taped during tutorial work, so that their non-typed statements could be matched up with what the computer saved. No one got the problem right on the first try, the average number of attempts was 6. Only 3 of the 18 students availed themselves of the circuit review linked from the tutorial.

Common mistakes included graphing an "ohmic vee" (I proportional to |V|), invoking the electric field as the source of the free electrons, invoking the electric field as the source of current rather than the thing that stops current.

The tutorial led them through a series of steps to finally get the correct graph.

1) Recognize that the current is always positive, changing an $I \alpha V$ graph to the ohmic vee.

2) Recognize that the current saturates at high magnitude of voltage, changing the ohmic vee to a graph that has asymptotes of I_{max}.

3) Recognize that the external E field is actually stopping the current, so there will be a negative voltage below which there is no current at all.

Practice problems ask leading questions to step the students along this path.

Once students realize that there's a frequency dependence involved, they tend to abandon intensity dependence briefly and focus only on the new concept. Eventually they merge the two concepts, with stopping voltage dependent on frequency, and saturation current on intensity.

The posttest results were generally pretty good, but students would at this point avoid classical responses even when the classical model was correct in that case. Unknown at this time if they're simply enamored of the new model, or if they've undergone a paradigm shift and no longer really even remember the old model.

The instructional materials were deemed effective, and the computer tutorials judged favorably.

It is possible that the students lacked a good grasp of how circuits worked, and it's clear that most did not have a good grasp of classical wave theory.

Direction for future work: do students totally discard the wave model after learning the particle model?

BE03: What Counts as Convincing? Examples From Physics and Physics Education – Nicole Gillespie, Knowles Science Teaching Foundation.

Historical Background: Consider Einstein (1905) regarding Brownian Motion, and the acceptance of the atomic model at that time.

In the 19th Century, Phenomenologists dominated the sciences. They disliked the kinetic theory of gases...equations existed that described what happened, why ponder the unknowable interior processes? Boltzman decided to ponder anyway, and he hit upon statistical techniques that the Phenomenologists thought to be unscientific mathematical hand-waving.

Einstein's 1905 paper made testable predictions about statistical mechanics, however, and went a long way towards convincing the scientific community that the theory of atoms was valid. Boltzman had the idea himself in 1896, but even HE wasn't convinced it was right, and others worked on the principles between 1896 and 1905.

So why was Einstein's proof accepted in 1905, when the previous efforts had not been? Was it merely a better proof? Had the standards finally shifted enough that it didn't have to be better, just in the public eye?

Moving on to students. What constitutes convincing proofs, evidence and arguments about the behavior of gases to a student?

The students in question were in a reformed course (CCLI A&I grant) involving "discussion/lab" studio sections twice a week and lectures three times a week. Algebra-based, the main clients were Biology/Pre-Med and Architecture majors.

Thermodynamics was studied for 3 weeks, building on mechanics/momentum/ energy concepts in order to attempt to deal with the Heat/Temperature and Heat/Work conflations.

96 hours of video were analyzed, asking the following questions: What discursive resources were used? What status did those resources have? What did statements **mean** to the students?

Example: Why does 40°F water feel colder than 40°F air? Statements ranged from "coldness is more concentrated in water" to "there are more molecules in water to collide with" to "water is denser, so molecules can't move as much."

One case study was presented, the compression of gas in a piston. 4 episodes following the progress of a group of students were presented, focusing on three topics.

1. Temperature – students knew what changes would happen, but weren't sure why. They often referred back to classroom demos and definitions.
2. Energy – this was invoked frequently. Initially, it was invoked regarding the addition of energy to a system, but later on they merely used energy equations. They wouldn't poke at the definitions.
3. Heat/Temperature conflation – initially students used definitions to simply state they weren't the same thing, but this didn't stick. Later on, the 1st Law of Thermodynamics helped this point get across.

Students knew the atomic model, but didn't really hold it accountable for explanations.

"Convincing" is local, not global. Argument from authority is usually successful in the short term, with "authority" including equations. Intuition was only seen as convincing if it agreed with authority.

Social elements are significant in answering the question, "What is convincing?"

BM – TIPERS (Tools Inspired by PER - Invited)

BM01: Where's the PER Inspiration in TIPERs? – David Maloney, IUPUI/Purdue.

Students need to work with ideas several times, in a variety of ways, to make the ideas their own. Especially if the students come in with alternative conflicting ideas. But even without conflicts, it can take a while to get comfortable with a new idea.

Some examples of TIPERs gleaned from a rapid flip through slides:

- What Is Wrong? Tasks – Students are presented with a physical situation in which one or more critical aspects are incorrect, such as a mathematical representation of a sum of forces that neglects to find the x and y components. Note: there will be the occasional task with no errors tossed in to keep them on their toes.
- Ranking Tasks – One of the most commonly-used TIPERs, this asks students to consider a group of situations and rank them from most to least, strongest to weakest, etc.
- Predict and Explain Tasks
- Energy Bar Charts – Using the representation developed by Alan Van Heuvelen to describe where the energy goes during a process.
- Conflicting Contentions Tasks – As seen in the PBI manuals, two or more student arguments are presented (preferably ones based on actual student interview responses or the results of Ranking Tasks), and users are asked to decide who, if anyone, is correct.

Here are some student difficulties identified by PER on the topic of electricity and magnetism.

- Magnetic poles are thought to be electrically charged...poles is poles.
- Electric and magnetic flux are both abstract and difficult to grasp.
- The vector nature of forces and fields.
- The fact that magnetic fields pretty much force a three-dimensional consideration of events means that students who have trouble thinking three-dimensionally have extra difficulty with B.
- Field and potential are often seen as equivalent, similar to problems with speed and acceleration in mechanics.

There are four types of inspiration for TIPERs:

1. Idea of repeated use in different ways.
2. Use task formats from actual research.
3. Use task formats from actual curriculum development.
4. Focus on PER-identified difficulties.

BM02: TIPERs: An Overview – Curtis Hieggelke, Joliet Junior College.

The strategy for developing TIPERs includes the following points:

- Base it on research into how students learn and think.
- Ease of use for both students and instructors is important.

- Make it usable in bite-sized chunks, so that it can be introduced gradually and less stressfully.
- The use of multiple formats helps at both the front and back ends, and makes things more robust.
- The task should allow students to practice working and thinking, not just rote exercise.
- Provide a conceptual base for problem-solving and understanding.
- Adopt a curriculum-independent approach as much as possible, so that the TIPERs are usable by a wide range of instructors..

TIPERs are implanted in the following ways:

- Embedded in lecture, with pre/post review.
- Group work exercises, including class polling and discussion.
- Laboratory work.
- Online work (such as Just In Time Teaching).
- Traditional homework.
- Formative and Summative assessment tools, notably investigating student models and ideas.

mTIPERs (magnetism-specific) cover the effects of **B**, how **B** is produced, and the relationship between **E** and **B**. There are 248 mTIPERs plus a Magnetic Assessment Tool (MAT) of 40 items, to be published in Fall 2005 through Prentice-Hall. Some of the tasks are single-issue, some are basic set tasks. The materials try to cover the most commonly presented topics, avoiding idiosyncratic topics that are only taught by a small minority of instructors.

eTIPERs (electric-specific) cover charge, charged objects, static electric fields and electric potential. The MEAT (Mechanics Electrostatics Assessment Tool) was developed to probe student understanding in this area. More on MEAT in the next talk. The CSEM was also used.

Here's a number of other TIPERs:

- Working Backwards Tasks – Also known as Jeopardy Problems, students are given a mathematical or graphical representation and asked to devise a problem that would result in that answer. These are grouped on their own rather than melded into the topic groups, to avoid tipping students off as to the topic of the task.
- Troubleshooting Tasks – Related to What's Wrong? Tasks, but in this case they're told there is an error and asked for fix it.
- Linked Multiple Choice Tasks – A single situation is used for several tasks, but with each question changing one or more things (i.e. now assume the charge is negative, what happens now?).
- Changing Representations Tasks – Students are explicitly asked to restate a problem in different representations. For instance, given a word problem, to show a diagrammatic representation, or a mathematical one.

- Meaningful/Meaningless Calculations Tasks – Given an equation and a situation, evaluate whether the equation tells you anything meaningful about the situation. Also tries to identify excess information.
- Qualitative Reasoning Tasks – Very closely related to the Linked Multiple Choice Tasks, but more emphasis on qualitative (where LCMTs may ask quantitative questions).
- Comparison Tasks – Given two different situations, how does some quantity differ?

A lot of data was collected using the CSEM, national results had a weak normalized gain of 0.29 (the CSEM normally has higher gains than the FCI, because students' pretest knowledge is much weaker...0.29 counts as weak on the CSEM even though it's more moderate to strong on the FCI).

Issues identified as problems for students included three-dimensionality, students approaching the topic as a disconnected pile of rules, lack of experience with non-permanent magnets, and a bias about the nature of poles.

When mTIPERs were introduced to an already-reformed class that had been performing around the national average, gains on the CSEM jumped to the 0.46 to 0.67 range (depending on section), with strong statistical significance and an effect size of 1.03. mTIPERs were concluded to work.

BM03: Qualitative Tasks for Introductory Electrostatics – Stephen Kanim, New Mexico State University.

When developing worksheets for the purpose of scaffolding students on traditional archetypical electrostatics problems, "Trivial" underlying issues turned out to be important after all. The "easy" material can be harder for students than we think. This inspired the development of eTIPERs.

Along the way, it was decided that an FCI-level instrument for electrostatics was necessary, and that the CSEM didn't meet their needs. Taking as a basic principle the idea that an effective E&M course reinforces and extends mechanics concepts, the MEAT was developed.

The MEAT has different versions for pretest and posttest. The pretest is solely concerned with mechanics principles that will be important in the E&M course, thus avoiding the problem of pretesting on topics about which you expect students to know next to nothing. The posttest includes the pretest problems, plus adds E&M items that are paired with their pretest counterparts. So, for instance, a central forces gravity problem on the pretest would be paired with a central forces electric force problem.

In the posttest, there was only small improvement on the mechanics questions, suggesting that the "reinforce" part of the course wasn't working as well as hoped. Student understanding of Work/Energy concepts started weak and stayed weak.

When comparing the paired items on the posttest, transfer was not so good. Students consistently did worse on the E&M items than on their related mechanics problems, sometimes much worse. The average score on the mechanics questions was about thirty percentage points higher than the average score on the E&M questions.

BM04: Skipped due to pain in arms and legs from sitting in uncomfortable chairs and taking notes at high speed all afternoon. Sorry.

BP – Plenary 1 – AIP's Gemant Award Lecture

BP01: Physics Influences Art: Evidence in the Surreal Painting of Remedios Varo – Alan Friedman, New York Hall of Science.

Varo was a Spanish exile working in Mexico City from 1942 until her death in 1963. She had a succession of husbands, mostly artists (which helped her get seen in the Surrealist scene, which was very much a men's club), but her last husband was a fellow "fleeing from the Nazis" sort who made a lot of money in the music industry and let her concentrate on her own art, rather than copying the style of another artist just to get noticed. She was inspired by science popularizer Fred Hoyle.

The mainstream critical line on Varo's work was that it was mostly influenced by mysticism, but Friedman advances the theory that much of her "mystic" work was actually attempts to somehow represent the scientific theories she was reading about in Hoyle's work. (Personal interjection: I found a lot of Friedman's evidence to be of the "fits the theory if you want it to fit the theory" variety, but enough seemed solid to lend at least plausibility to his interpretation.)

Friedman's first exposure to Varo was a mention in Pynchon's <u>The Crying of Lot 49</u>, which he read in the early 70s. But when he tried to find this mysterious artist, all evidence seemed to have vanished, with the few listed catalogs of her work having been checked out of libraries and never returned! It took him until 1982 to finally track down actual pictures of her work. In 1986, the New York Academy of Sciences was the first to display Varo's work in the United States.

Varo's paintings tended to be very small, and she used a brush with only a single hair on it to do much of the painting. This gave her a great deal of control (and avoided the issue of brush-clogging), but of course meant that even a small painting took a long time. Most of her earlier work was simply copying the style of her husband at the time, seeking acceptance in the boys' club that was Surrealism, but once she moved to Mexico her own style started to emerge. She did some commercial art work while her husband was getting his business started, the results of which gave some hints of what was to come. But it wasn't until later in life, when she got interested in Hoyle's writings and her husband's wealth let her paint what she wanted, that her talent really bloomed.

Her later works often included imagery linked to cosmology, such as frequent use of stardust, inclusion of odd uses of coordinate axes, or a painting ("Exploration of the Source of the Orinoco River", 1959) that could be seen as a visualization of steady state cosmology, a river kept ever full by a tiny source hidden in a tree. Music was often used to represent an organizing principle or theory. In general, when she depicted scientific themes, it was with a feeling of wonder and enrapturement, unlike the terror evidenced by many of her contemporaries.

Of course, the mystical angles can't be completely ignored. Varo didn't see science as the only possible way to interpret the world, but it can be argued that she felt it was a particularly **good** way to do so. The belief that the universe must, at its root, make sense was attractive to her.

For more on Varo, see Janet Kaplan's biography, a 2002 catalog of her works edited by Walter Gruen, or http://www.hungryflower.com/leorem/varo.html (which I Googled for just now).

Tuesday, January 11, 2005

CC – Student Conceptions in Conceptual Physics (invited)

CC01: Investigating Student Learning with Visualizations in a Physical Science Course – S. Raj Chaudhury, Norfolk State University.

Norfolk State is a historically black (95%) and female (65%) institution, which has been developing BEST (Bringing Education and Science Together) on a CASTL (Carnegie Academy for the Scholarship of Teaching and Learning) grant.

This particular project in BEST looked at how students moved among representations, asking the questions, "How do we help students best **decode** and **deconstruct** these representations?" The representations included diagrams, numerical mathematical, symbolic mathematical, verbal, etc.

Initial work looked at proportional reasoning (ratios etc.) using the verbal and visual representations. For instance, the same problem presented both as a word problem and as a diagrammed problem.

On the verbal problems, it was all or nothing. If a student had any success at all, they got the entire problem correct. There was more variation in the visual problems, with partial successes seen more often. Also, those who succeeded at the verbal problems did even better on the visual problems. These results came up time and again.

Conclusion: students have better visual than verbal skills, it'd be a good idea to try to capitalize on their visual skills.

This lead to LSD – Learning Science through Diagrams. (Me, I'd have called it LSTD.)

The first example of LSD discussed in this talk was teaching students about the structure of the Solar System, explicitly teaching them with diagrams and the spoken word, but little else. The initial student models were commonly what was referred to by researchers as the "bracelet model," in which the ordering of the planets is incorrect and several share each orbit, like charms on a bracelet.

LSD instruction explicitly used four diagrams.

1. A "real" orbital diagram, seen from about 45 degrees down from the system axis, with planets as dots, orbits as bright lines, and trying to evoke a three-dimensional view.
2. A chart showing the relative sizes of the planets, but not their orbits.
3. A side view showing the radius and tilt of the orbits.
4. An old-fashioned geocentric diagram (presented for contrast).

When students were asked what was different about the geocentric model, the most common answer was "Pluto is missing".

Pre-instruction, about 35% of the students gave a more or less correct model with the right order of planets and generally right distribution (i.e. terrestrial planets all close in, gas giants spread way out, Pluto being wonky). Post-instruction, 70% did a good job of reproducing the diagram 1. A few science fiction fans had models with the asteroid belt and some of the moons included.

Post-instruction, 25% drew the size comparison model instead (diagram 2), and most of the rest went with diagram 3's orbital tilts.

The second diagram example was the path (somewhat abstracted) of a ball thrown up in the air, essentially a strobe picture showing height, velocity and acceleration as three layers of detail. So they started with just position, then added velocity, then added acceleration. This was the **only** text they used for studying one dimensional motion with acceleration, but with all three layers in place, it's essentially visualized calculus. (Warning: do not **tell** students that it's calculus, they'll get intimidated.)

The targeted group scored a 62% on exam items related to this topic, while the control group scored 35% on the same items.

However, in interviews about the diagram, once the word "science" hit the table, students shifted to using textbook definitions and jargons, even if they didn't understand them, abandoning the comprehension they'd gained from the diagrams. A frame problem.

Future research goals include assessing actual understanding of the diagrams (make sure they get the point), trying animations instead of static images, and using palmtop computers in the classroom to deliver these animations and let students manipulate them individually.

CC02: A "Zen" of Physics Teaching – Dewey Dykstra, Boise State University.

This talk is looking at PER for non-STEM students, a population that varies widely on almost every measure when compared to STEM students. They typically have "constrained electives" in science, meaning they have to take some sort of science, but are relatively free to pick the specific class. They tend to end up in large lecture courses that lack a lab requirement.

Conventional instruction routinely fails these students. They come out of the course essentially unchanged from their initial condition, if not worse off in attitude. However, this brings two silver linings:

- The students have already been written off by just about everyone ("These students will never learn to think." – unnamed source), so you can implement new ideas without worrying about complaints from other faculty and administrators, as long as enrollment stays up.
- Because these courses are rarely pre-requisites for any other course, you're not constrained in terms of the material you have to cover. If you want to spend all semester examining energy concepts and never touch kinematics, no worries!

Because the goal doesn't have to be "make sure these students have covered this list of topics they'll need next semester", you can focus more on educating instead of training. Students should leave with a new understanding of the world, believe they can make sense of it, have the skills needed to work with others, and have good attitudes towards science.

Content-driven instruction typically follows the training paradigm, aimed at giving the "canon" to the "deserving". But this view, if it's ever appropriate at all, is wholly inappropriate to non-STEM students. They aren't the "deserving," and they don't really need much of the "canon".

Rather, we want these students to gain both experiential and explanatory knowledge and the correspondence rules to link it all together. Ideally, we'd like the STEM students to get this too, but as mentioned above, it's a lot easier to fiddle around with non-STEM courses. See M. Jammer's Concepts of Force, reprinted in 1999, for an example.

The Physical Review Special Topics in PER online journal will have results in more detail, or see http://www.ipn.uni-kiel.de/aktnell/stese/stese.html for more, but here's a few highlights.

The standard STEM course at Boise State, with a low pre-test score and a content-driven canon, had FCME gains of about 0.15. The understanding-driven non-STEM course had equally low pre-test scores, but an average gain of about 0.6. Whether a student had previous high school physics experience made no difference in this study.

So, here's the titular Zen of Physics Teaching:

- Don't shoot for the canon, just try to make sense of things.
- Students do better when they're not trying to **guess the answer you want**.
- Teach by koan? (Went by this point a little too fast for me to catch it.)
- Fit explanations to experience.

Regarding the above boldface, students trying to guess the answer that the teacher wants are not actually trying to understand the material, they're playing an entirely different game. By not focusing on the canon, students can be encouraged to try to understand things.

CC03: Student Understanding of Energy in Courses for Non-Science Majors – Michael Loverude, CSU Fullerton.

Why study non-STEM students?

- They're there.
- They're often articulate and skilled in language use, so interviews may be more fruitful.
- They're important to the future (i.e. they go on to be policymakers).
- There's little pressure to cover the canon, except in teacher-training courses.
- They are similar to high school students in a lot of respects, which is useful in helping broaden the impact of the research for the AAPT's membership.

The course being discussed here used an Energy-First curriculum and a systems approach. The Grade 3 science standards in California include a fair number of energy concepts, so this was good for any teachers in training.

Students tend to think of energy, heat and current as physical "stuff" (i.e. caloric, phlogiston, if not by those names). On the other hand, physicists tend to see energy as an abstract construct unifying disparate physical processes...but find the "stuff" model a useful simplification.

Research set out to see if the "stuff" view was robustly held by students (i.e. did they really think that something that cooled off got less massive as heat-stuff left?).

Does mass change? About half said it did, for various reasons. Stuff is used up in energy transfer, while thermal expansion is due to the object being packed full of stuff. About 20% thought heating increased mass, as more stuff was being added. Another 20% thought heating made things lighter, because hot air rises, expansion makes for lower density, etc. (Mass/density conflation possible here.) The responses were often inconsistent in their reasoning.

Results from Summer 04 talk also discussed very quickly, such as the idea that in the "ball rolling in bowl" oscillator that gravity sucked energy out.

It was added at the end that STEM majors also tend to hold these views.

CJ – Probing Physics Understanding (contributed)

CJ01: What Don't Your Students Know? The Most Astronomical Misconceptions – Jeffrey Bennett, University of Colorado at Boulder.

This talk is not intended to cover things like seasons, phases of the moon and other basic misconceptions, it's aimed at misconceptions of modern astronomy.

0th Misconception: Belief that science is just an accumulation of facts and beliefs, either a "stamp collecting" of trivia or a sort of secular religion. This covers all science, not just Astronomy, and one must teach **process** to have a hope of overcoming it.

1st Misconception Group: Problems with scale and context.

- View of space as crowded with stars and planets ("Kirby Space"), rather than mostly empty.
- Conflation of star systems with galaxies (something abetted by science fiction, where writers don't seem to know the difference either).
- Belief that 10^{11} stars is an awfully small number for a galaxy.
- Belief that we're all under a cloud of impending doom because of what science predicts will happen in five billion years.
- Confusion over scale of planets and orbits from standard abstracted diagrams.

2nd Misconception Group: Space and Earth are different concerns.

- The rules are different, what works on Earth doesn't necessarily work the same way in space, and vice versa. Like Lovecraft's cosmology, without the horror.
- A lack of connection between the two environments.

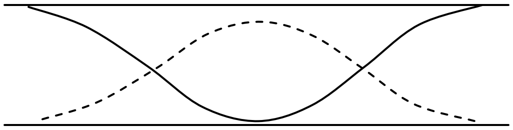
3rd Misconception Group: Grab bag.

- Since everything is moving away from us, we must be at the center of the universe.
- The Big Bang happened in a pre-existing spacetime, with matter exploding outward to fill it.
- Supraluminal speeds are possible.
- Scientists just made all this stuff up, theories aren't anywhere close to facts (i.e. a misunderstanding of what the terms theory and fact actually mean).

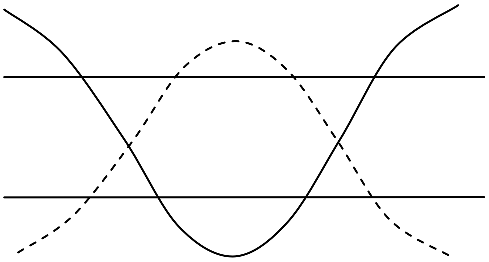
For more, contact Jeffrey.bennett@comcast.net

CJ02: Probing Student Understanding of Standing Sound Waves – Jack Dostal, Montana State University.

This is preliminary data from the early stages of a project.



When physicists consider the above diagram (Word Draw is horrible, sorry), we see it as an abstract representation of the longitudinal displacement of gas molecules as a function of position.



When we see this second diagram, we have no problems with it. It's inelegant, but still meaningful.

However, while students can label the nodes and antinodes, and give the definitions we want, they do not find the second diagram to be sensible. When asked in interview, they responded that in the case of the second diagram, the air was having to go through the walls of the tube. The lines aren't so much seen as transverse waves like on a string as they're seen as the path taken by air molecules as they blow through the tube. This drove the development of a standing wave assessment tool.

The desired model: longitudinal displacement along direction of the tube, average positions of the air molecules not changing, ignore thermal and edge effects.

The goal of the assessment instrument is to see if students are learning the desired model and using it consistently. So far, students are getting post-instruction scores in the 35% to 50% range.

One common problem seen in the assessment tool and related interviews is that nodes are seen as places of constant pressure, rather than places of no oscillation. This pressure maybe higher than average, as molecules move to the node and get trapped, the sound holding particles in place (related to how sand piles up in the nodes on a drumhead?).

CJ03: Research as a Guide for Developing Curricula (Tutorials) on Waves and Physical Optics – Mile Kryjevskaia, University of Washington.

Developed 2 tutorials intended to help students get up to speed on transmission, reflection and reflection as waves change media.

Used transparency overlays with wavefronts of different wavelengths and the principle that wavefronts must stay connected to help with refraction.

The high number of variables (and subscripts) can confuse students, and the "simple" equations like $v = f\lambda$ are actually harder than the complex equations.

Wave materials from Physics By Inquiry workbooks were adapted into Tutorial form, with the goal being to build students up to the wave equation starting with basic pulse activities.

The new Tutorials improved both the incidence of correct answers with correct reasoning and of valid reasoning on incorrect answers. Another cycle of modification and refinement is planned.

CJ04: Addressing Student Difficulties in Understanding Phase and Phase Difference – Homerya Sadaghiani, The Ohio State University.

Interviews were conducted and the results used to develop both multiple choice and open response instruments. The data from these instruments was then used to develop tutorial-like worksheets on the subject of waves. The group being studied was students taking regular and honors calculus-based third quarter introductory physics (waves and modern physics).

Phase was involved in the items students found most difficult.

Phase and position/velocity were often conflated (i.e. phase oscillates rather than always increasing with time).

Students had difficulty relating phase to y(t) graphs.

Students generally could not successfully link phase differences to path length differences.

In the area of interference, neither path length difference nor phase difference were linked to the order of interference fringes (m).

CJ05: Gender Differences in Representations of Electric Circuits – Jill Marshall, University of Texas at Austin.

Thesis: men and women draw and interpret circuit diagrams differently.

A spectrum of operations was established: iconic/concrete at one end, symbolic/abstract at the other, and "figural" in the middle.

Men tended to make diagrams more abstract and angularly geometric, while women made diagrams that were more curved and "organic", but both genders tended to stay mostly within the "figural" section of the spectrum. The differences aren't strong enough to really call "abstract vs. concrete".

Study was too small to get more than some hints, however, and the death of the chair of the Institutional Review Board threw things into disarray and prevented the second phase of the study from being carried out.

CJ06: Examining Student Understanding of Electric Circuits – MacKenzie Stetzer, University of Washington.

Abbreviations for this talk: CC = Complete Circuits issue. BBW = Battery, Bulb, Wire task. One battery, one bulb and one wire are used to try to complete a circuit.

Building on the work of Engelhardt et al on the BBW task, it was decided that complementary tasks were necessary to filter out the effect of students not understanding how the interior of a bulb works. So they started working on trying to figure out what was going on.

Short circuit arrangements turned out to be much more popular in multiple choice questions than in free response.

Most students with incorrect diagrams did fail to understand how a bulb's connections work.

Students with correct diagrams usually knew how bulb connections worked, but not all. About 15% of the students drew a correct circuit diagram but an incorrect "how the wires hook up inside the bulb" picture. These students probably don't really understand CC.

The BBW may overestimate robust understanding, but it does provide insight.

CJ07: Development of Student Reasoning in an Upper-Level Thermal Physics Course – David Meltzer, Iowa State.

The course in question was a roughly equal mix of classical thermodynamics and statistical mechanics, taken by students who had all had some exposure to thermodynamics in previous classes. It was taught as a blend of lecture and IE, and they're working on developing materials.

The students were strongly split into high/middle/low brackets, making it hard to generalize.

A diagnostic was administered at the end of the introductory modern physics course, then again at the start of the thermo/stat mech course.

The misconception that work equals $\Delta(PV)$ independent of path remained consistent between administrations. However, the view that ΔQ was also independent of path lost favor, and some of the partially correct explanations gained favor.

Some items on the instrument considered heat engines, and posed various situations while asking if they were physically possible (and why). Students generally caught violations of the 1st Law of Thermodynamics, but none caught violations of the 2nd Law. The posttest expanded on these questions, and students did better (about 50 to 60 percent correct with correct explanation).

Students had difficulty with entropy in irreversible processes.

Progress in correcting difficulties was generally slow.

CJ08: Students' Ideas about Entropy and the Second Law of Thermodynamics – Warren Christensen, Iowa State. (Sequel to CJ07)

The introductory modern physics course discussed in CJ07 was studied in Fall 2004, and a diagnostic instrument given as a pretest. The concept focused on in this talk is the matter of what happens to entropy (S) in a non-closed, spontaneous process.

In general, the change in S for such a process can be positive, negative or zero. Only ΔS for a totally closed system (such as the entire universe) is always positive. Only 15% of students said $\Delta S_{\text{universe}}$ increased, while over 70% claimed that it was zero (about half of these invoking a conservation principle). Only 4% got the entire item (open and closed systems) correct.

A similar item was given on the final exam, tweaked to fit the instructor's style of exam problem. The responses were consistent with the pretest, if a little better (30% correct for a closed system).

Unfortunately, interview data suggest that even this small improvement is overstated. While students were more likely to pick "entropy increases" when forced to make just one choice, most interview subjects thought "entropy remains constant" was almost as good of an answer. And they generally thought entropy was defined as disorder and randomness.

Compared to the upper-level students discussed in CJ07, these students still don't realize that entropy in an open system may be undeterminable, but still do a little better overall.

DE – Practical Pedagogy (contributed)

DE01: Achieving Large Conceptual Gains in a Traditional Introductory Astronomy Course – Brian Buerke, Albright College.

In a modified traditional course, large gains were found on the Astronomy Diagnostic Test (TPT 31 162-167 (1993), an old version of the ADT). Gains pre-reform were around 0.20, while after reform they were between 0.40 and 0.50, a change large enough that it's probably real, even though the ADT version that was used has recognized flaws.

The labs were very project-oriented, involving the construction and later use of astronomical instruments (like sextants and telescopes).

Lecture included typed-up notes distributed to students, but these were not complete lecture notes. Rather, they had blanks for equations and sample exercises, which students had to fill in as they went. This allowed a certain amount of IE, while providing structure for students with poor note-taking skills.

DE02: Student Impressions of Sample Exams – Carol Koleci (pronounced Ko-LET-si), Worchester Polytechnic Institute.

Considered the classic situation of "I read the book, took notes, aced the homework...and bombed the exam." In other words, the failure of transfer.

Sample exams were given in class, with 100% given for any attempts, the scores counted in the homework grade. These were peer-graded in class, and a survey for additional extra credit was appended to the sample exams.

The construction of the course meant there had to be three real exams in the course of 7 weeks.

Problems on the sample exams are similar to those found in Halliday, Resnick and Walker. Students generally did better on the parts of the real exams that corresponded tightly to problems on the sample exams, but this improvement did not generally require that students recognized the similarity (i.e. the problems were similar on deep structural levels, while students generally match on surface features or equations). On exam 3, students who did notice the similarity tended to do better, but this was not statistically significant, and could also have been related to the idea that better students are more likely to sort problems based on deep structural elements.

All students cited practice exams as the most helpful thing in the course. The data also suggests that peer grading may be helpful, but more research is needed to be sure.

Also, calling it a "sample exam" instead of a homework assignment seems to have been partially responsible for the success of the effort.

DE03: Continuing the Battle Against Post-Exam Syndrome – Kathy Harper, The Ohio State University.

Post-Exam Syndrome: frustration with performance on an exam hurts performance on later exams, a self-fulfilling prophecy of failure with no attempts on the student's part to figure out **why** their score was low.

To combat this, students were given a tersely graded version of their exams and tasked with correcting their errors and explaining why the errors were errors. The exam would be discussed in class only after this self-correction cycle had taken place. This made assessment more formative, especially since students were graded on how well they corrected and explained...the old "they won't do it unless they're graded on it" situation.

The perception of usefulness was dependent on the instructor, so they're trying to change the process so that it's less susceptible to this.

Students need to have time to process things, or there's no effect (i.e. handing it back and telling them to correct it in class NOW doesn't help).

The effects have been small, but promising. At the least, it seems to help retention and discourage the "empty the brain after the exam" effect.

In Winter 05 quarter, they will examine the instructor effect more closely, and use a larger treatment group.

DE04: Including Hands-On Activities in Cooperative Learning Environments – Eric Page, University of Rochester.

Cooperative Learning Environments (CLEs) are institutionalized study groups of 5 to 8 students with a mix of ability levels. CLEs are assigned times to work together in workshops run by teaching assistant facilitators (both graduate and undergraduate, not always physics majors). The goal of the CLEs is to provide a low-stress (or even fun) learning situation, and **engage** the students.

Learning, retention and student happiness have been measured since the CLEs were instituted in 1999 for the algebra-based course (as well as other courses across the university, a total of 23 faculty in various departments use them now). Attendance is not (supposed to be) graded, but it does correlate positively with performance.

CLEs have hands-on activities where possible, and students purchase a $10 "toy box" of equipment to bring to the workshops.

Students have shown now strong preference for any type of activity, but the participation has been better in hands-on, invigorated groups.

DE05: Cancelled

DE06: Introducing An Energy Theme Into A 9th Grade "Physics First" Classroom – Christopher Smith, Barrington (IL) High School

This freshman-level course was designed with an Energy-First theme in mind, and that theme is carried through to biology and chemistry in later grades. It includes the following elements:

- Identification of energy storage locations rather than defining U, K, etc.
- Description of the transformation of energy from one form to another.
- Use of energy pie charts for cases of closed systems (switched to bar charts once the idea of Work was introduced later on).
- Identification of where dissipated energy goes.
- Developing conservation principles.
- Labeling all energies as $E_{\text{something}}$, rather than using other letters like U, P, T, V, etc.

Students would start with a whiteboard activity on energy storage and categories of energy. They would find situations that matched a given transformation (such as E_{stretch} to E_{moving}) and then draw it out on the whiteboard. There would be no calculations, instead pie charts (and later bar charts) would be used to track energy. Discussion was seen as more important than exact right answers.

Students worked on a Rube Goldberg device at the end of the energy unit.

By the end of the energy unit, students had relatively good common ground on energy, and used the subject throughout the rest of the course as well as in later courses. This grounding helped them later when W, Q and friction dissipation were introduced.

Aside: **g** was introduced in units of N/kg as a field strength, since kinematics (and the m/s^2 units) came later in the course.

DE07: No-show.

DE08: Adapting and Implementing Hierarchical Learning Ensembles: A New Pedagogy for Team-Building and Group Decision-Making in Undergraduate Engineering and Science Education – Jennifer Freeman, John Carroll University (Cleveland, OH).

Funded by an NSF CCLI A&I (Adaptation and Implementation) grant to address the gap between academic and work environments. Specifically, how the academic cohort does not really match up with the diversity of ages and backgrounds found in the workplace.

A project was devised that would let high school students, undergraduates and graduate students work together. The ensemble had two Graduate Teaching Assistant "managers" (one from Biology, one from Education), four undergraduate student "middlemen" (two each from Bio and Education), and four high school student "employees." They formed two teams, each team getting one Biology major and one Education major as middlemen.

The project was to develop a sort of bone spackle, a realistic bio-engineering task, over the course of one week in the summer. Each group had preparatory exercises before this week started, to get them ready for the project (i.e. the high school students studied the properties of Jell-O$^{\text{TM}}$ since the bone spackle would be gelatin-based).

Various team-building exercises and cooperative structuring elements were enforced, such as teams being required to eat lunch together, write up daily reports, etc.

At the end of the week, each team made a sales pitch for their material.

Student reactions included:

- Real science is HARD.
- Everyone has something to contribute, from top to bottom.

- Each discipline learned something from the other (education types learned the science, science types learned group work and organization, for instance).
- The energy level in lab was high, and confidence was boosted at all levels.

Committee on Laboratories Meeting

I am on this committee now, which met at the same time as the Research In Physics Education meeting, so I have no notes on the RIPE meeting, sorry.

See Tom Foster for complete minutes on this meeting, I'm just going to include a few pieces of possible general interest from the notes I took.

The Programs Committee has now established firmly that you can only give one talk per national meeting, and only one invited talk per calendar year (this second point being important in the matter of planning invited sessions!).

The Laboratories Committee has some overlap with the Apparatus Committee, but Labs is mostly about the curriculum of the labs rather than the equipment, and Apparatus is the other way around.

Sometimes a specific session is sponsored in advance for CCLI dissemination talks, but often such sessions spontaneously arise from contributed talks (or, as seen above, the CCLI dissemination talks just get slotted into the most relevant other sessions).

Rereading my notes, I find I volunteered to help get speakers for an invited session on the impact of the Americans with Disabilities Act at **Syracuse** (Summer 2006), not **Anchorage** (Winter 2006). This should make it easier to find people willing to come talk. Whew. And since I also agreed to be invited myself, that's a load off in case circumstances prevent me from making Anchorage.

Wednesday, January 12, 2005

EB – Plenary 3

EF01: The Dark Century: Matter, Energy, and the Future of the Universe – Virginia Trimble, University of California.

Science is:

- **What** we know (i.e. the universe is big)
- **How** we know it (i.e. measurements of size)
- **Why** things are that way (i.e. Big Bang Inflationary Cosmology)
- **Implications** for the future (i.e. universe cooling and thinning)

(general survey of astrophysics omitted)

We can see out to five gigaparsecs without noticing any edge effects, so the universe must be at least twice that size.

About 73% of the mass of the universe is dark energy/quintessence/λ/whatever you want to call it. This has been confirmed by multiple measurements (including supernovae luminosity distances and cosmic background anisotropy angular sizes).

The more money you want for your research, the less science knowledge is had by those who fund you.

(At this point, I was feeling ill and stopped taking notes, sorry. This also made me miss the morning session I had intended to attend after the plenary.).

FC – History of PER (invited)

FC01: A Retrospective on Physics Education and Research: From Aristotle to Arons – Len Jossem, The Ohio State University.

The slides from this talk are available on CDROM, and were zipped through at an intentionally too-fast pace. Email Jossem@mps.ohio-state.edu for a copy.

Socrates started off by annoying people in an attempt to teach them, setting the stage for the rest of us. However, his actual method was not the Socratic Method we use, it was a top-down authoritative transmission of the canon. We've taken the basic idea of asking questions and set it to a much different task.

Aristotle said ideas "arise" many times, not just once, but are often lost. Roger Bacon added that those ideas need to be backed up by experiment, and the instrumentation developed to pursue this Baconian ideal kickstarted Galileo and his contemporaries.

Comenius in the 1600s championed education for all, and Michael Faraday a bit later on came to be seen as a model for effective lecturing.

Maxwell once said, "Nowadays we have too much to teach and too little time to teach it." This seems to be a time-invariant statement. Maxwell also advocated the use of multiple representations.

Since 1970, the number of physics abstracts per year has more than tripled, with the upswing starting in the 1960s. (For education reference, Bloom's Taxonomy came out in 1965, just as things were seriously ramping up.)

With more physics being done, there's more incentive to teach well, but there were very few failure-analysis studies of teaching done in the 60s or early 70s.

Unfortunately, during that time there was also a shift away from internal funding of research by universities, and towards external funding (such as NSF or corporate funding). This has shifted loyalties away from the university and reduced incentive to improve teaching.

FC02: From Arons/Karplus to Redish/Zollman, the Physics Education Reformation – Robert Fuller, University of Nebraska at Lincoln.

Prior to the work of Arons, the focus was purely on transmission of the canon. Teaching was seen as an art rather than a science, and content was king.

The reformation got rolling with Renner and McKinnon in 1971, followed by Karplus and Arons nailing their theses to the door in 1976 (AJP 44, 396).

The "theology" of this reformation came from Piaget's work in genetic epistemology (genetic in the sense of pertaining to a genre, not in the DNA sense), intellectual development and equilibration/self-regulation. The driving force was the semi-clinical interview, the radical proposition that if you want to know what a student is learning, you might do well to ask them.

The main thesis nailed to the door was that schools were not helping students develop formal reasoning skills, with only about a third of students possessing these skills. This needed to be fixed!

The Reformation Era shifted the focus from the content to the learner, with the premises that teaching could be treated as a science, and that therefore research into teaching would be worthwhile.

The Modern Era started in the 90s, kicked off by the work of Redish and others. It drew from cognitive science and other fields, brought in multimedia methods, and started training people specifically in PER via the use of Graduate Research Trainee grants. The PER Supplement to AJP was launched, and PER people started to get on various review boards **as** specialists in PER. New journals started to be launched.

The Post-Modern Era is beginning now as people with PhDs in PER are starting to take up the reins from earlier researchers (who came into PER from other fields). The various methods introduced in the earlier eras are being combined, with a mixed method of both qualitative and quantitative nature arising. PER results are finding more toeholds in teacher training and general education physics courses, and there's a greater political savvy shown by PER researchers (especially compared to reformers like Arons, who tended to use vinegar rather than honey).

(Fuller presented lists of people he felt fit the Reformation, Modern and Post-Modern eras, but there was a lot of audience disagreement on some of the names, so I didn't even try to copy the lists down.)

Ontogeny recapitulates phylogeny. Just as the field as a whole as progressed, individual researchers generally need to pass through all four phases (Orthodox, Reform, Modern, Post-Modern) in their professional development.

FC03: Fostering Change in Science Education: Creation, Implementation, Evaluation and Research – Uri Ganiel, Weizmann Institute of Science, Israel.

PER is a softer science than other branches of physics, live with it. There's no **hard** proof available on the same level as can be found in, say, Atomic/Molecular/ Optical Physics.

We are in the middle of the Information Revolution, and science education needs more reformation. The 1960s post-Sputnik crash courses didn't stick, and in any case were based on WWII-era ways of doing science. When the momentum died down a few years later, things stopped working.

Mode 1 Reforms – "Fire and Forget", set up the program and then assume it'll do fine on its own.

- Melba Phillips said, "The trouble with problems in physics education is that they don't stay solved."
- Ken Wilson described Mode 1 as "Five years and out." Political leaders leave before any of their policies come to fruit, and funding cycles are generally only about 5 years. So once the money runs out, there's no longer the political backing to get it renewed.
- Projects die before they've really had a chance to take root.

The necessary ingredients for development are:

- Subject matter expertise (necessary but not sufficient)

- Continuity and persistence past the 5 year mark.
- **Applied** research on teaching and learning.

We have had many dire warnings about educational standards over the years, but the resistance to a national syllabus remains high. Not that a national curriculum can't be uniformly **bad**, mind you.

If the dedication is there, Mode 1 can transition to a sustained effort.

A continuous and reversible cycle of research, curriculum development and implementation (Mode 2? He never said) would be better.

Teachers are central to any innovation or reform effort, the politicians need to know this. They also need to know that a longitudinal systems approach, vital to success, will be slow...but it **must** be steady. Inconsistency kills.

Witticisms

Here's a couple of things I wrote in the margins of my notes that were somehow inspired by the talk I was listening to, but didn't really merit inclusion in my notes for that talk.

Just when you think you have it nailed, you realize it was a screw.

You may catch more flies with honey, but you kill more maggots with vinegar.