Thanks Keith for that wonderful introduction. In fact, Keith inspired me to start singing along with my guitar playing, because he did it with such gusto, and so I continued in that tradition … without the corresponding skill.

So, I’m going to talk about some work on bringing Cognitive Psychology to schools, and really, in some ways the title should be turned around to ‘bringing schools to cognitive psychology’ because there’s a two-part story here. On one hand, we are applying laboratory Cognitive Psychology research to improve school learning. On other hand, what our Pittsburgh Science of Learning Center (see LearnLab.org) has been doing is using technology and interactions with schools to do better Cognitive Psychology— in particular with more external validity that extends the internal validity of laboratory work. So that’s the general theme.

Keith mentioned the impact of our cognitive tutoring technology. The most successful has been a full course for high school algebra that is based on computational models of student thinking and learning that come out of Cognitive Psychology research. As Keith mentioned, the Cognitive Tutor Algebra course has been widely disseminated. We spun off a company in 1998, Carnegie Learning Inc., which has continued to disseminate the course to thousands of schools. It was a very field-based research project from the beginning, where we interacted with math instructors and math educators and followed the national council of teachers of mathematics’ standards in designing the course. It brought in the research base from both Cognitive Psychology and Artificial Intelligence.

Given this is the APS meeting, I wanted to give you a quick example of the Cognitive Psychology research behind the Cognitive Tutor Algebra course. This is my favorite audience quiz. The question here is, ‘what prior knowledge do algebra students have, and can we predict, given our hypotheses about prior knowledge, which of these problems is going to be most difficult for a beginning algebra student?’ So you see, these problems all have the same mathematical content, but one is stated as a story, one a word problem, and one is an equation. Given shortage of time I won’t make you go through the quiz, but real quick, who thinks the story problem is going to be the most difficult? How about the word problem? How about the equation? So that’s a pretty even distribution. That’s a good sign, a more even distribution than usual!

Teachers usually say the story and the word problem are the most difficult. That’s what I thought when I designed this study. But in fact, the equation was the most difficult. We replicated this result in a number of studies. These are beginning algebra students with beginning algebra problems. Basically, there’s a kind of language learning explanation here, where algebra can be thought of as a second language. Given a simple algebra problem in English, a student can understand it and show some thinking strategies to
solve it without translating it into the “Greek” of algebra. But, if they are given the problem in the Greek of algebra, rather than in English, they have greater difficulties.

We have done formal surveys like this with teachers. Most teachers, and if you look at most math books, they are designed as though you start with the equations because that’s what’s thought to be easier. Then you go into the story problems. So these results changed the way we designed the Cognitive Tutor Algebra course. There’s a lot of emphasis on having students analyze real-world problem scenarios, use different kinds of representations and tools, and all along the way there’s this tutoring system that can solve the problems, that has knowledge of typical student mistakes, can follow along, catch misconceptions, give students feedback, give them hints when they’re stuck. It can also track the concepts and skills that students are applying here and give students feedback to adapt new activities to help them on the things they need most, and move students along who have already mastered the concepts and skills and don’t spend extra time on those.

So those were the tutors, and in that research we did large-scale field studies of their effectiveness in enhancing student learning. Often these were quasi-experimental designs, but many such studies have been done over the years and often in urban schools. This slide shows a result from one of our early studies in the Pittsburgh area. In general we saw very large results on problem-solving assessments and multiple representation use, and statistically reliable but smaller results on standardized tests. So, building on that success and that’s the “bringing Cognitive Psychology to the classroom” part, when we created this new Pittsburgh Science of Learning Center about two and a half years ago, one of the ideas there was to use Cognitive Tutors, and technology-enhanced courses in genera, as a research platform. So that’s the second part of the story: I want to describe how we can facilitate better psychology research on learning (and computer science research on advanced technologies) through this research platform.

The Pittsburgh Science of Learning Center (PSLC) has these interlocking research and development activities shown in this slide. The key idea is this notion of what we call in vivo learning experiments. These are experiments that are part of a real class, and they’re usually a couple weeks long -- much longer than a typical laboratory study of learning. Typically, the control condition is ecologically determined, that is, the ‘as-is’ instruction in the current course. The researcher has no choice or option to make a straw man control condition. You have to beat the real thing. The treatment condition(s), then, could involve changing the teaching practices or the material or the technology. So that’s the study part of PSLC’s interlocking activities. It is essentially a subject pool for learning researchers to more easily in vivo experiments, but unlike the usual Psychology Department subject pool, this one involves students in a real course and the studies are executed in the context of that course.

So what are these courses? We have agreements with sites using different courses, two in math, two in science, and three in language. The site agreements are like a research hospital agreement, where we work out, through a committee process, ways in which studies can be run in those classes. You can come to our center and have a study run (see learnlab.org), just as a medical researcher can go to a research hospital and do a study a
new medical treatment with real patients. Each of these “LearnLab” courses is in multiple high school or college sites. While the Algebra Cognitive Tutor course is over 2,000 schools across the country, we don’t have such agreements with all of those schools, just about four or five of them. But we have enough sites that we are satisfying the demand so far for research studies. As demand increases, we will increase the number of sites.

Let me give you a quick sense of some of the technologies. In the French LearnLab course, there’s a unit on French culture and intercultural awareness. Students watch videos of French people interacting, for instance this French boy and Arab shopkeeper, and then discuss their observations in an on-line discussion board. But, students are often not noticing critical features in these videos -- like that the shop is open when French-owned stores would not be and some corresponding racial tension between the boy and shopkeeper. So, a computer tutor was created that is interleaved with video watching and prompts students to extract these critical features. Technologies in other LearnLab courses include a physics tutor, an algebra tutor, a chemistry virtual lab, etc.

This diagram is meant to represent how PSLC facilities researcher-school interactions. Rather than each researcher going into a school on their own, the LearnLabs are kind of central resource where you can come to do a school-based study and we have the connections with the schools to make it happen.

You are probably wondering about this toothbrush image here. A major component of the PSLC is an effort to a shared theoretical framework around learning principles to improve academic learning. How many of you have heard the toothbrush story of theories in Psychology? Just Art and a couple of you. “Theories in Psychology are like toothbrushes: Everyone has one, no one uses anyone else’s.”

So it is a tough issue, even for our set of researchers with this amount of money, getting people to agree on theories or even on shared vocabulary. And it’s not just that you agree, it’s that you really use it in the same way. For instance, when we say ‘explicit instruction’ for instance, we may think we all mean the same thing, but subtleties pop up. We find that one person means verbal instruction and another means any active effort (like pointing) by an instructor or instructional media. Using the wiki technology, like Wikipedia, is a great way to resolve those subtleties. We get people to put their positions in writing in a live and easily accessible form. Others can than review, correct, or extend. And the whole history of changes in stored. So we have built this PSLC wiki where every one of our studies is summarized in a structured format. Each study page has sections for independent variables, dependent variables, findings, explanations, etc. These pages use terms in the PSLC theoretical framework and these terms are hyperlinked to glossary that provides encyclopedia-like descriptions of these terms. The wiki an evolving, organic document that now has about 250 pages that were created in the last nine months by PSLC members. We are planning to open this up in the next year or so to the research community at-large.
So part of this theoretical framework for what we call “robust learning”. Robust learning is the idea that our courses are not just after learning as reflected in “normal post-tests” or final exam -- tests just assess those things that you were instructed on and right after the instruction. Instead, robust learning measures address long term retention and transfer, and even accelerated future learning.

The PSLC is seeking a theoretical framework that gets beyond these grand debates and false dichotomies that are shown in this slide. We want to get beyond toothbrush theories and get beyond knowledge-lean laboratory studies. We want to see whether principles from the laboratory apply in the knowledge-rich environments of academic learning. It is likely that learning is significantly different in situations where participants have lots of relevant prior knowledge, like classroom learning, than it is in knowledge-lean situations, like that in many laboratory experiments where stimuli are designed to make participants blank slates and, purportedly to isolate the learning processes. But, the processes that are so isolated may well be quite different from those operating in academic learning whereby drawing on a rich prior knowledge base is a matter of course.

So, we want instead of these debates and dichotomies and toothbrush theories, an experimental basis for resolving those debates with a shared set of concepts and terms. Studies in PSLC are associated with one or more of our three research clusters. This table shows the studies in the “coordinative learning” research cluster. Each row is a different study. The columns are different instructional variables being employed in one or more of these studies. These variables include instructional techniques like prompted self-explanation, tutoring feedback, personalization, feature focusing (which I illustrated with the French Culture tutor), and others. A ‘v’ in a cell means that the study (in the row) varies that instructional technique (in the column), that is, the treatment condition employs it and control condition does not. A ‘b’ means both the control and the treatment employ that technique. So you can see that many of these instructional techniques are being experimentally tested in variety of different contexts. That is, the treatment is competing against different kinds of control conditions (some that already employ innovative instructional techniques). This kind of cross-context variation greatly increases the external validity and generalize-ability of the results.

The columns also indicate whether the study involves or benefits from a deep domain analysis of the “knowledge components” should and do acquire and/or from a detailed analysis of how instructional variations change learning processes. Further columns indicate which dependent measures of learning are employed including our three robust learning measures I mentioned above.

Let me give an example of couple of these studies. A key motivation is to help students avoid “shallow learning”, like this geometry student who seems to think that if angles look equal in a diagram, he can conclude they are equal. We see these kind of shallow inferences all the time in academic. To get more robust learning, one of the kind of hypotheses that we and others have explored, is to interleave example study with problem solving, that is, to have students go back and forth between studying solution examples
and attempting to generate solutions on their own. The theoretical rational is that such instruction minimizes, on one hand, the shallow procedural encoding that can result from too much isolated problem solving practice and, on the other hand, the limited attention and illusions of knowing that come from too much example study.

One of these studies was done in the Chemistry LearnLab course using a new Cognitive Tutor that was created using our Cognitive Tutor Authoring Tool or CTAT. A quick aside: CTAT is available for free at ctat.pact.cs.cmu.edu and facilitates tutor authoring whether by advanced AI programs or even non-programmers. You can build a tutored activity in an hour or so.

Back to the Chemistry experiment on interleaved example study and problem solving. Students in the treatment condition watched a video describing a correct solution of a chemistry problem, which was followed by some comprehension monitoring prompts where students had to provide an explanation of keys steps in the solution they observed. Next they solved a similar problem with the help of the tutor. Then, they studied another example and solved another problem and so forth. Students in the control condition simply solved problems with the tutor. The study was run in at Chemistry LearnLab sites at one college and two different high schools. The results revealed a large difference in learning efficiency, whereby the interleaved condition took significantly less time to learn than the problem-solving condition. There was no difference between the conditions in performance on the post-test, but an important and significant reduction in time to learn.

Another PSLC study explored the same issue in the Geometry LearnLab course. Actually, in this study, some German researchers took our geometry tutor and translated it to German. In the interleaved example condition, students first are given the numerical answer to each step in a problem solution and they have to explain why the answer is correct by choosing an explanation from a menu. Next, they must start providing the numerical answers themselves. In the problem-solving condition, students must come up with the numerical answers to each step themselves all problems. Here again there was a big learning efficiency result, with the treatment taking significantly less time to learn and doing just as well on the normal post-test items. But, in this study there was also a conceptual transfer result, such that the interleaved example students were better able to answer difficult questions about the geometry concepts.

Students in this study were requested to “think aloud” while they worked through the instruction. This data supports the interpretation that the interleaved example students were better engaging declarative memory of geometry principles and concepts than the problem solving students. Their verbalizations were more likely to address “why” to perform a step whereas the problem-solving group said more about “how”, indicating more procedural and surface-level encodings and associated knowledge components. Such knowledge works fine on a normal posttest, so that explains why you get the same result on the normal posttest. If the researchers did not use robust learning measures, like the conceptual transfer test, we might conclude no difference in learning. But instead the transfer test indicates deeper declarative encodings of the knowledge that allow for more flexible knowledge application.
Many, many other such studies are being performed within PSLC. One in the Physics LearnLab explored whether instruction should provide students with explanations of physics solutions or ask students to provide them. Result: Ask students to provide them, don’t give them. An interesting study in the English LearnLab: If I give students texts, they can learn vocabulary by example, and they do so more likely if I give them texts on topics they are interested in. A Chinese LearnLab study was inspired by machine learning results on co-training, where a learning system given two different kinds of sources of input can learn without feedback. The study found human learners in some ways behaving like such machine learning systems and in other ways differently.

Some conclusions: Educational practice should be research-based as medical practice is. We need to test educational practices to determine when they work and when they don’t. We need to connect the “toothbrush theories” coming from different labs and experimental paradigms and create a more coherent theory that has practical implications. The Pittsburgh Science of Learning Center is taking steps toward these goals. We’re focusing on this robust learning issue on tests that matter, on a replicable, practice-based theory. We are providing building this LearnLab facility to help bridge the research-practice gap and allow people to do careful, rigorous principal testing in controlled studies, but in real courses as a part of real classes.

(New Speaker) We have time for a couple questions.

(New Speaker) I just wondered if you had any, either in your research or in communication with other people, any connection with younger grades, middle school or elementary school, and how approaches might work. I know this is all high school.

(Lecturer) We do have a Bridge to Algebra course where we have been working with middle schools. There is a lot of other work going on. My colleague David Klahr is doing scientific discovery strategies at the middle school and elementary school math level. I could give you other such examples.

(New Speaker) What are the differences on these tests, so that racial group differences or male/female group differences…does this decrease the differences for the cognitive tutor?

(Lecturer) Well, if you look at the treatment results of a disadvantaged group and compare them to the control advantaged group then we see gap reductions, but basically what we see is that everybody goes up. I think that’s kind of consistent with the general observation that school is bringing students up, and its summer where the gaps widen, where there’s differences in home environment, which maybe relates to the last discussion and we can do some things about that, maybe with educational games.