2003 National AAPT Meeting – Summer
Notes from August 4-6
Madison, WI

The following notes were taken by Dr. Van Domelen during the regular sessions of the 127th National Meeting of the American Association of Physics Teachers. Only one Ceremonial session is included, and the PERC is not included (as the PERC Proceedings should do that job well enough).

Abbreviations: There are a few concepts that came up frequently enough in the talks I attended that I gave them their own abbreviations.

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

Monday, August 4, 2003

AI – PER: Student Characteristics and Behavior

AI01: Belief versus Understanding ("Scientist Answers"): Effect on Assessment – Timothy McCaskey, University of Maryland

Students do not generally believe the answers that they give on the FCI. They will give the answer they think we want, but not truly accept it as reality (i.e. they KNOW N3L, they just think it's bogus). Males tended to have slightly higher scores, while females tended to have a greater split between their FCI answers and what they really believed.

The FCI was given twice to each group. The first time was a regular administration for scoring. The second time, they were asked to put a circle around the answer they believed in and a square around the answer they thought a scientist would give.

A "split" was defined as a response where the "my answer" and "scientist answer" disagreed. Males averaged 8 out of 30 items with splits, women slightly more.

Looking at the N3L items on the FCI, only 15 out of 116 students got all of them right, and most of these students picked "my answer" options that were incorrect.

Conclusion: many students "lie" on the FCI, telling us what they think we want to hear, not what they believe.

AI02: Factors Affecting Attitude and Performance Variations in Active Learning Courses: Effects of Gender on Performance in a Biology Course Taken After Reformed Physics – Wendell Potter, UC-Davis

A physics course (Physics 7) was reformed to focus on critical reading and thinking skills, "making sense" of phenomena in terms of widely applicable models. It operated on a Discussion/Lab (i.e. studio) core of 10 hours a week, plus one 80 minute lecture with weekly quiz (that emphasized problem-solving and what to do when you don't have enough info).

By forcing "think about it" questions, Physics 7 reduced GG. A hard criterion-based grading system where 80% was seen as acceptable performance was used. They found that the lower the average scores, the smaller the GG.

MCAT scores improved for all that took the reformed Physics 7 (in all areas, not just the Physics part), and the GG narrowed there as well. Scores on Intro to Anatomy showed similar improvement and reduced GG for those who had taken reformed Phys7. Unreformed Phys7 had a positive effect on all scores (compared to those who did not take Phys7), but did not reduce the GG.

AI03: Gender Differences in Student Responses to Workshop Physics – Eric Anderson, Avila University (Kansas City, MO)

The course under study is a small upperclassman course in algebra-based physics using Minnesota's cooperative system, a workshop environment and several IE methods. GG persists, despite improvements in FMCE scores.

The MPEX revealed a strong if not very significant correlation between attitude and FCI gains for female students ($r = 0.46$, $P<0.10$), no correlation to speak of for males. Conclusion: attitudes and expectations make a bigger difference for women.

- Females may start with some better attitudes, but in some cases the course hurts their outlook.
- The differences may suggest that there are some useful resources that may be tapped for teaching women.

AI04: Locus of Control – Chance or Self-Attribute? – Vince Kuo, University of Minnesota

Locus of Control is a statement about how people see the relationship between their efforts and the things that happen to them. Internal LoC (also called "Mastery") tends to consider everything to be the result of their own actions, while External LoC (aka "Helplessness") sees everything as being caused by forces beyond one's control. An analogy is that Internal is like a bowler (skill matters) while External is like a craps shooter (skill irrelevant, except where it pertains to dealing with what comes). Internal

LoC tends to result in a higher motivation, since the person believes that their efforts make a difference.

[Aside unrelated to the talk: some studies have shown that we do not have consistent LoC. Specifically, males tend to be Internal with regards to successes and External with regards to failures (I succeed because I'm good, I fail because the teacher's being unfair), and females tend to be the opposite (I succeed via luck, I fail because I didn't try hard enough). And, of course, situations can also trigger a shift in LoC; you're more likely to be Internal in an area where you consider yourself to be an expert.]

Question: do physics courses tend to make students more External?

They use the Rotter I:E Scale to measure LoC. It's a 29-item test that is easy to use and has been highly validated in other contexts. Students are presented with two statements and asked to pick the one they most closely agree with. Some items deal with Social things (I vs. E in politics, for instance), others with Personal things (like academics) and a few are red herrings to keep the subject from guessing their way to a preferred result.

College students as a whole tend to be Internal, males more so than females. However, as the years and decades have gone by, the amount by which they are Internal has fluctuated, perhaps due to the social climate.

In a class using IE, Internal LoC leads to higher achievement. In a class using traditional lecturer-dominated methods, there is no correlation between LoC and performance.

UMN collected a raft of data (I:E test, CSEM, FCI, etc) and looked at it.

- All students were Internal on the Personal scale.
- The classes became slightly less Internal over the course of the semester.
- Males started out more Internal than female, and the GG widened by the end of the semester.
- In their calculus-based courses, Internal students were always higher achievers. In the algebra-based courses, this happened to a lesser extent.
- All differences were minor, but statistically significant.
- The political/social scale shifted from neutral to External over the course of the semester, probably due to Gulf War II: Electric Boogaloo.

AI05: Investigating Learning Style Preferences for Conceptual Physics Students – Steven Sahyun, University of Wisconsin at Whitewater

Used a learning-style assessment tool called VARK – Visual, Aural, Read/Write, Kinesthetic. VARK was given as extra credit in the course, and those who did not participate tended to have grades of C or lower.

The class as a whole was pretty strongly K, with VAR levels about equal to each other. The national VARK scores were similar, but with a little less K and a little more R. The difference between K and any one of the VAR elements in pairwise comparisons was always significant.

Conclusion: Kinesthetic learning is preferred. "I do and I understand" – Confucius

AI06: An Epistemological Intervention with Big Results – David Hammer, University of Maryland. Talk presented by Rachel Scherr, Hammer had travel troubles.

The course under consideration is a Physics for Bioscience Majors one. It does well on gains for MPEX and MPEX2 (almost all courses actually have losses on these surveys).

A student codenamed Louis was examined and his situation seems reasonably generalizable.

Louis scored a 36% on his first exam, and two weeks later scored 84% on a makeup exam. Between the two he went to Hammer during office hours, and he was interviewed after the makeup exam (Hammer wishes he'd thought to record the office hour!).

The advice Louis was given in office hours was "Try to explain it to a 10-year-old". He realized that the way he had been studying (memorization cramming that worked for Organic Chemistry) was not useful here, and that making up his own analogies, even somewhat flawed ones, helped him a LOT.

Hammer's claim: Improvement hinged on Louis's epistemology (how he thought about stuff). The short timescale shift to a different mode was simply an activation of a previously available mode (albeit unstably).

Scherr's claim: Improvement also hinged on traumatic failure [Aside: an old saw says that to teach a mule something, you first have to hit it in the head with a two-by-four. This will get its attention.]. Louis had been given the advice earlier in the semester, but it wasn't until the dramatic failure of his existing epistemology that he was ready to accept a new one.

AI07: Epistemological Gains in a Large Lecture Class – Joe Redish, University of Maryland.

UMD is exploring the "hidden curriculum" of epistemology. Student expectations control the class regardless of our desires.

The MPEX starts by asking "What do you think you're going to do in this class?" It's what they think they think, not what they really think. [Aside: "You'll PAY to know what you REALLY think!" – J.R. "Bob" Dobbs] As a result of this, MPEX scores almost always go down. Can this be changed?

Yes.

The course (the one in AI06, I think) was changed to emphasize the VALUE of improved thinking. It was difficult to compare to a control group, since both the population and the instrument (MPEX2 now) had changed, but they managed to finesse a sort of pseudo-control. The reformed class's attitudes improved instead of getting worse.

In the second semester of the course, Magnetism and Electromagnetism were dropped in order to focus more on other topics, and it was harder epistemologically. Students who had taken the reformed first semester showed no further gains on the MPEX2, but those who had taken the unreformed first semester caught up to those who had.

AI08: Physicists' Epistemologies of Quantum Mechanics – Raymond Hodges, University of Maryland

"Common Sense" is usually a pretty good epistemology for physics, but it falls down when you get to Quantum Mechanics. What IS a good epistemology for QM?

Experts tend to think first in terms of formalism, then move in one of several ways:

- Models: realizing the map is not the territory
- Recipes First: but recipes tend to fail when pushed too far
- Experiment-Driven: appeal to the experiments

Details of a "good" epistemology vary from expert to expert.

AM – Ceremonial 1

AM01: Energy – A Basic Physics Concept and a Social Value – John Roeder, Calhoun School (New York City). (Didn't catch the name of the award, and it's not in the Announcer that I could find.)

This was essentially an autobiographical talk about the Energy Education work he did that got him the award given in this session.

Main point: energy is neither produced nor used; it is transformed from more useful to less useful forms (i.e. heat).

In teaching about energy, he used the invariance of $\mathbf{F}\cdot\mathbf{d}$ to motivate definitions of work and gravitational potential energy on a roller coaster, and then got kinetic energy from taking the coaster downhill. He discussed various snags in his program and curriculum, most of which he claimed were fairly easily resolved.

At the current time, renewable sources of energy plus nuclear energy only provide about 10% of our nation's energy, falling short of even the pessimistic goals set in the 1970s.

AM02: Assessment as the "Hidden Variable: in Conceptual Physics Achievement – Michael Zeilik, University of New Mexico. Excellence in Introductory Undergraduate Physics Award.

By saying that the effects of assessment are hidden, he means that we don't notice these effects enough, not that we don't notice them at all.

Three educational strategies have robustly done well in the 20th Century, with effect sizes consistently better than a difference of 0.3 standard deviations between control and experiment groups:

- Active Learning/Mastery Learning (0.5 SD)
- Cooperative Learning (used in many disciplines, 0.5 SD)
- One-on-one tutoring with trained tutors (2.0 SD)

Everything else attempted by PER has simply not worked well, robustly or consistently. This includes:

- Lecture.
  - It just reinforces memorization habits and can give the illusion of competence (aka The Feynman Effect)
  - "Educational Seduction" showed by the "Dr. Fox" experiments in med school. Charismatic nonsense is rated higher by students than good teaching.
- Computer/Web-Based instruction. While sometimes more convenient, it doesn't seem to help with learning in any measurable way.

Active Learning + Effective Classroom Assessment = "Sticky Learning" (Deep conceptual learning)

IE makes it sticky, but requires alternative forms of assessment.

There is a need to merge PER with AER (Astronomy Education Research) into PAER to avoid duplication of effort and better disseminate findings. Talk shifts gears here to discuss results of some AER at UNM.

The Astronomy Diagnostic Test (ADT) has been nationally normed, so was used to measure the effect of instructional changes at UNM. UNM classes did better on the ADT than the national norm, and also closed the GG.

The Student Assessment of Learning Gains (available on www.flaguide.com) was used to assess the classes, as it is designed to avoid the Feynman Effect (where a really good presentation can fool people into thinking they actually learned something, when in fact they didn't).

There was no GG in student self-assessment of their own abilities pre-instruction, but a large GG in terms of the math background students came into the class with. Male students had more and higher-level math, and were more confident in their abilities.

The use of cooperative quizzing helped reduce the GG over time. Students were allowed to change their answers after a small-group discussion.

Students in the study tended to expect higher grades than they actually got, MUCH higher in some cases.

By the end of the course, there was no GG for the FCI gains. UNM also created a new General Relativity/Cosmology inventory, and it showed no GG for students who completed the reformed course.

They concluded that gender sometimes matters. It does matter for Content, Concepts, Connections and Context. It may or may not matter for Self-Confidence and Competence. Good formative assessment yields good results, and unlearning is important ("Kill the (old) teacher").

BH – Innovative Teaching

BH01: Development of the Physics Teacher Self-Assessment – Dave Van Domelen, Kansas State University.

BH02: Why Are You In Physics? – Kastro Hamed, University of Texas at El Paso.

In the past 13 years, there has been a downward trend in the number of physics majors, and this has been placing pressure on many departments. Most studies of this problem have focused on the students who leave, and have treated all students quantitatively as experimental subjects. They have lacked a voice for the students who stayed in the major.

"What are the qualities of 'loyal customers'?"

Qualitative interview protocols were used. Students were studied at both UTEP and at Georgetown University, significantly different populations.

UTEP:

- 14 out of 22 physics majors decided to be physics majors before entering college.
- An interest in research and a desire to use mathematics were the biggest reasons for retention among physics majors.

Georgetown:

- 7 out of 7 physics majors had taken Bio, Chem and Physics in high school, as well as calculus. Many had taken AP Physics and AP Calculus.
- Many liked getting their "Why?" questions answered, and disliked both Bio and Chem.
- Families did not encourage them to major in Physics, and most discouraged the students.
- Most of the majors were somewhat turned off by the more abstract upperclass theory courses.
- None of the physics majors planned to attend graduate school in physics. A few planned to go on to engineering, most into law or business (family influence).

Moral of the story: talk to your students! Talk to their high schools! More contact, more communication is better, if you know what they like about the major you can better keep them in. Things like an active SPS and a friendly environment go a long way towards retention.

BH03: Helping Students Get The Most Out of the Introductory Course – Calvin Kalman, Concordia University

One of those speakers who whipped through PowerPoint slides too quickly to take any meaningful notes. He handed out copies of his slides, which are largely too small and cramped to make much sense of, but here's a few things I picked out.

The course made heavy use of "reflective writing" and small group discussion of same. An important point about reflective writing is that it is not essay writing, do not ask for complete sentences and polished theses…it's all about jotting down things as quickly as they come to you.

In Critique Exercises, students are asked to clearly lay out the contrasts between two differing perceptions of physics principles. They need to create convincing arguments for both the standard Newtonian framework and for an alternative conception.

BH04: Group Work in a Freshman Engineering Honors Physics Class – Bill Reay, The Ohio State University

The course under consideration is PHYS132E, the calculus-based honors engineering electricity and magnetism course (132H is the physics major version, 132 without any letters is the regular engineer-oriented calculus-based course).

The Engineering Department asked for group work in the "business style" rather than something like the Minnesota cooperative work method. The groups that were formed stuck together across all the courses in the engineering honors program (you have to take all the honors sections or none).

Students felt that these fixed groups helped them learn physics, and even more that being in assigned groups helped them build teamwork skills. "Leave no one behind" attitude.

The class's group project was to create a giant inkjet printer. They were not given specific training in group work, however, and as a result, groups generally did not meet outside of class (1 out of 17 did) and there was little structure. A mid-term interview of the groups helped kickstart things, but it was too little too late.

One conclusion of the semester's work was that the students need multiple group projects over the course of 131-2 to help bring home the various group work skills before they get to the third quarter robot competition project.

BH05: The Mystery Gap and the Process of Eliminating It – Shahzad Riaz, Ann Arbor (adjunct faculty somewhere in the area, he didn't specify)

Students don't know how to get from common sense to an equation, they lack the big picture. Riaz assumes that student common sense **can** be linked to a correct equation, however.

BH06: Divergent Thinking – Alan Van Heuvelen, Rutgers

Problem-solving skills are extremely important to both physics and everyday life. Such skills include:

- Defining the problem well
- Thinking divergently
- Characterizing a problem in multiple ways
- Evaluating as you go
- Etc.

This talk concerns the divergent thinking skill. Nondivergent thinking can lock you down into a path to failure, and you won't be able to get off that path until you have failed: it comes with a high opportunity cost. Divergent thinking can generate "offramps" that will save you from wasted time and effort. Therefore, we need to help students develop skill at thinking divergently.

- Make students come up with many ways to do a thing (an experiment, an evaluation, whatever)
- Brainstorming
- Have students pick criteria and then evaluate something according to those criteria.
- Encourage multiple hypotheses and then have them design experiments to test these hypotheses.

BH07: Physics 100: Introducing Engineering Students to Physics – Sara Wright, University of Illinois at Urbana-Champaign (talk given by Tim Stelzer)

The motivation for this new course: 20% of students in the intro calculus-based physics class get a D or less. That percentage rises to 68% among African-American students, despite very high standards for admission to the Engineering College.

Physics 100 was created as a half-semester prep course running from week 5 through week 12 (so that students could get in after realizing they were in over their heads in the regular engineering physics course), covering the early material more slowly for the at-risk students.

The focus of the course is on epistemology, and algorithmic thinking is eschewed.

The target students were not generally aware that they were underprepared (only about 4% were aware of it going in). A web-based self-evaluation was used to help aim at-risk students at Phys100, which has gotten about half of the underprepared students into the course. They would like to make this a mandatory placement exam in the future, for now it's voluntary.

Each week's study has the following elements:

- Online interactive "lectures" to get around schedule problems (by the time someone places into Phys100, the rest of their schedule is already worked out, making it almost impossible to find a time everyone can meet multiple times a week).
- Online "preflight" pseudoquizzes where students get credit for the attempt.
- Online homework. Some traditional WebAssign-like problems, some interactive examples with lots of help and walkthroughs.
- Two hours of face to face small group discussion.
- Online quiz, with feedback available after the due date.

Thanks to the large online component, they have a great deal of data that they're just starting to analyze.

BH08: Methods of Efficient Study Habits and Physics Learning – Nouredine Zettili, Jacksonville State University

Poor study habits in this talk are defined as inefficient habits. Quantity of studying is meaningless if the quality is low.

Premise: Poor study habits are the main barrier to success in physics courses.

Remedy: Start the semester with a lecture on good study skills and the research that shows why they're necessary.

One cause of inefficient studying is staying too long on one topic. Efficiency drops off as one topic is focused on, only rising slightly as the self-imposed end of the period approaches. If you switch topics every so often (50 minutes on, 10 minute break, then new topic is recommended), your efficiency doesn't decline much before you "reset" it. Cramming is really bad, because it exacerbates the boredom effect.

Additionally, recall drops off over time. The ideal times to review material to keep recall high are: 10 minutes after first exposure, 1 day later and 1 week later.

Finally, breakfast is very important. A hungry student is an inattentive student. And a quick sugar rush from soda or candy won't last all hour, so encourage students to get at least SOMETHING substantial like a powerbar or piece of fruit before class, no matter how early it is.

Tuesday, August 5, 2003

CI - PER: Student Perceptions and Responses

CI01: Learning: A Piece of the Puzzle – Rob MacDuff, Arizona State

The intent of this work is to create a cognitive model for problem-solving that is both plausible and testable. This is split into Knowledge and Understanding parts:

Knowledge – Distinction-making in PS, Coordination of distinctions into cognitive metrics, coordination of metrics into concept spaces (yes, it was a jargony talk).

Understanding – Transformation of concepts within and between concept spaces.

Students need to be able to pick out relevant distinctions from irrelevant ones.

Students are not consciously aware of the PS process. You need to slow things down and make students aware of the process.

CI02: "Earthquakes and Ice Water" – Spontaneous Generation and the Use of Analogies in Elementary Science Discussions – David May, University of Maryland

Analogies can be very useful, but they have not been much researched in terms of how students, especially young kids, USE the analogies.

A case study was made of a 3rd grade class, with the session under investigation involving a student's use of analogies in a classroom discussion of how earthquakes are caused.

Quick summary of the case study, since it seems to be showing up in every third Maryland talk this year. ☺ While discussing earthquakes, a student (Skander) came up with the idea that earthquakes were caused when a rock falls into the lava inside the earth, pushing it up and making the ground shake. When challenged on this by fellow students who claimed the rock would instantly melt, Skander amended his analogy by bringing in the idea of icewater…even if the ice melts, it still raises the water level by adding more water. Other students stood by the idea that when something melts it is totally gone.

Students in general generate analogies, validate them, construct new knowledge using the analogies, and then use the analogies to communicate this knowledge to others (all this often without realizing it consciously). Once the analogy is out there, others can criticize it, expand on it and generally manipulate it to try to come to an understanding of the material at hand. The analogy may not be correct, but it is still useful.

Conclusion: children do use analogy in a sophisticated way, we want to identify and encourage this.

CI03: The Development of Evidence-Based Instruments to Probe Students' Scientific Investigation Ability – Xueli Zou, California State University at Chico

This work was motivated by a need to teach higher-level reasoning skills (and show that this had been done), specifically the requirements of the new ABET standards. How do we measure if they gain these skills?

Observed students in an ISLE classroom, using a written survey linked to the conceptual framework of the class. In the future they hope to also use a computer-supported performance-based instrument.

The early work identified differences between experts and naïve students, such as the ability to tolerate "messy" data. The written survey tried to determine the students' positions on these scales. Form I was given only as a post-test and Form II was more context-dependent and was used as both pre-test and post-test.

Sadly, neither survey correlated with anything useful, further work is needed.

CI04: Context Cues, Associations and Learning of Physics – Lei Bao, the Ohio State University

Virtual Experiments using computer animation, force-feedback joysticks and Virtual Reality goggles were employed in this work. The VE system is superior to prior computer experiments because it takes place in three dimensions plus haptic (kinesthetic) cues through the joystick, and it's generally better at showing details (like slowing down during collisions to show the deformation). However, Bao went through about three talks' worth of slides, and there was no way to follow his actual points, sorry.

CI05: Zdeslav's talk.

CI06: Probing the (In)Effectiveness of the Conventional Introductory Astronomy Lecture – Tim Slater, University of Arizona (Ed Prather presenting in Tim's place). tslater@as.arizona.edu for more info.

Arizona has anecdotal "evidence" that lecture mode doesn't work. A number of things other than IE were attempted in order to make lecture work, but they had no real effect.

Conclusion, it's not so much what the instructor does that matters, it's what the **students** do. Students must be engaged or there will be no improvement.

U of A switched to lecture-tutorials for their large-class settings, using materials from the CAPER program. I forget what the C stands for, but the rest is Astronomy and Physics Education Research. 20 minutes a day of tutorials in with lecture had a large effect, and students liked it.

CI07: Impact of Homework Weight on Student Learning – Todd Ruskell, Colorado School of Mines

Used the LON-CAPA (http://www.lon-capa.org) system to give and grade homework. Huge amounts of statistics were available and were analyzed, but almost no significant effects anywhere (the few statistically significant results could very well have been the one in twenty random hits you get from a $P<0.05$ test). Introduction of the PRS may have also messed things up.

Based on possibly shaky numbers, a weighting of 10% for homework was advanced as optimal.

CI08: Writing and Peer Ranking in Introductory Physics – Rhett Allain, Southeastern Louisiana University.

Allain used peer evaluation of essays to get at the partial correctness hidden within wrong answers, and to expose students to different ways of thinking. Evaluation tasks lead to processing at a higher cognitive level.

Students had three types of evaluation:

Rating – give a score. Students often scored everyone high, so this was less useful.

Ranking – put them in order from worst to best, so you can't give everyone the same high score.

Review – a full critique.

The peer assessment was web-based, which allowed for anonymization (strip the names from essays before mailing) and saves a lot on photocopying. It also lets the instructor seed the packs with "exemplar" and "badlar" essays that are either really good or really bad, to make sure everyone gets one each of both types. The bad ones are frequently well-written but factually incorrect. Essays are otherwise "binned" into broad categories (rather than scored) by the instructor, and a student will get one from each bin plus an exemplar and maybe a badlar.

The key is feedback, not just from the instructor but also from other students. Students develop their ranking criteria as a class exercise. It's also important to give focused questions for the essays…questions that lack focus lead to essays that are hard to rank, because you get into the apples-and-oranges situation.

DC – PER: Curriculum Assessment and Student Difficulties

DC01: The FCI and Nonphysics Students – Laura McCullough, University of Wisconsin at Stout

This study is a start at examining the FCI scores of people in Literature and Social Sciences courses who are unlikely to ever take a physics course. UWStout lacks traditional academic majors, having "vocational majors" instead, so it's hard to be sure who will and won't take physics, but this group seems pretty unlikely to take physics.

The sample group is 150 students (100 female, 50 male), 46% are first-semester-ever Freshmen. The surveys were anonymous but with demographic data collected. 76% of the sample had algebra as their highest math, 41% had high school physics. Yet the average student had 3.18 years of high school math and 3.5 years of high school science, implying a lot of low-level "consumer math" and "earth science" courses.

Pretest scores averaged 20% for the women (with no reliability) and 32% for the men (with moderate reliability). These scores were only slightly lower than those for students taking physics at Stout.

The strongest distracters were "Force In Direction Of Motion" and "Acceleration = Velocity". N1L and "cut the string" circular motion problems were relatively easy for this group.

There was a significant GG, most prominent on the "path" problems (6, 12, 14).

DC02: Conceptual Learning in Physics for Liberal Arts vs. Algebra-Based Physics – Lawrence Day, Utica College.

Compared a course using Hewitt (PLA) and one using Giancoli (ABP). PLA followed a historical layout: Aristotle to Galileo to Newton to Einstein.

Both courses were taught in passive lecture mode, no real difference in pretests (on what, I didn't catch…probably FCI), ABP beat PLA on posttest. Strong correlation between posttest scores and final grade.

DC03: Initial Investigation of Students' Conceptions of Star Formation – Janelle Bailey, University of Arizona.

Early work on a star formation inventory of sorts. Examined Astronomy 101 students and students in a general education "Stars" course.

Three questions:

Q1) What is a star?

Q2) Where do you think stars come from?

Q3) How does a star form?

Q2 and Q3 are roughly the same question, just asked differently. Q3 tended to elicit more detailed responses.

A correct answer to Q3 had four elements: Gas, Collapse, Temperature Increase, Fusion.

55% of students started with a region of gas.

48% of students invoked some sort of collapse, with 19% specifying gravitational.

11% of students mentioned the temperature increase, but tended to misuse terms like heat, temperature, etc, which obscured the results.

Only 8% of students mentioned fusion.

Pre-instruction, 44% of students were fully incorrect, and only 1% were fully correct. The plan for the future is to ditch Q2 and take the time to perform interviews.

DC04: Incorporating a Tutorial Approach to an Advanced Mechanics Course for Physics Majors – Bradley Ambrose, Grand Valley State University (Michigan)

Used a mix of actual University of Washington tutorials and new ones inspired by the ECR style to cover vector fields, ∇ operators, etc. Administered pretests on various topics, both basic stuff (that students may have managed to avoid learning despite succeeding in physics courses) and more advanced topics.

No actual results yet.

DC05: Paula's talk, PowerPoint Gone Wild!

DC06: Teacher Understanding of Propagation and Resonance Phenomena in Sound – Katherine Menchen, University of Maine

Quoted work by Linder, Wittman and Hrepic.

Research performed at Grand Valley State, UMaine, UWash. All courses involved were Physics By Inquiry and used the pretests from the propagation section.

A sample item was a tuning fork touched to the end of a meter stick. Common incorrect responses included:

- Frequency is proportional to distance ("it fades")
- Frequency is dependent on the medium

At this point, things went south. Someone leaned on the light switch, leading to a momentary panic as everyone tried to figure out what was going on, and then the fire alarm went off and that was pretty much it for the session as far as I was concerned.

Junior PER Crackerbarrel

The main focus of the crackerbarrel ended up being communication and publication, with a side order of discussing the nature of AAPT committees.

Item: To be more of a community, perhaps we need a better way to communicate the non-research stuff that makes people feel like part of a group? Like a gossip server, or something like that. I suggested a webboard like phpbboard or ikonboard, where we could easily split it up into sections for serious research-related stuff, for career development chat, and for just plain chit-chat that would otherwise either clog up the serious mailing lists or have to wait for once or twice a year. The idea was eventually passed along to Bob Beichner for possible inclusion in his proposed Physics Central project.

Item: Anonymous peer review isn't very practical due to the closeness of the professional community…we can tell who wrote what pretty easily, or at least tell which research group a paper comes from. Might want to consider alternative review formats?

Item: We need more journals! There's too much of a bottleneck right now, and while there's plenty of places we **can** try to publish, it's somewhat diffuse. Plus, these alternative journals are often not interested in highly technical PER work, preferring things that can be put to use in the classroom immediately.

Wednesday, August 6, 2003

EH – PER: Teacher Preparation, Curriculum Development and Evaluation

EH01: Preparation of Tomorrow's Physics Teachers: Theory and Practice – Eugenia Etkina, Rutgers

Reformed program focusing on **pedagogical content knowledge**, students were future teachers in both physics and biology, in the Educational Master's Degree program (Ed.M). Since the reform, enrollment in the Ed.M With Physics certification program has jumped from 1 to over 10.

Used the ISLE method. Students proposed and/or were given models to analyze and accept or reject in a cycle of refinement. Certification involved six physics methods courses using the ISLE cycle, including a research internship in X-Ray Astronomy, followed by a clinical practice that included student-teaching.

Students initially emphasized content, organization and presentation, but eventually switched to pedagogy (what students know, how they think) after the coursework. The focus shifted from the instructor to the student.

EH02: Addressing the Needs of Pre-College Science Teachers: A Research-Based Approach – Donna Messina, University of Washington. (Note: this talk was really about two talks' worth of material crammed into one, so I probably missed important details as I frantically took notes.)

Comparing two different ways to preparing pre-service teachers in the Seattle school district via the Teaching/Learning Partnership between the schools and the Washington Physics Education Group.

Reduced time-on-task of the second version had a negative effect. They found it's not enough to focus on pedagogy, you must present sufficient content…can't assume they learned it all in college.

EH03: Developing a Research-Based, Inquiry-Oriented Dynamics Module for K-12 Teachers – MacKenzie Stetzer, University of Washington

Goal: help teachers understand physics content at a deeper level than the students need, relearn the material using Inquiry methods (pursuant to fixing the problems exposed in EH02).

Method: Develop a module (i.e. enough chapters to form a new book) for Physics By Inquiry that addresses N3L and the idea of force. ECR, especially the Force Follows Motion concept.

The simple tool used for measuring forces was a rubber band attached to a ruler at one end and a string at the other, a simple spring scale called a "pull meter" in the PBI module. The module seems to have done pretty well.

EH04: Revision of Introductory Physics Lectures – Ingrid Novodvorsky, University of Arizona (another very fast talk, hard to follow).

The sequence under consideration is a 4 semester calculus-based program for majors: Mechanics, Thermodynamics, E&M, Quantum Mechanics.

It had standard Lecture/Lab/Recitation timeslots in the course catalog, and they were pretty much stuck using them. Lectures were made more interactive, with questions assigned before each lecture based on readings. Conceptual questions were posed in lecture, usually 2-4 in a session with each question more difficult than the one before. The PRS was used to collect responses.

Many examples were quickly flashed through.

Gains were pretty good, there was evidence of "sticky" learning. Students did equally well on quantitative and qualitative items, and they liked it.

EH05: Research as a Guide for Developing an Inquiry-Based Curriculum on Mechanical Waves – Mila Kryjevskaia, University of Washington

17 densely-packed slides. About all I got out of it was that a PBI module on waves was being developed. Or maybe several modules. The impression I got in EH03 was that one module was an entire book, but in this talk the term seemed to mean chapter.

EH06: Teacher's Implementation of a Reform-Based Middle School Science Curriculum – Clarissa Bercovich, Center for Research in Mathematics and Science Education in San Diego.

"The Challenge of making sense of students' making sense."

The goal was to identify how teachers' knowledge and beliefs affected the quality of implementation of the Center's curriculum, CIPS (Constructing Ideas in Physical Science).

CIPS is an ECR-style curriculum, with "making sense" used instead of "resolve".

Teachers using CIPS for the first time were videotaped to identify difficulties. They tended to get the Elicit and Confront down pretty well, but had major issues with Resolve.

- Too many answers were given when students stumbled.
- They would draw out the right answer from one student, then stop and move to the next topic.
- Much of the talking was social chatter rather than working towards understanding.

- Alternate ideas were not discussed, merely dismissed. Students providing alternative hypotheses would be ignored in favor of students who could provide the right answer.

- They did not refer back to experiments.

The tapes were later shown to the instructors, who were interviewed.

- Teachers seemed to lack the ability to identify the influence of prior student ideas.

- Teachers' views of learning were gradualist rather than a sort of punctuated equilibrium. In other words, keep at it and they will slowly learn, rather than there being an "Aha!" moment that breaks a student through to a new understanding.

- Teacher view: Smarts + Time + Right Idea = Learning!

- Teachers were "panning for gold," not mining deeply for it.

Conclusion: Teacher development should focus on properly implementing the Resolve step.

EH07: Inquiry-Based Physics for Preservice Science Teachers – Andrew Wallace, Angelo State University (Texas).

Texas has recently changed the requirements for teachers, especially science teachers.

- Courses not covered in the high stakes testing have vanished.

- Teacher certification is designed to "teach to the test" for the TEKS test.

- It is now impossible to get certified in just Chemistry, Physics or Physics/Math, or to get any kind of secondary certification in any of these specialties in Texas. "Physical Science" now covers them all, making it necessary to be proficient in all three areas for certification. This "solves" the shortage in physics teachers by reclassifying it as part of "physical science" and assuming chemistry teachers can cover physics.

A new degree program has started, designed to service people attempting to get one of these new certifications. A Physics/Chemistry combined degree is being offered, with a common "core block" and then extras in one of the two fields to establish you as a major in one and minor in the other. There is no separate Physics or Chemistry minor anymore, the only way to minor in one is to major in the other and take the unified degree program.

Lecture has been reduced in all courses, replaced with more interactive methods (taking advantage of the overall shakeup to implement PER results). Multiple forms of assessment, including those from KSU's IDEA center, are used.

Students tend to have trouble with the technical and programming requirements, especially those coming from a Chemistry background. Java and HTML courses do not make you a programmer.

State funding has not increased commensurate with the new requirements, surprise, surprise. Training of in-service teachers is lagging behind due to shortages of manpower and money.

I left before EH08, sorry.

FB – PER: Qualitative and Quantitative Probes of Student Thinking

FB01: The Validity of Normalized Gain: Finding a Valid and Useful Measure of Instruction – Andrew Heckler, the Ohio State University.

This was not actually a critique of normalized gain as a measure, rather a statistical analysis of the robustness of <g> and how correlated it might be with confounding variables. A lot of data mining and simulations that had a large number of statistically significant results, but a much larger number of insignificant results, making me suspect NO results were real.

FB02: Reliability and Validity of Multiple Choice Exams in Large Introductory Physics Courses – Michael Scott, University of Illinois

Back in 1996, UofI switched to multiple choice exams for large classes due to concerns over fairness in grading with large sections. Since then, they've gotten a positively frightening N value for use in even-odd reliability testing of the test items (contributed by 51 professors and taken by 12000 students so far).

Reliability is good, with uncertainty being only about a quarter of a letter grade (i.e. if a student took two different tests on the same material using different items drawn from the question bank, the scores would likely be within 3% of each other).

Validity was tested by giving free-response versions of the exams to 33 volunteers and then interviewing them. The correlation between a student's multiple choice score and the grade given by "committee grading" on the free-response test was 0.78. Correcting for attenuation and heterogeneity (i.e. the volunteers tended to come from a biased sample, as they self-selected), the correlation approaches 1.0.

Conclusion: the multiple choice tests used at UofI are valid and reliable.

I asked whether students liked the MC tests, and the response was along the lines of "They haven't known anything else to compare to." Something of a non-answer, I'm afraid.

FB03: Using Concentration Analysis to do Research and Improve Learning – Genaro Zevala, Technologico de Monterrey, Mexico.

This work was based on the Redish/Bao research, using multiple choice items and expanding the categorization from two letters (HH, ML, etc) to three. A great deal of matrix algebra followed, out-Bao-ing Bao, but not seeming to add anything meaningful. Too much labeling for too little content.

FB04: A Method For Accounting For Errors in Model Analysis Theory – Steven Sommer, Southern Illinois University at Edwardsville.

This was a computer simulation exercise in which the programmer assumed errors were possible and then crunched numbers repeatedly to see what effects the errors would have. The method used found a quadratic relationship between initial error and the error in the final result was quadratic, but the method is impractical for matrices larger than 3x3.

During Q&A a member of the audience pointed out that a Monte Carlo analysis would do a much better job. Sommer was generally unable to answer questions of "what is it good for?" because he lacked physics and education backgrounds, and was doing this purely as a self-described "math geek".

FB05: Using Model Analysis to Investigate Individuals' Understanding of Lunar Phases – Rebecca Lindell, Southern Illinois University at Edwardsville.

"Understanding of Lunar Phases" includes things like being able to relate the phase to the moon's position relative to Earth and Sun, the direction phases change, etc. Students have a poor understanding of these things.

Study set out to determine what, if any, incorrect initial models the students held. Due to the large size of the classes, qualitative interviews could only scratch the surface, so it was decided to use quantitative methods and then use Model Analysis to attempt to get more out of these results.

Model Analysis was applied to the Lunar Phases Concept Inventory, an instrument designed specifically for use with Model Analysis. There were 10 options per multiple choice item (maximum…some items had less).

The results indicated that while some facets needed more instruction, students had correct views of others.

FB06: Cancelled (Leith Allen was unable to attend or have someone speak in her place).

FB07: Investigating Students' Learning of Electricity and Magnetism Through Model-Based Diagnostic Instruments – Substitute speaker whose name I didn't catch, the Ohio State University.

(Note: slides were badly reproduced and rather busy, making them hard to read. Apologies if I misinterpreted what I was seeing.)

Models are dynamic and generative representations of real, hypothetical or imaginary situations. Comprehension should involve the construction of good models.

Question: What are the models constructed in E&M? What context features affect the choice of models by students?

Answering this question would give us good predictive power. For instance, students tended to misapply their model of a conductor to insulators, so we might predict certain student difficulties based on this incorrect choice of models.

FB08: Imagistic Model Construction Process in Electricity – Melvin Steinberg, Smith College.

Process is essentially a reinvention of ECR, with new jargon. Advocates a gradualist position of design under constraint, descent with modification. Changes to student models should be smooth and unthreatening. Basically, the "confront" step should be less confrontational.

Students tend to think in terms of images (supported by brainscan data), so his method employs more visual representations, such as a reinventing of the "color for voltage" circuit diagrams seen in Physics By Inquiry vol 2.

FB09: The Use of Weekly Journals in Physics By Inquiry – Yuhfen Lin, the Ohio State University.

Background: In the phys106-7 Physics By Inquiry courses at Ohio State, students are given a weekly journal assignment. Sometimes it's about the material covered in class, sometimes it's about career plans or views of how the course is progressing. Among other things, journals are used to solicit feedback from the students without having them draw attention to themselves in class. Additionally, there's homework, pretests from the PBI manual and a Question of the Day.

The journals have been determined to be a good source of formative feedback in both directions. But…students often fail to recognize that this is happening.

Part of the feedback in the journals reveals that while students are impressed by PBI and would use it in their own classes, they would remove all the things that make it different from traditional lab (they would give The Right Answer, make everyone follow at the same pace, etc). The students still equate teaching with the reciting of Right Answers, even after experience with PBI.

The students **did** learn via PBI, but it was despite their prejudices. Those prejudices were not altered by their success in learning a new way.

FB10: Speaker did not show up.

FB11: Linguistics and the Epistemology of Physics – David Brookes, Rutgers.

The theory part of the talk was moved to a PERC poster.

According to Max Black, metaphor is a filter that **creates** similarity rather than merely invoking a pre-existing similarity. Language creates the ways we think about things.

Physics education can suffer from inaccurate metaphors, but benefit from good metaphors. In other words, metaphors are very important, not just "chrome".

Examples were presented of both good and bad metaphor use by students (did not invoke Skander from CI02, but Skander's earthquake metaphor's a good example of one that's between good and bad).

We often seek "compromise" metaphors to bridge the gap when two competing metaphors are equally good at describing a phenomenon (i.e. wave and particle metaphors for subatomic stuff: Schrödinger used a "smeared particle in a wave" while Bohm et al used "discrete particle guided by wave.").

FB12: Gestures as Evidence of Student Thinking About Physics – Rachel Scherr, University of Maryland.

Gestures can help us understand student thinking as it happens. And not just gestures with pre-coded meanings like thumbs-up and so forth, more important are spontaneous and possibly unconscious gestures. There is a rich literature on, for instance, different ways we shrug.

A video of students working on a problem was presented, with emphasis placed on the gestures taking place during their discussion. This showed how important gestures can be to understanding the subjects.

- Gestures tend to be more holistic than speech.
- Gestures are often vital to understanding spoken words (i.e. a transcript would be totally misinterpreted without knowing the gestures that go along with the words).

- Gestures are often more compelling than words, convincing others of an idea where simply telling them fails.

- Instructor intervention may elicit helpful gestures from students. However, it can also interrupt gesture-based lines of thought, so be careful!

- A gesture has to come at the right time to have an effect (video showed a key gesture appearing several times before it had its impact).

Moral of the story: Pay attention to student gestures! Before we can articulate something in words, we often gesture our way towards understanding. Don't dismiss "handwaving answers" too quickly.

FK – PER: The Practice of Analysis as a Window on Theory (Invited Session)

FK01: Empirical Investigations of Student Understanding and the Development of Instructional Materials – Paula Heron, University of Washington.

Due to the length of session FB, I missed the beginning of this talk (there was about a minute to run across the hall, but after sitting in FB for three hours I had to take a few minutes break first).

The gist of the talk seemed to be that students can have a correct idea but be unable to logically defend it, but I may be misreading a minor point as major.

FK02: The Questions We Ask and Why: Methodological Orientation in PER – David Meltzer, Iowa State University.

Objectives: Find ways to help students learn better, help them improve reasoning and problem-solving skills.

Model: Redish's archery target. Consider an archery target, a series of concentric colored circles. The outer ring is white, the middle ring is gray and the center bull's-eye is black.

White: Random noise, areas where the student knows nothing at all.

Gray: Partially known information, sometimes correct, sometimes complete.

Black: Well known information, student is always correct and well-connected.

The Gray zone is the most useful to consider. Black is difficult to improve since it's already very good, and while it's easy to probe the Black zone, it's not too useful, since you won't find anything the student has trouble with. White is hard to do ANYTHING with, because it's so diffuse and even empty…there's nothing to find and nothing to work with. The answer to "What are student difficulties" is "everything" in the White zone. But the Gray zone, the boundary region between knowledge and ignorance, is fruitful both for probing and treating.

Therefore, one should try to map the Gray boundaries. Find where the stable islands of Black might be, probe the linkages between pieces, etc. See how students in the Gray zone respond to guidance and collect an ensemble of data.

This model was applied to a Thermodynamics project (CCLI grant, calculus-based Engineering Physics course). Thermodynamics was found to be a good Gray zone for the sample group.

32 volunteers out of a 494 person class were interviewed, a self-selected group of high achievers.

Interviews did focus on:

- Baseline questions
- Diverse contexts (such as cyclic processes)
- Learning difficulties
- Gauging stability and resilience of student views

Interviews did not focus on:

- Where students got their knowledge
- Learning styles

A large chunk of the sample was unable to answer fundamental questions raised in class. There was a strong belief in Work = 0 for any cyclic process, confusing a state process with a cyclic one. Students exhibited learning difficulties rather than having true alternative theories. There were **no stable alternative models** found in the interviews. The "Net work is zero" view was dislodged once students were asked to draw P-V diagrams, so it was not a robust view. **Common ≠ Robust**.

It is necessary to address incorrect ideas, but while they may be somewhat resistant to change, they are not full-blown alternative theories. Rather, we tend to force vague ideas into focus on the spot. Our cloudy understanding may simply "precipitate" into the same incorrect statements from person to person.

Some common "precipitates" are to confuse Heat with State-Process Energy, or to confuse change in Temperature with change in Heat/Energy. These statements are "new" to each student at the time they are made, and even if every student makes such statements, they are arrived at independently.

(Gears shift a bit here.)

Any interview takes place over a period of time. How does one deal with changes in the student over the course of the interview? **Assessment is a dynamic look at a dynamic process**. It is necessary to have "ensemble" values for the results of your interviews: statistical summaries, qualitative assessment of data elements, etc. Student understanding is metastable, but however hard this can make interviews, interviews are still **real data**.

FK03: Jargon, Insight and Researcher Beliefs: A Theory-Based PER Study – Andy Johnson, Black Hills State University (South Dakota).

Alternate title: The Role of Theoretical Considerations in PER

Theoretical perspectives influence all research.

Theory – System of knowledge and beliefs about <u>something</u>.

Theoretical Perspective – Tied more into personal beliefs and ontology. This defines both What Is, and how things are and should be connected. Will be abbreviated TP from hereon.

TPs rest on foundational assumptions; they're "frames" in the sense of the talks at Redish's birthday session at the Winter 2003 meeting.

It is difficult to step outside our own minds and into a new TP.

Constructivist TP (CTP) elements:

- Everything we know, we figured out.
- Sense-making is important.
- You cannot say that any two people share a common model.

A strong CTP can deny much of what we need to do in research, as CTP is fundamentally at odds with Objective Realism (i.e. There's real stuff out there that we can measure, and everyone will measure the same thing).

A more research-friendly TP is held by many researchers, and is characterized by Johnson as the "Not-A-Theory" Theory. It has the following elements:

1) My observations are good enough.

2) Any probing I do is non-perturbative enough.

3) Any information I ignore was unimportant anyway.

Johnson claims all three of these points are invalid.

1) It's hard to invalidate a model you already have, either as a student or a researcher. We will see what we expect to see, through the filter of our possibly unacknowledged model. Our observations are not good enough.

2) Cognition is strongly context-dependent. Explanations are generated on the spot depending on the question, the perturbations drive the responses. "Knowing" is an emergent process.

3) It's too easy to miss the really important stuff (like, say, gestures). We have to narrow our focus **before** we know where the focus should be, and we use our TP to determine this.

Objectivity is impossible, entanglement with a TP is inevitable. So be aware of your TP, get a good one, and make sure people know both your TP and your reasons for using it. There is literature on TPs that can offer much to PER, and we can use TPs as a tool rather than as a mask.

In part two of the talk, Johnson discussed how he picked three TPs and endeavored to "get inside" them as much as possible while analyzing some student data. All three (Facets, Distributed Cognition and Sociocognitive) gave different but still useful results. To use a metaphor that he didn't, the more blind men you have examining the elephant, the more likely you are to figure out that it **is** an elephant. TPs can be our internal "blind men".

FK04: Discussion Panel – Paula Heron, David Meltzer, Andy Johnson. Querents noted where I could identify them.

Q1 (Rachel Scherr): Asking other two panelists if they have TPs. Meltzer agreed that he had a "working model" if not necessarily a full TP, as did Heron. Neither was eager to label their viewpoints as actual theories, and claimed Johnson had an overly broad definition of theory. Heron added that she had a point of view, guiding principles and a perspective, just not a theory.

Q2: How much do your goals affect your TP (assuming for the moment that you do have one)? Heron claimed her goals were too modest to do much to any perspective, but they did help her orient herself. Johnson brought up a sort of "which came first, the chicken or the egg?" analogy for goals and TPs, they tend to feed back on each other.

Q3 (David Hammer): Asked Heron for clarification of something from earlier in her talk. Heron declined on the grounds that it would take too long and offered to clarify later in private communication. She also warned against overgeneralizing from her talk.

Q4: More comment than question, "We have met the enemy, and they is us." Thanks to our TPs, we are as blinkered as our students.

Q5 (Bill Reay): Why isn't there PER being done on graduate-level courses? General consensus was that it's hard to do, the sample sizes are really low, we're still a new field, and there's a lot of resistance from faculty against messing with the precious upper level courses. Noah Finklestein pointed out that it **is** on the agenda, however, and I noted that some has been done (such as applying Physics By Inquiry's ideas to graduate Electrodynamics at Ohio State…it was shot down by the resistance angle).

Q6 (Noah Finklestein): How do models evolve? "Wow, deep." Meltzer said that as goals and data evolve, so do the TPs. Follow-up Q: What about incompatible TPs? How do you reconcile them? Johnson replied that there's the world, and then there's the way you look at the world…different slices through the orange. The TPs may not actually be incompatible when you step back, since they are both based on the world.

Q7 (Seth Rosenberg): Are TPs the same as Kuhnian paradigms? (Seth also, as an aside, thanked Andy for the talk, noting that it's dangerous to let TPs remain unexamined.) Johnson replied: not really, as paradigm is social while TPs are personal. TPs and paradigms can interact, but aren't the same thing. Disclaimer: he's not attacking any particular research tradition or group with his comments about the Not A Theory theory.

Q8 (Joe Redish): TPs are largely descriptive, might different TPs be more complementary than contradictory? Johnson said "Yep." This was also kinda the point of the second part of his talk.

I did not take notes at the PERC, for two simple reasons. One, there's already the PERC proceedings, which should cover most of the juicy bits. Two, it's a **lot** harder to take notes at poster sessions and discussion groups, and I was already getting kinda worn out from writing the preceding notes.

See you at the Summer 2004 meeting!