Interactive Support for Mathematical Reasoning and Metacognitive Judgments of Learning

Ken Koedinger
Director of the Pittsburgh Science of Learning Center

Human-Computer Interaction & Psychology
Carnegie Mellon University (CMU)

Chasm between learning sciences & educational practice

Dimensions of today's challenge

- Empirical dimension
  - Lots of rigorous principle-testing lab studies
  - Lots of realistic classroom design research
  - *Too few experiments combine both*
- Theoretical dimension
  - *Almost as many theories as there are results!*
- Practical dimension
  - Great ideas in field, ignored in lab
  - Strong scientific results being ignored in field

Pittsburgh Science of Learning Center (PSLC) Purpose Statement

Leverage cognitive theory and cognitive modeling to identify the instructional conditions that cause robust student learning

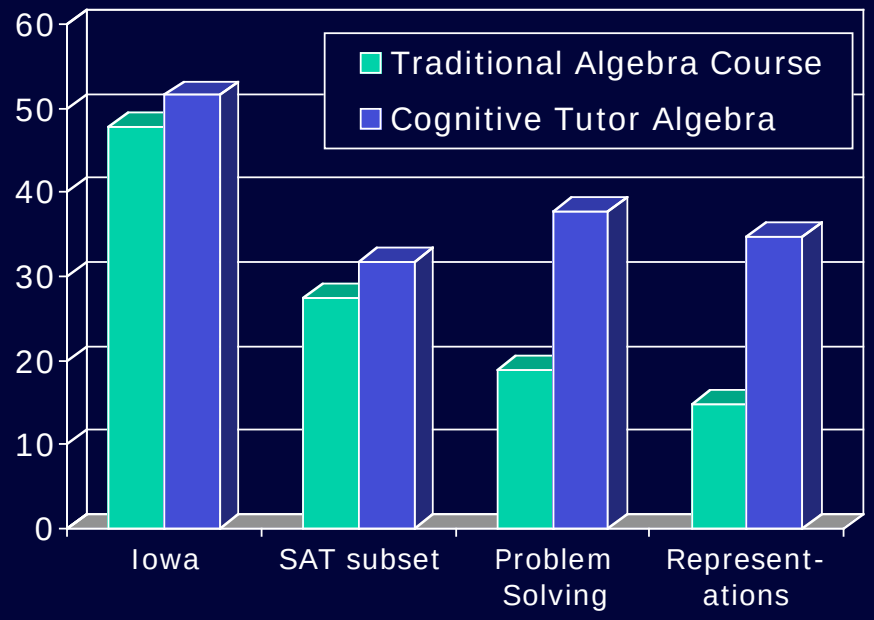
Overview

- Background: Cognitive Tutors & PSLC
- Assistance Dilemma
  - Argument for less
  - Argument for more
- Example-problem dimension -- cognitive & metacognitive issues

Real World Impact of Learning Science

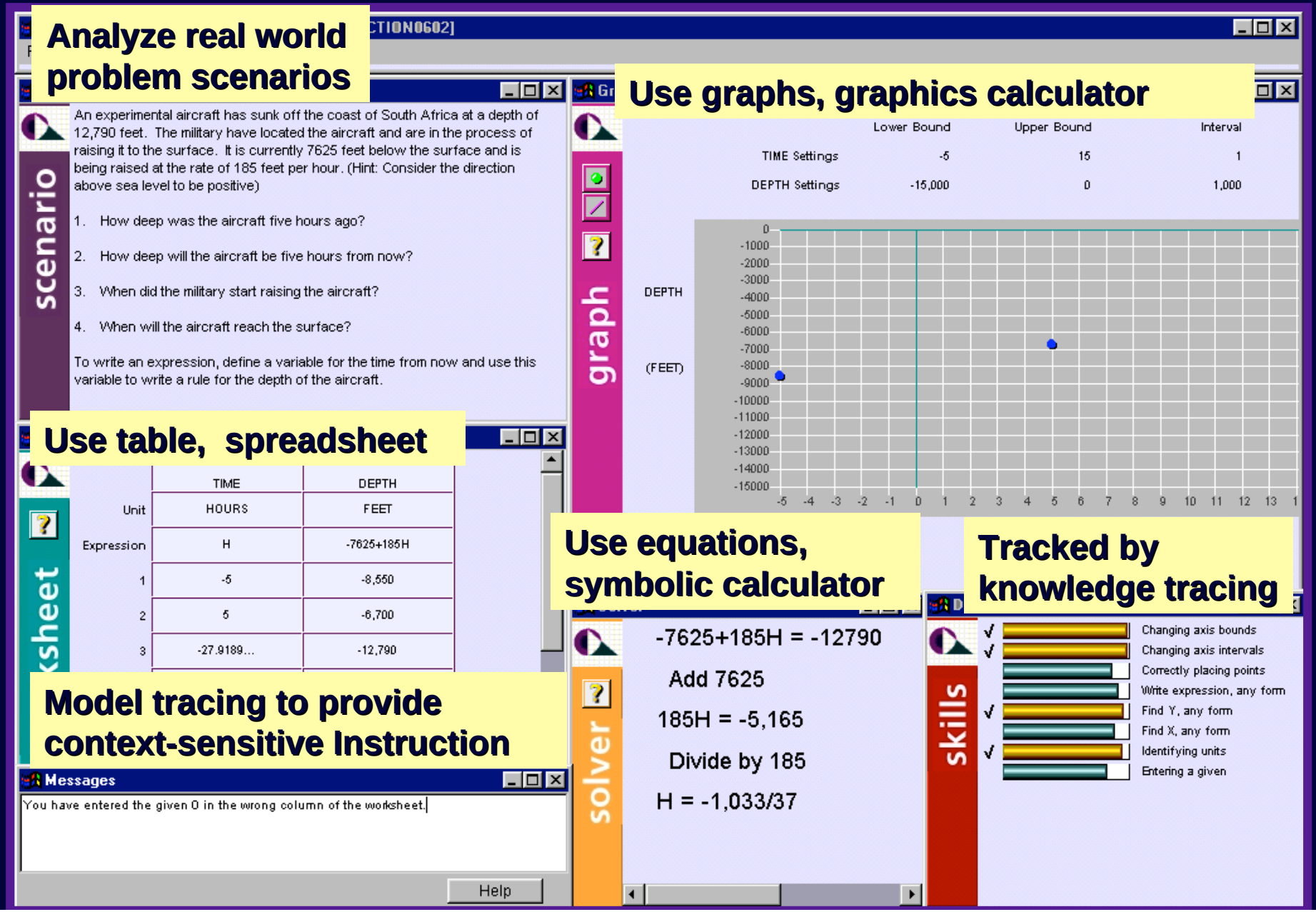
Cognitive Tutor Algebra course

- Based on cognitive theory & AI models of student thinking & learning
- Most widely used & evaluated Intelligent Tutoring System
  - > 2600 schools
  - > 10 full-year field studies demonstrating better student learning



Koedinger, Anderson, Hadley, & Mark (1997). Intelligent tutoring goes to school in the big city.

Algebra Cognitive Tutor Sample



Prior achievement:
Intelligent Tutoring Systems bring learning science to schools

A key PSLC inspiration:
Educational technology as research platform to generate new learning science

Logic of Pittsburgh Science of Learning Center (PSLC)

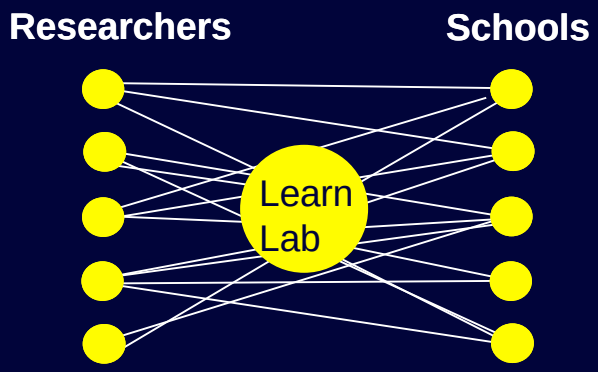
- Support experimental studies that
  - Test fundamental principles, not whole courses
  - Are internally & externally valid
- Create a theory of “robust learning”
- Leverage technology & computational modeling

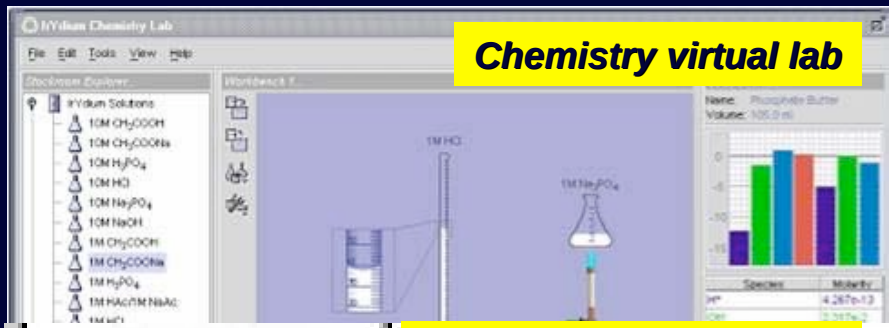
A Bet About the Future

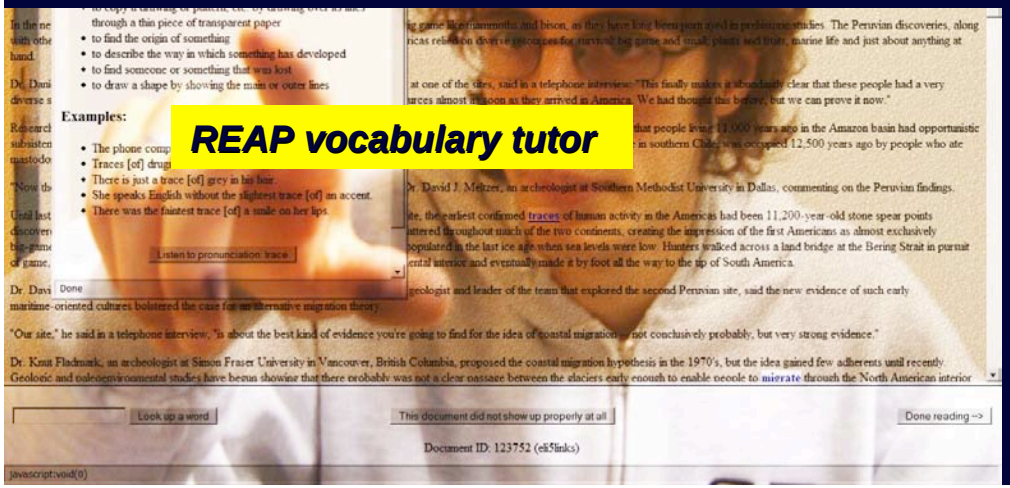
- The key to the 21st century university:
- Technology!
- And not because of the direct benefits of technology
- But because of vast data on learning & fast feedback to instructors, designers, administrators, researchers …

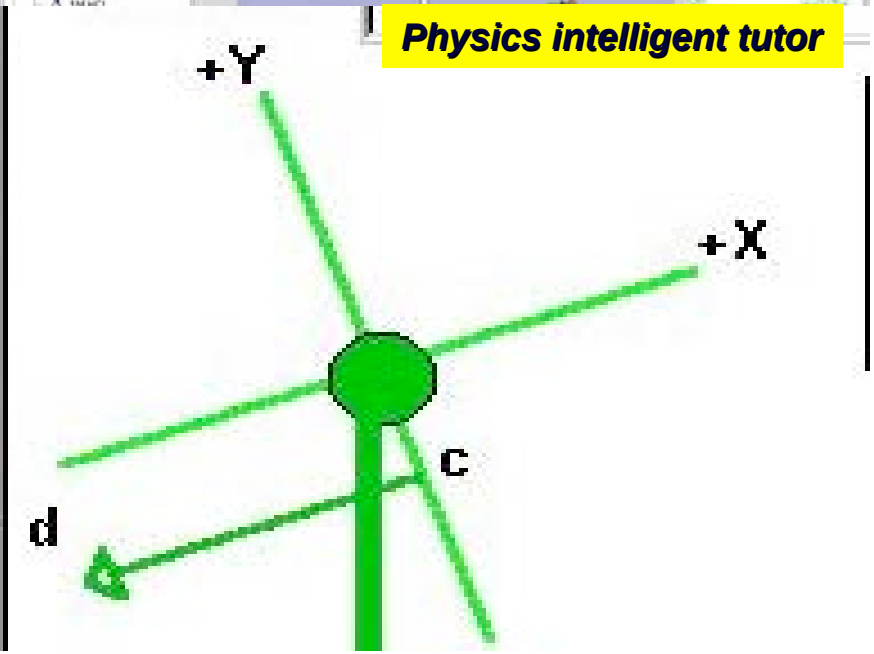
LearnLab: Like a research hospital for learning

- Tech enhanced courses in Science, Math & Language
- Agreements with schools
  - *Instrument* for continuous embedded assessment
  - *In Vivo Experiments* test principles for achieving robust learning









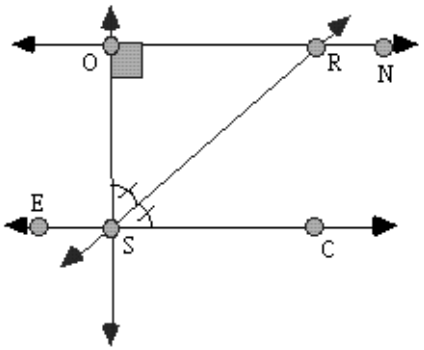
Example *In Vivo* Experiment on "Self-Explanation"

- Self-explanation: Have students explain steps in solutions
- *In vivo* experiments: Tightly controlled principle-testing experiment embedded in a real course

Aleven, V. & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26(2)*

Problem Solving Condition
(Ecological control: Tutor as it was)

External Angle & Parallel Lines



Given: ON ll EC. If the measure of Angle SOR is a right angle, find the measure of Angle SRN.

m<SOR	90
m<OSC	90
m<OSR	45
m<ESR	135
m<SRN	

Hint Done

Messages

Some reasons dealing with parallel lines are highlighted in the Glossary. Which of these reasons is appropriate?

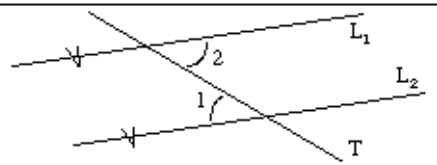
You can click on each reason in the Glossary to find out more.

More...

Glossary

Converse of Isosceles Triangle (Theorem)
Isosceles Right Triangle
Triangle Sum (Theorem)
Linear Pair
Linear Trio
Parallel Lines --- Corr. Angles Are Cong
Parallel Lines --- Alt. Int. Angles Are C.
Parallel Lines --- Alt. Ext. Angles Are C.
Parallel Lines --- Int. Angles on the Sa.

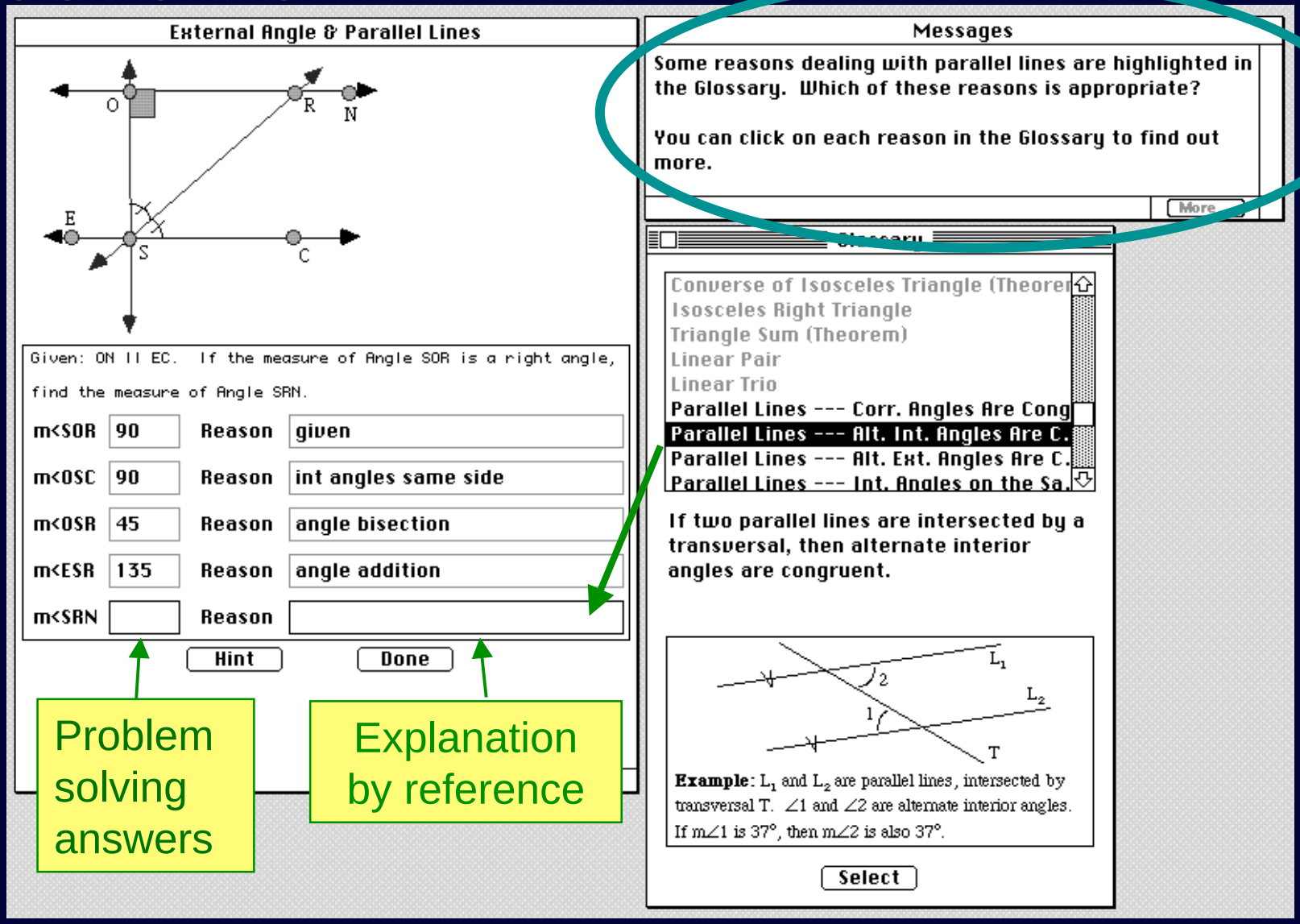
If two parallel lines are intersected by a transversal, then alternate interior angles are congruent.



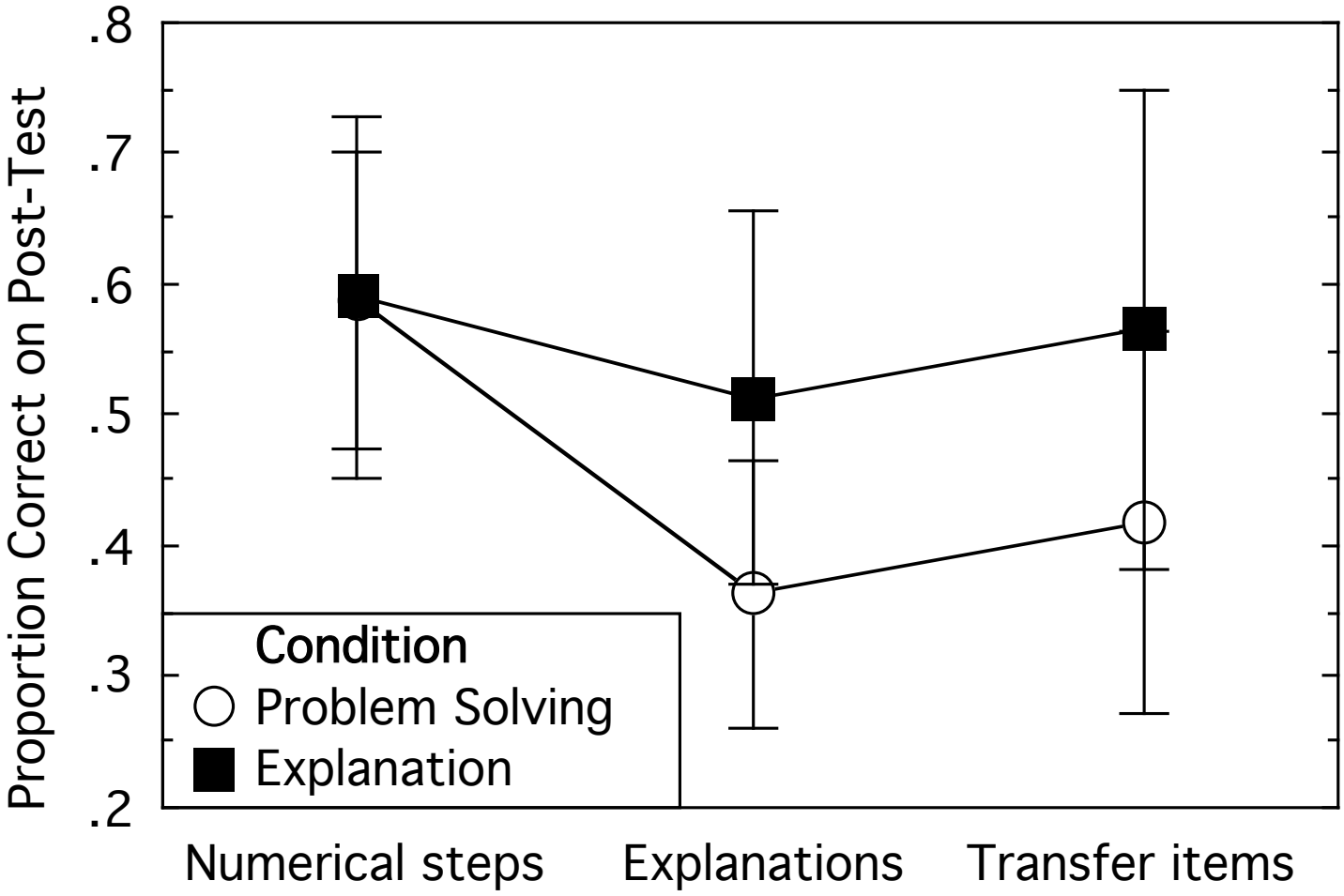
Example: L_1 and L_2 are parallel lines, intersected by transversal T. $\angle 1$ and $\angle 2$ are alternate interior angles. If $m\angle 1$ is 37°, then $m\angle 2$ is also 37°.

Select

Explanation Treatment Condition



Self-explanation yields better transfer



The *"Assistance Dilemma"*

Instructional designer's dilemma:

When should instruction *provide* students with assistance

vs.

When should it *withhold* assistance & elicit student knowledge construction?

- Fundamental unsolved problem in the learning sciences
  - Defines a design space for instruction
  - Experiments & theory to find areas of space where robust learning is maximized

Assistance Dilemma: Whether to give or to receive?

	Benefits	Costs
Giving information or assistance	Efficiency of communication	Shallow processing Does not engage LTM
Withholding information or assistance	Generation effect Engages & structures LTM	Cost of errors Floundering, confusion, wasted time

The no pain, no gain argument => more assistance

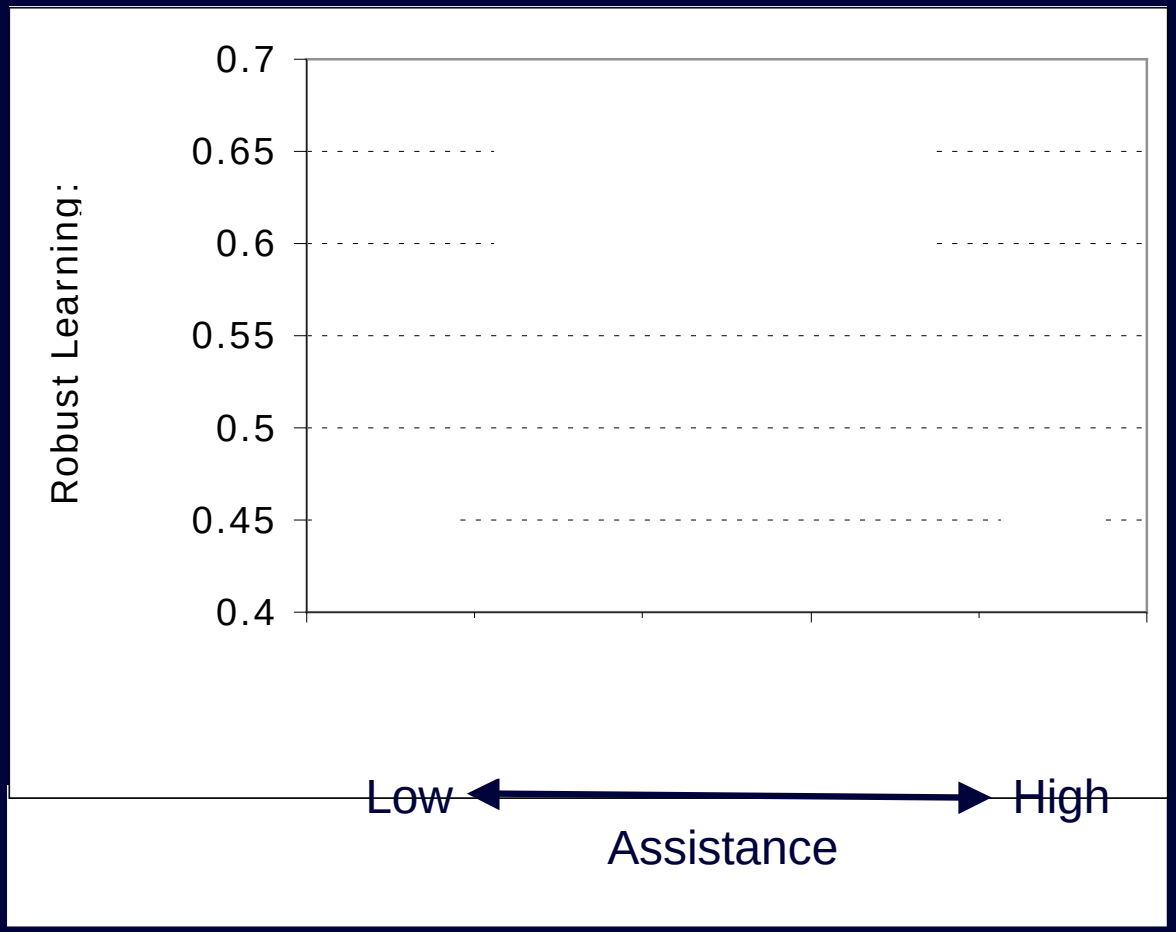
Many researchers & results argue for *lower assistance* place greater demands on students

- "Desirable difficulties" (Bjork)
  - Interleaved practice, delayed feedback, wide spacing
- Abstract examples (Kaminski & Sloutsky)
- Prompting for self-explanation
- Invention as preparation (Schwartz)

Testing effect

- “Tests enhance later retention more than additional study of the material, even when tests are given without feedback.
  This surprising phenomenon is called the *testing effect”*
  – Roediger & Karpicke, 2006

Testing Effect Example
Karpicke & Roediger, 06



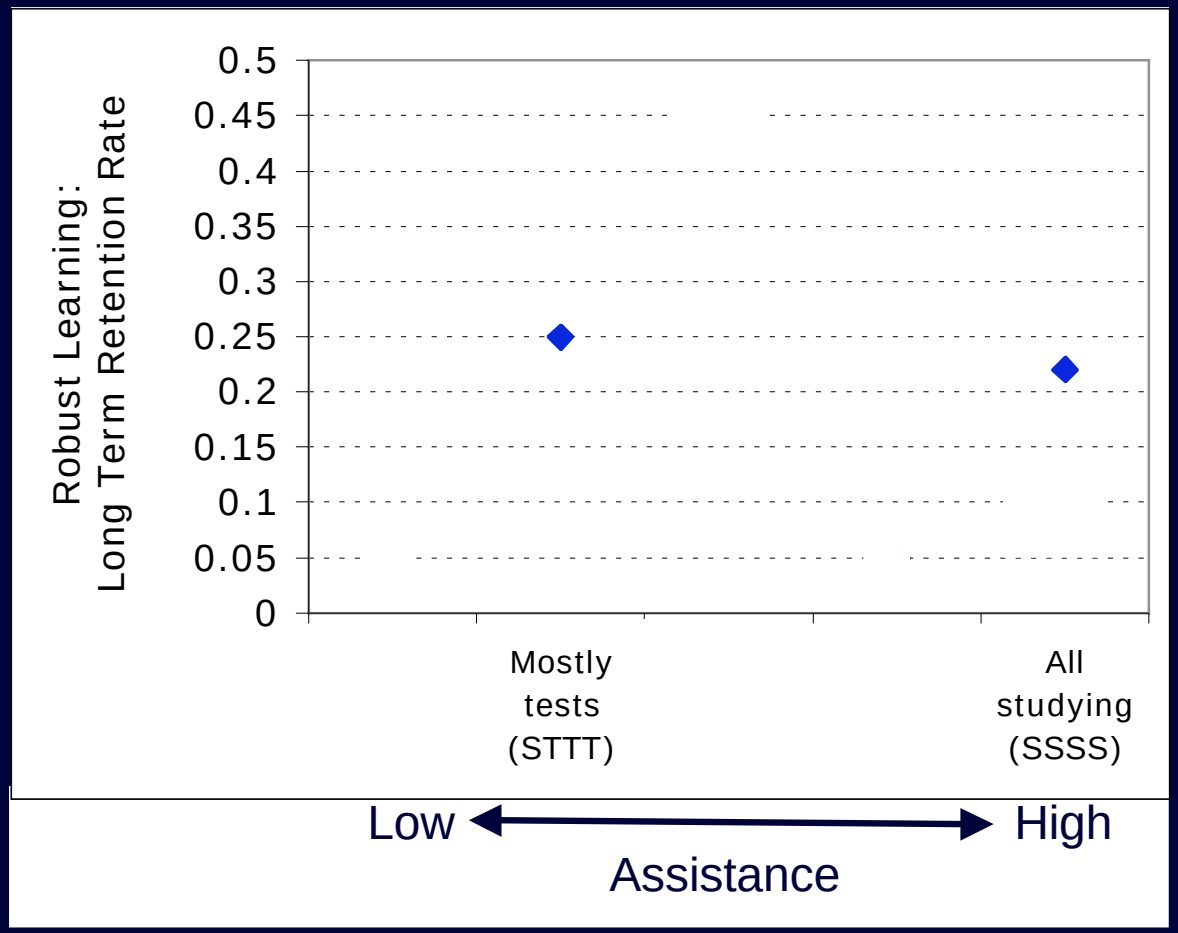
Memory task: Study & recall a list of 40 words

- "Tests enhance later retention more than additional study of the material" (R&K, 06)

Yes | and no

Testing Effect Example 2

Thompson, Wenger, & Bartling (1978)



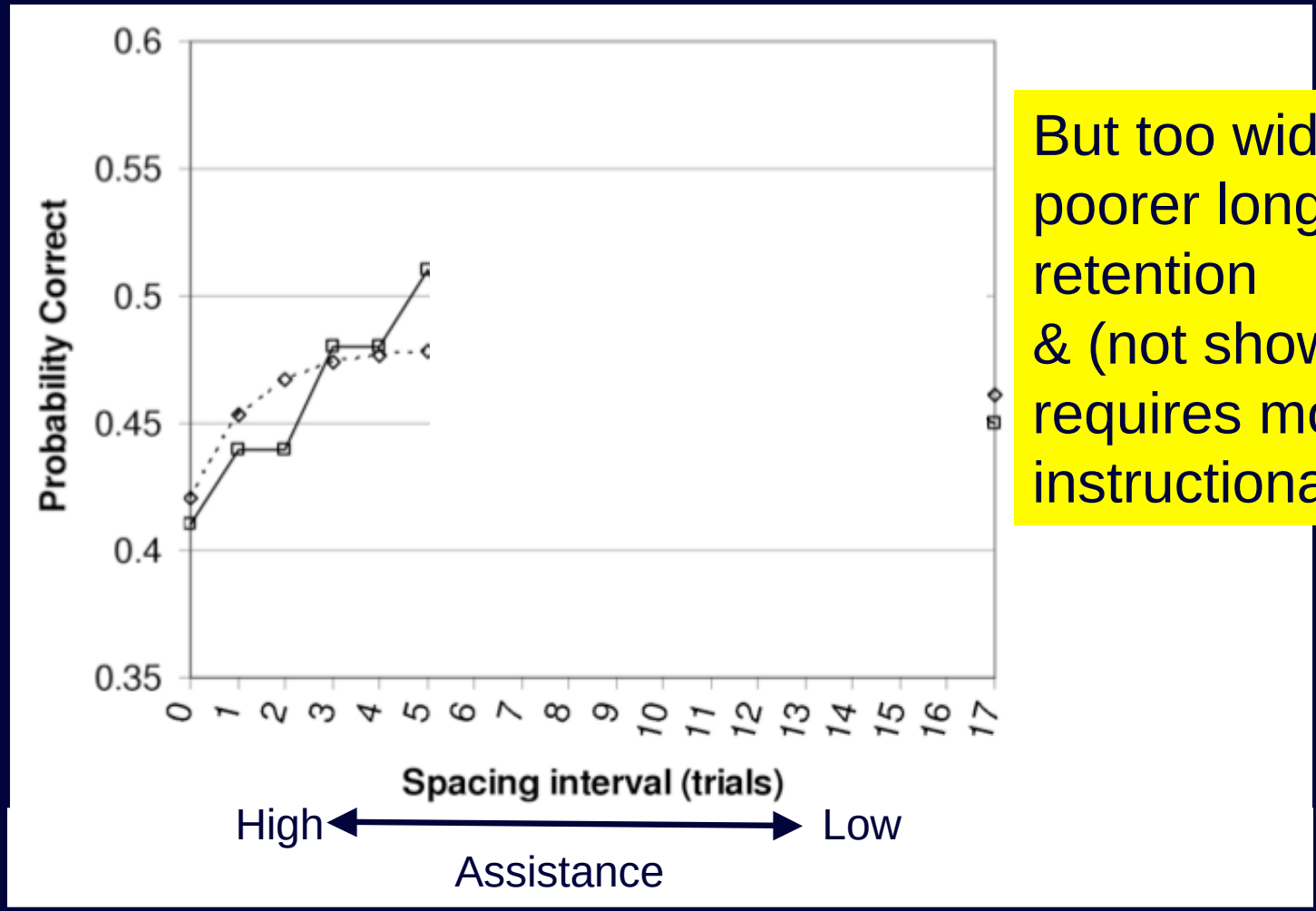
Memory task: Study & recall a list of 40 words

F trials were tests where failure yielded feedback: a study trial

- "Tests enhance later retention more than additional study of the material" (R&K, 06)

Yes? | and no

Spacing Effect: Wider is better (lower assistance) …



But too wide yields poorer long-term retention
& (not shown) re-study requires more instructional time

Pavlik, P. I., Jr., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. Cognitive Science, 29, 559-586.

Is it “no pain, no gain”?

Or “less pain, more gain”?

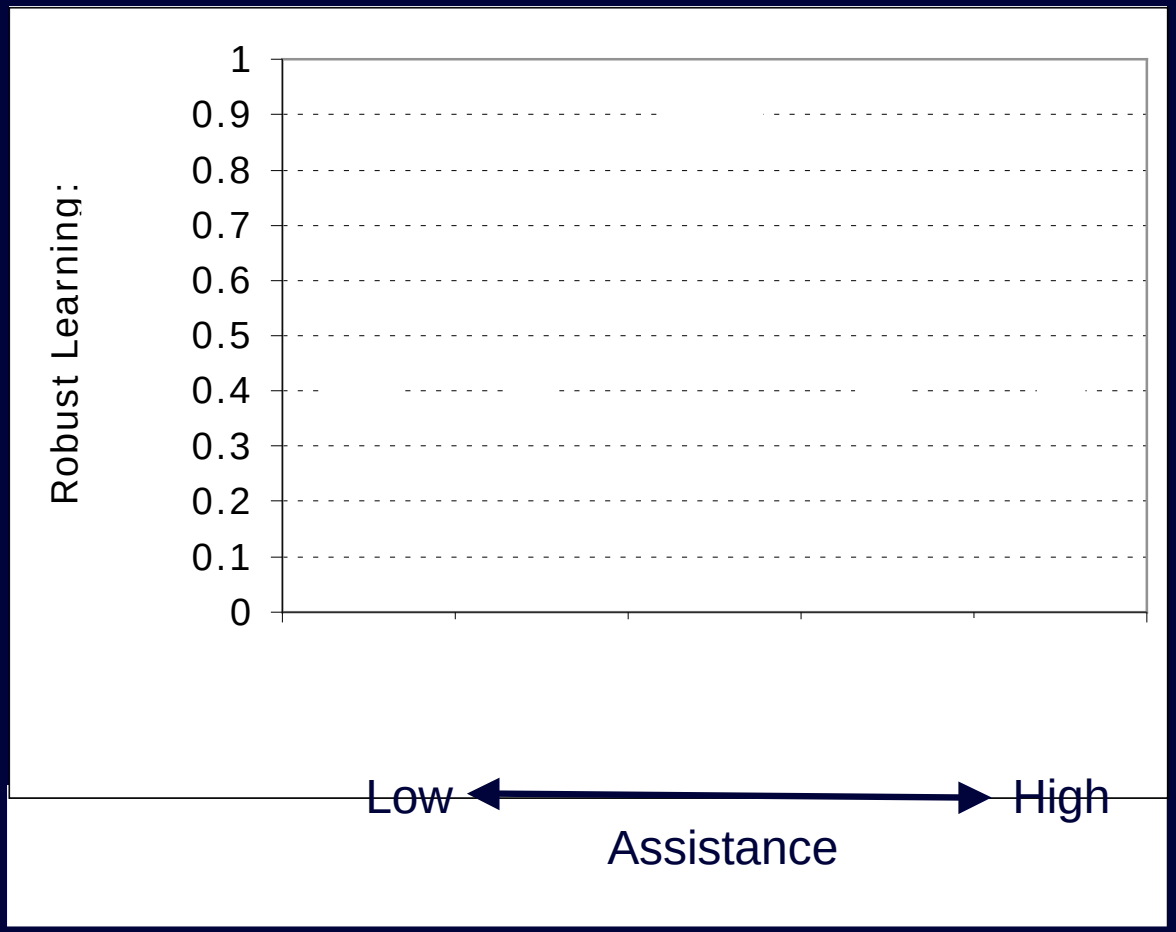
- Arguments for higher assistance …

The less pain, more gain argument => more assistance

Other researchers & results argue for *higher assistance*

- Direct instruction (Behaviorist, Klahr, …)
- Cognitive Load Theory (Sweller, Mayer …)
  - Reduce “extraneous” cognitive load
  - Mayer’s multimedia principles
    - Modality, continguity principles ….
  - Sweller et al.
    - No-goal problems, worked examples …
- Provocative paper:
  - Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, … (Kirschner, Sweller, & Clark, 2006).

Worked Examples
Sweller & Cooper



Comparison

- 8 Problems

Solve (a+b)/c = d for a

Vs.

- Example-problem pairs

Example:

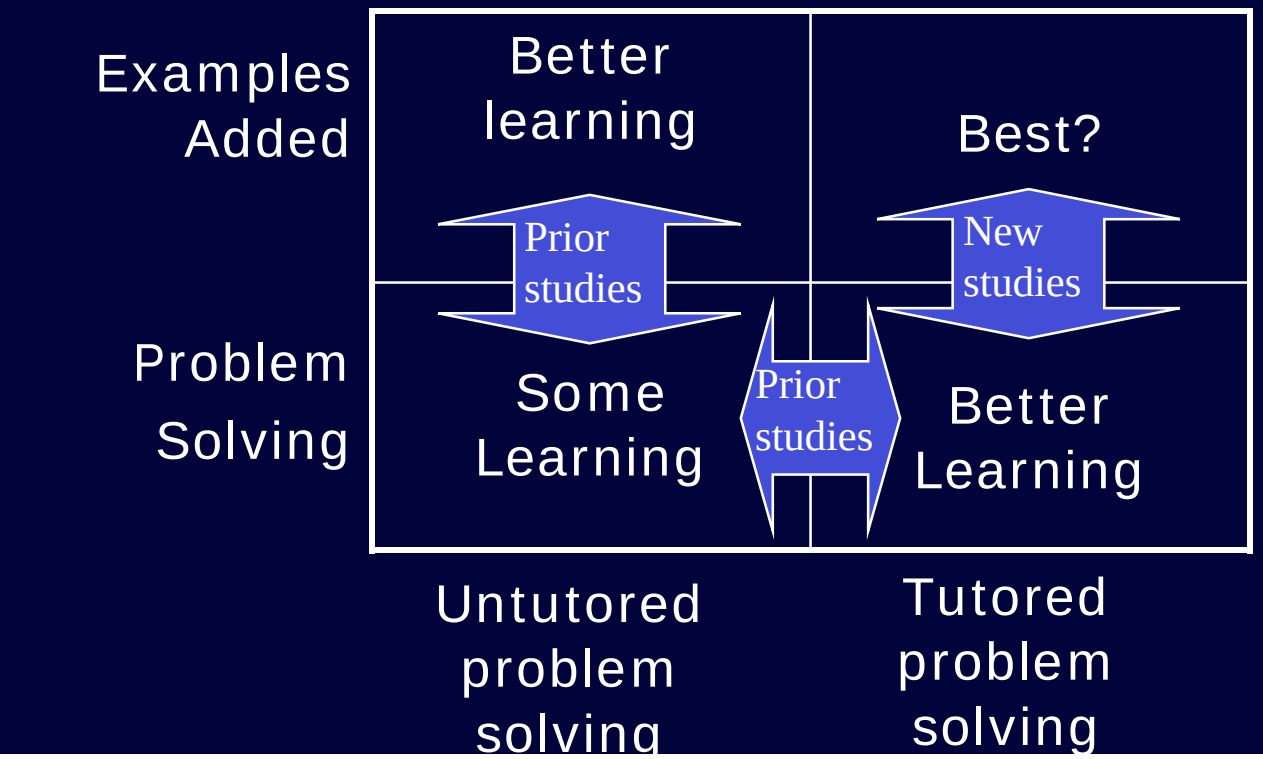
(a+b)/c = d
a+b = dc
a = dc - b

- "a worked example constitutes the epitome of strongly guided instruction" (K, S, C, 06)

Yes | & maybe?

Can worked example principle beat a strong control?

- Prior work: Added worked examples to *untutored* problem solving
- New: Add examples to *tutored* problem solving



PSLC *In Vivo* Studies of Worked Examples in Intelligent Tutors

- Chemistry tutor studies
  - Replacing half problems with worked examples yields more efficient learning -- same outcome in ~20% less time
- Algebra
  - Less time, better long-term retention
- Geometry tutor studies
  - Not only less time, but better conceptual transfer

Carnegie Learning's Cognitive Tutor

File Edit Tutor Go To Window Help

Circles problem Unit 1 Section 1

Major Arc BFD 2

Look Ahead Problems Look Back Progress

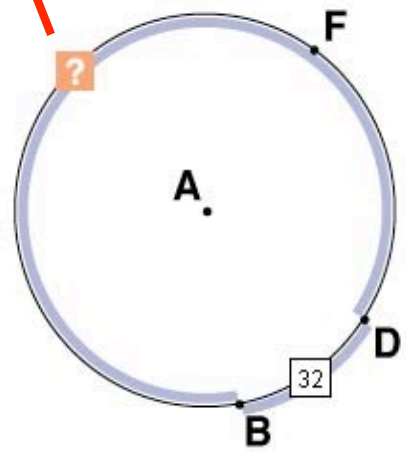
DIAGRAM

Given is circle A with arc BD.

If the measure of arc BD is 32°, what is the measure of arc BFD?

m Arc BFD =

Rule =



Ecological Control = Standard Cognitive Tutor

Students solve problems step-by-step & explain

Carnegie Learning's Cognitive Tutor

File Edit Tutor Go To Window Help

Circles fixed Unit 1 Section 1

Major Arc BFD 2

Look Ahead Problems Look Back Progress

DIAGRAM

Given is circle A with arc BD.

If the measure of arc BD is 34.7°, what is the measure of arc BFD?

Treatment condition:
Half of steps are given as examples

m Arc BFD + m Arc BD = 360 degrees

m Arc BFD = 360 degrees - m Arc BD

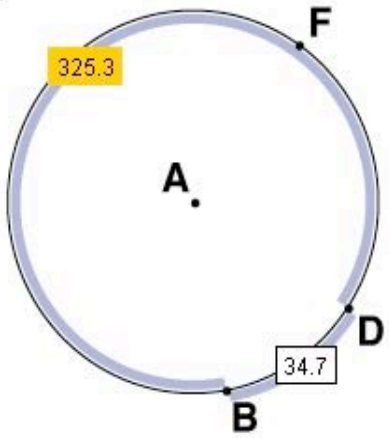
m Arc BFD = 360 - 34.7

m Arc BFD = 325.3

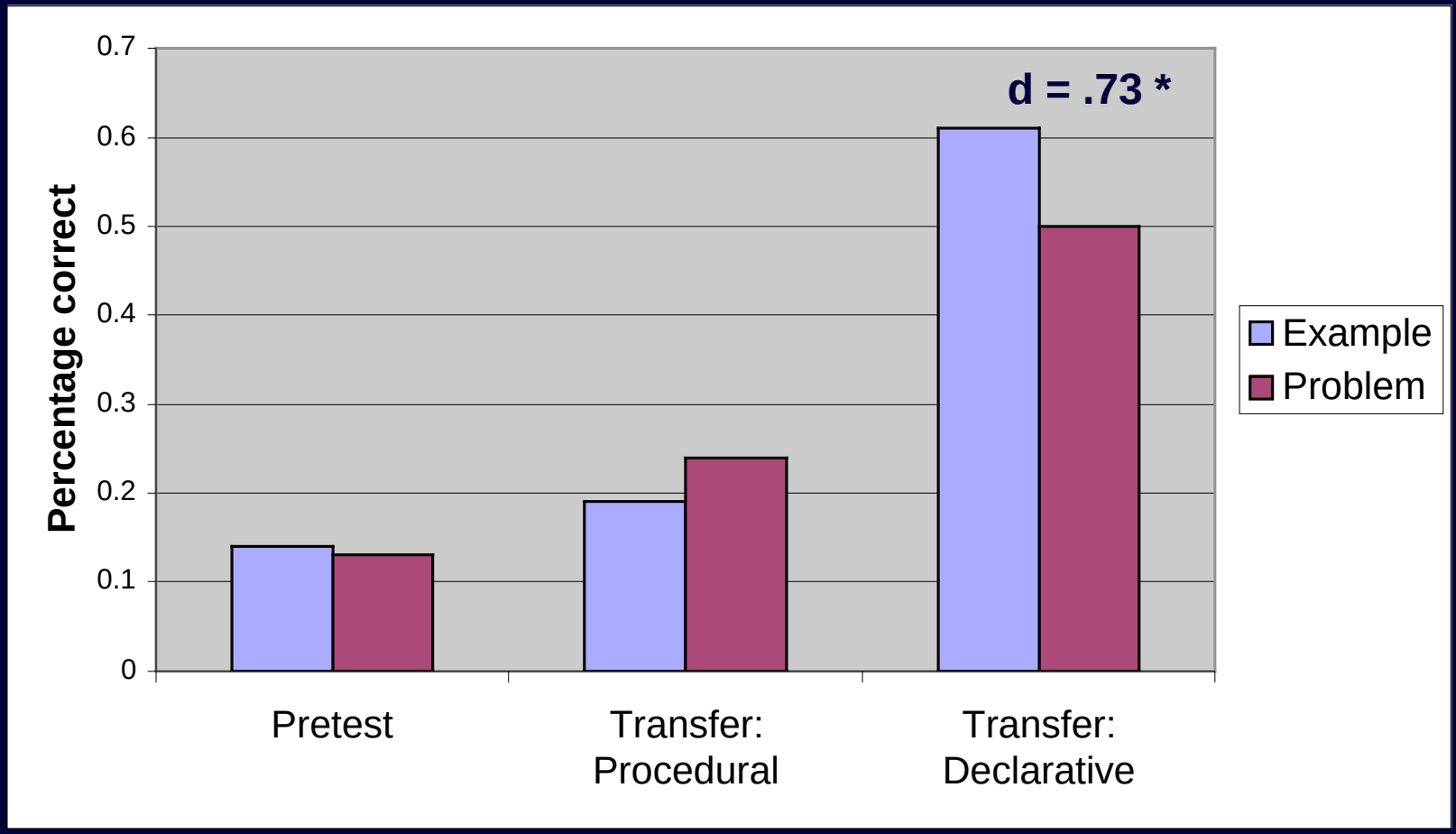
Rule =

Worked out steps with calculation shown by Tutor

Student still has to self explain worked out step



Result: Better conceptual transfer



Think aloud data: Indicate different modes of thinking

- Problem group: Explained how
  - "one can compute the measure of arc EF by subtracting 33.3 from 360"

  => more procedural learning

- Example group: Explained why
  - "This is a major and minor arc, this means the sum of both is 360 degrees"

  => more conceptual principle learning

Theory: Does Cognitive Load Theory Account for Result?

- Tutors are designed to minimize load
  - Tutor provides step-by-step feedback, often gives goal structure to students
    - Mediates goal processing load attributed to problem solving
- Nevertheless students seem to benefit
  - Better to present examples before rather than within problems
- Why?

Why examples before rather than within problems?

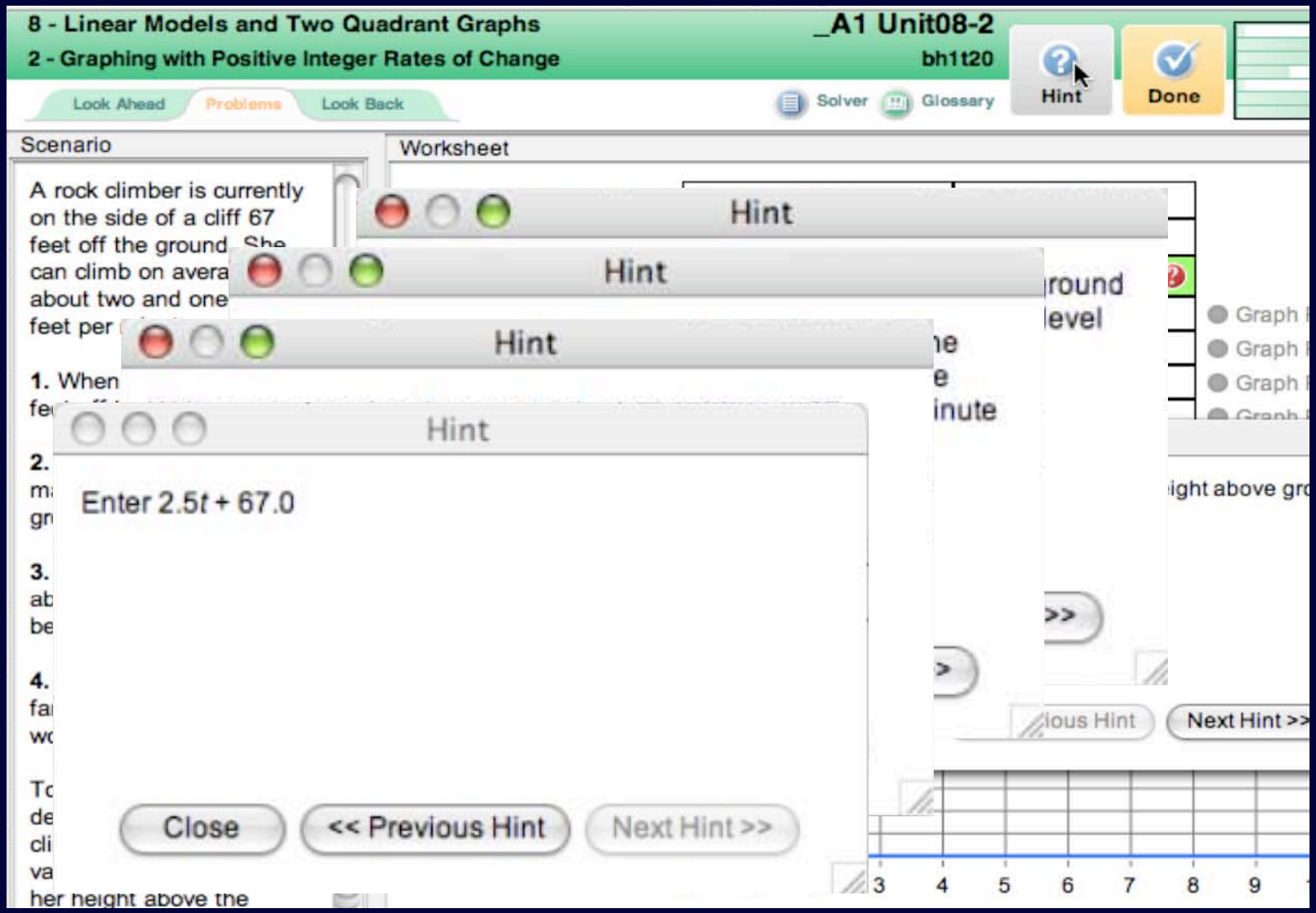
- Cognitive Load Theory not adequate
- Alternative: Metacognitive frame of mind evoked by problem solving
  - More engineering than science mode of thinking (Schauble; Miller)
  - More performance oriented than learning oriented (Dweck; Elliot)

Tutor Data Mining Evidence

- What happens when students get a "bottom-out" hint?
- …

Shih, Koedinger, & Scheines (2008). A response time model for bottom-out hints as worked examples. In *Proceedings of the First International Conference on Educational Data Mining*. **[Conference Best Paper]**

“Bottom-out” hint in Cognitive Tutors => on-demand example



Individual differences in self-explanation

After bottom-out example:

- Some students enter answer & move on
- Others seem to reflect, engage in self-explanation
  - Measure: spend more time than usual

- *Students who reflect more, learn more!*
  (R = 0.48)

A step toward resolving assistance dilemma ... with very broad impact implications!

	Instructor options: Give	Instructor options: Elicit
Explicit rules	*Lecture*	*Self-explanation*
Implicit: Example solutions	*Worked example*	*Homework*

- *Current instruction*: Gives rules & elicits solutions
- *Better instruction*: Gives solutions & elicits rules/concepts

Final Thoughts

- Cognitive Tutors bring learing science to schools, but also …
- Ed Tech as “Hubble telescope” for learning research!
- Assistance Dilemma is fundamental unsolved problem in learning science
  - Inverted-U function worked out for some dimensions of assistance
  - Using in adaptive systems that optimize for robust learning

END 1

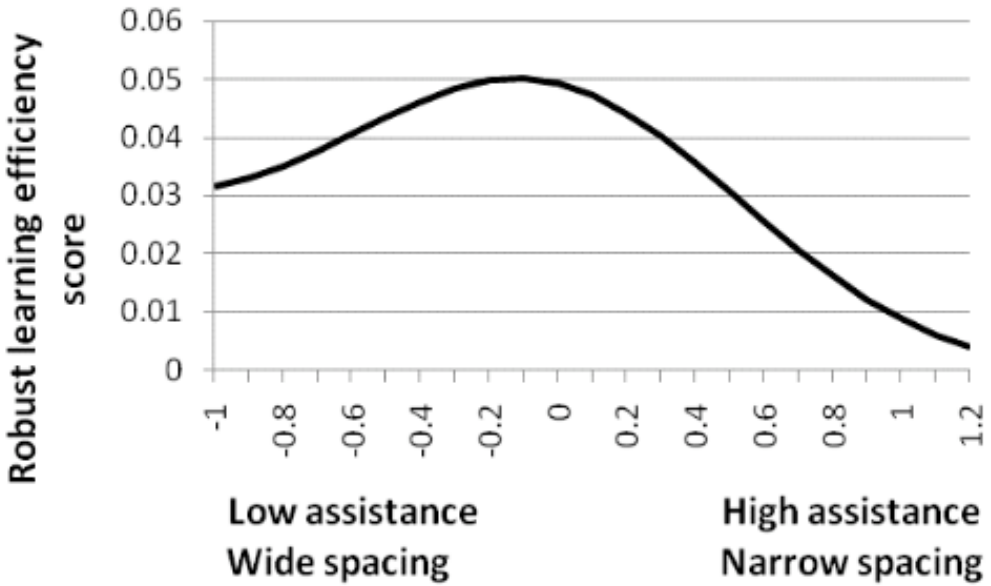
Assistance Dilemma Summary

Instructional support	*Poor* learning outcome	*Good* learning outcome
High assistance (less demanding)	crutch	scaffold
Low assistance (more demanding)	undesirable difficulty; extraneous load	desirable difficulty; germane load

Need predictive theory => engineer more effective instruction!

Formulating trade-off => adaptive technology to optimize learning

$$eff_m = \frac{p_m b_{suc} g_m + (1 - p_m) b_{fail} g_m}{p_m (t_m + fixedsuccosts) + (1 - p_m) fixedfailcosts}$$



eff_m = efficiency of robust learning
$p_m * b_{suc} * g_m$ = learning from success
$(1-p_m) * b_{fail} * g_m$ = learning from failure
$p_m (t_m + fsc)$ = success time
$(1-p_m) ffc$ = failure time

m = activation of fact
p_m = probability of recall success
b_{suc} = gain from success
b_{fail} = gain from review after failure
g_m = long-term increase in activation
t_m = time of recall
fsc = time for success
ffc = time for failure

General form of “assistance formula”

For each *learning event:*

Robust learning efficiency gain =

$$\frac{p * \text{benefit-of-success} + (1-p)*\text{benefit-of-failure}}{p * \text{cost-of-success} + (1-p)*\text{cost-of-failure}}$$

p = Probability of success *during* instruction

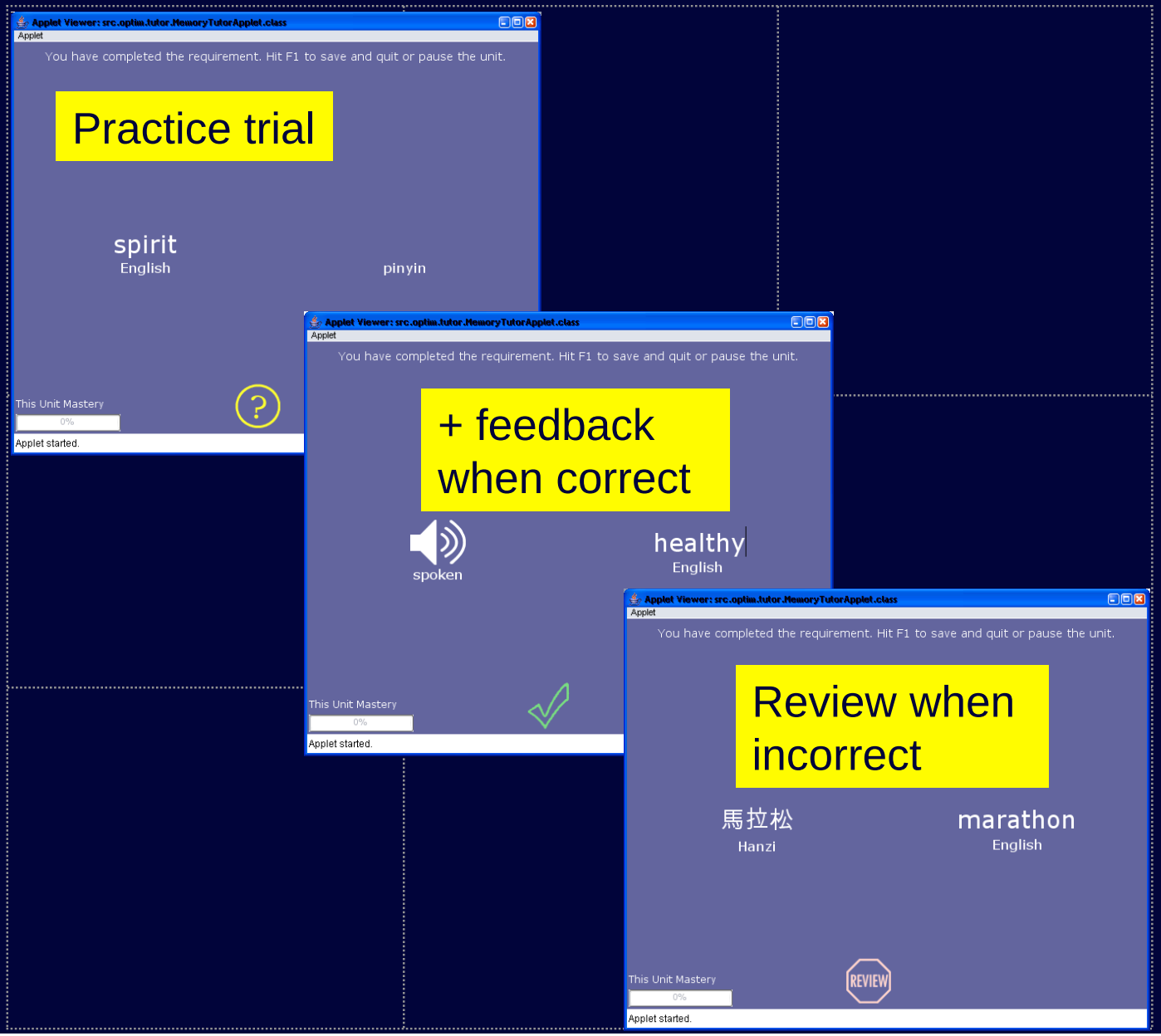
Pavlik's Tutor for Fact Practice

Different practice tasks:

English->Pinyin

Audio->English

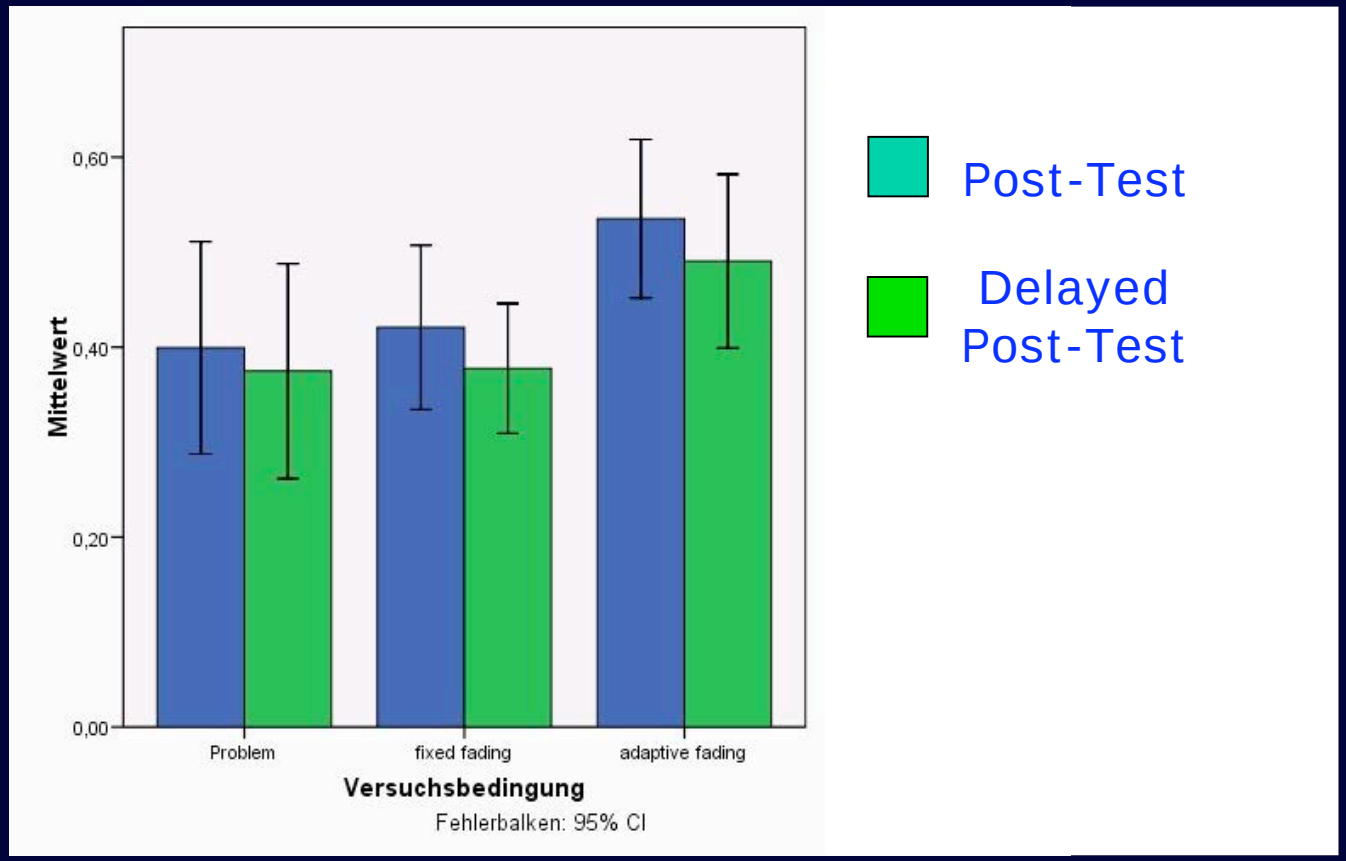
Hanzi->English



Adaptive Fading of Examples

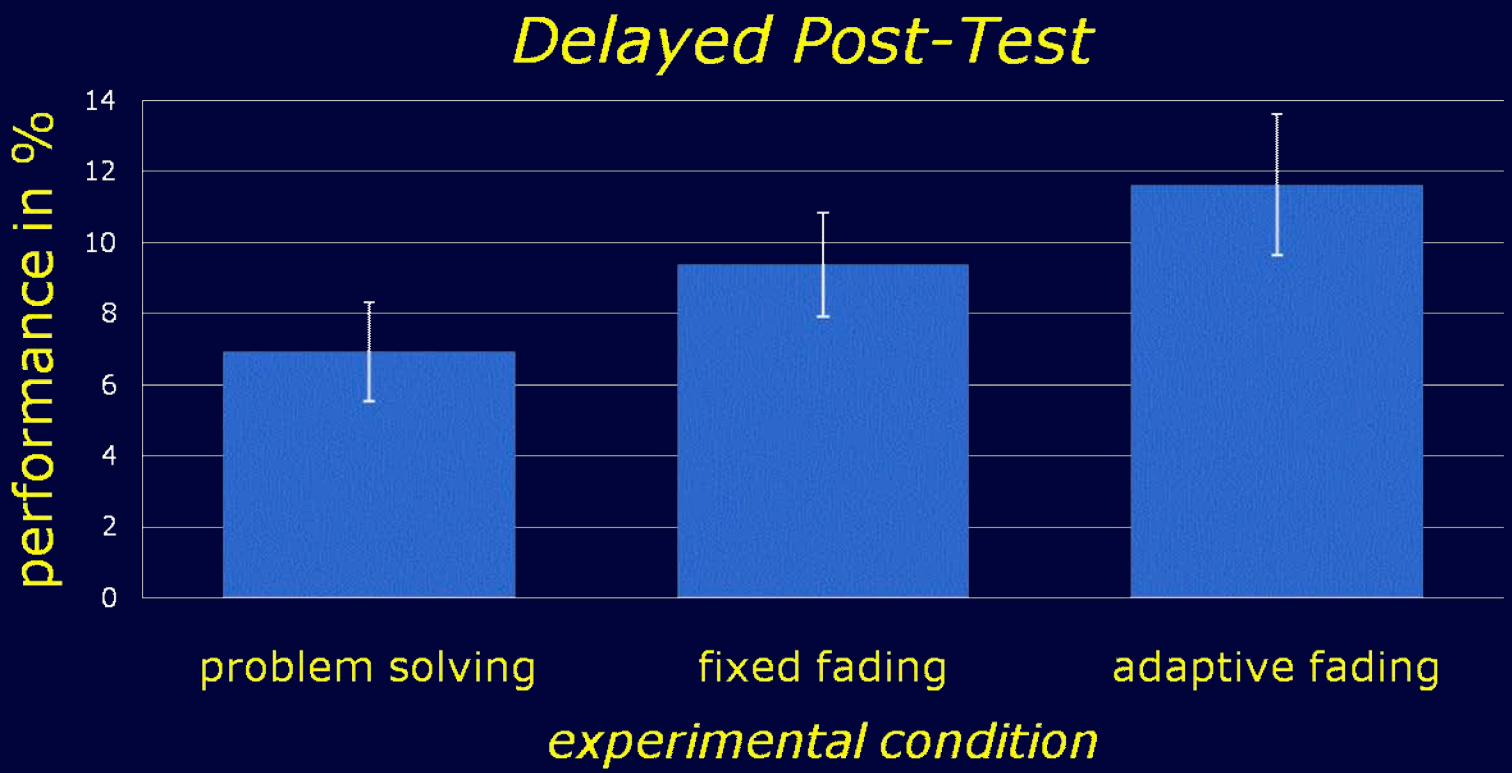
- Fading based on quality of self explanations of worked out value steps (assessed by Tutor)
- Students who self explain well receive fewer examples than students who self explain poorly

Results



- adaptive fading examples > fixed conditions

Results: In Vivo study



- Result is robust in classroom environment: adaptive fading examples > problem solving

END 2

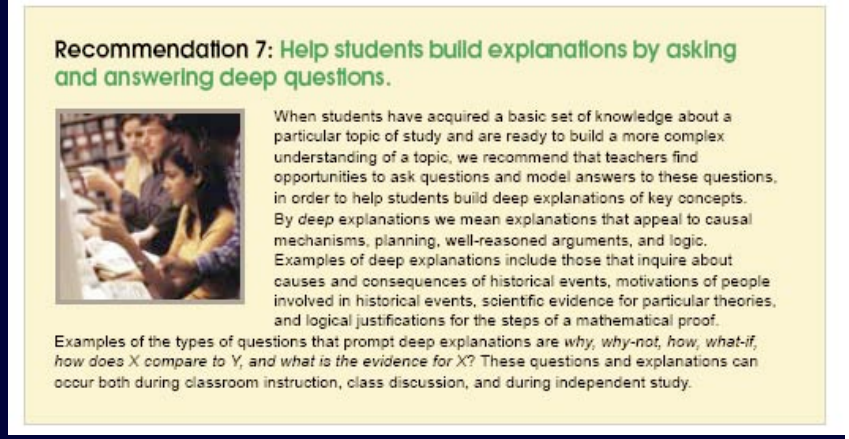
Examples <> Study trials
Problems <> Test trials

What's the difference?

- Testing effect studies focus on *facts*
  - Learning process is memory
- Worked example studies are about learning *general rules* & procedures
  - Robust learning requires *category induction* as well as memory

Progress on Principles of Learning to Guide Practice

- PSLC wiki: See learnlab.org
- IES: Cognition Practice Guide
- APS group & web site



Recommendation 7: Help students build explanations by asking and answering deep questions.

When students have acquired a basic set of knowledge about a particular topic of study and are ready to build a more complex understanding of a topic, we recommend that teachers find opportunities to ask questions and model answers to these questions, in order to help students build deep explanations of key concepts. By *deep* explanations we mean explanations that appeal to causal mechanisms, planning, well-reasoned arguments, and logic. Examples of deep explanations include those that inquire about causes and consequences of historical events, motivations of people involved in historical events, scientific evidence for particular theories, and logical justifications for the steps of a mathematical proof. Examples of the types of questions that prompt deep explanations are *why, why-not, how, what-if, how does X compare to Y, and what is the evidence for X*? These questions and explanations can occur both during classroom instruction, class discussion, and during independent study.