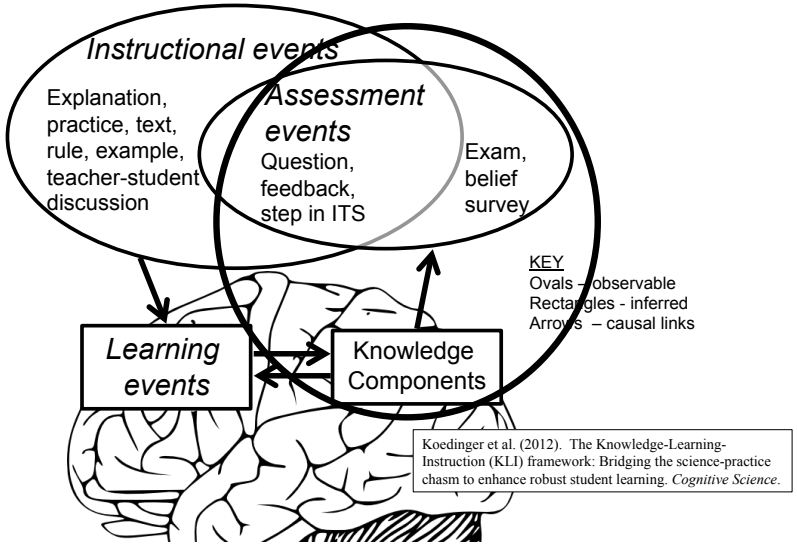
Cognitive Task Analysis: Think Alouds and Difficulty Factors Assessment

Ken Koedinger
HCI & Psychology
CMU Director of LearnLab

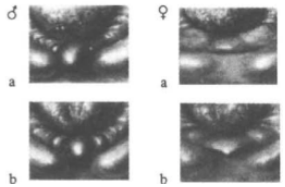
KLI: Infer KCs from repeated assessment events



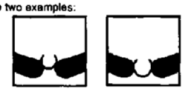
Koedinger et al. (2012). The Knowledge-Learning-Instruction (KLI) framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*.

Unpacking & repacking expertise: Chick sexing

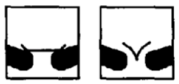
- Experts don't know, what they know
  - 98% accurate after years of on-the-job training
- Interviews led to design of "pictures in which *critical features* of various types were indicated"
- After just minutes of instruction, novices brought to 84% accuracy!



Male chicken genitals tend to look round and fullish like a ball or watermelon. Here are two examples:



Female chicken genitals can take on two different appearances. They can look pointed, like an upside down pine tree, or flatish. Here are two examples:

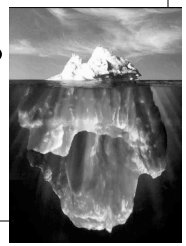


Biederman & Shiffrar (1987). Sexing Day-Old Chicks: A Case Study and Expert Systems Analysis of a Difficult Perceptual-Learning Task. JEP: Learning, Memory, & Cognition.

Do you know what you know?

Why data is important to improving student learning

- If we knew everything about students' learning challenges, we would not need data
- But, there is a lot we *do not know* about student learning
- In fact, there's a lot we don't know about our *own* learning
  - You've had lots of experience with the English language
  - You might say you know English
  - But, do you know what you know?



Overview

- Cognitive Task Analysis
  - What is it? Why do it?
- CTA methods
  - Difficulty Factors Assessment
  - Think Aloud
- Hands-on exercise

Cognitive Task Analysis

- Techniques to specify *cognitive structures & processes* associated with task performance
  - Structured interviews of experts
  - Think alouds of experts & novices performing tasks
  - Computer simulations of human reasoning

Cognitive Task Analysis Improves Instruction

Studies: Traditional instruction vs. CTA-based

- Med school catheter insertion (Velmahos et al., 2004)
  - Sig greater pre to post gain
  - Better with patients on all four measures used
  - Example: Sig fewer needle insertion attempts!
- Other examples
  - Radar system troubleshooting (Schaafstal et al., 2000)
  - Spreadsheet use (Merrill, 2002)
- Meta-analysis, 7 studies: 1.7 effect size! (Lee, 2004)

Isn't knowledge analysis done for long-standing academic domains?

- Hasn't all this been worked out?
- Surely by now we understand the content of, say, Algebra?

No! CTA has a lot to offer even in "well treaded" academic domains

Story Problem

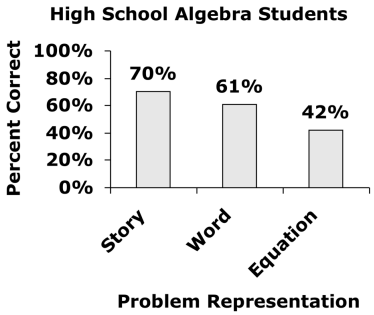
As a waiter, Ted gets $6 per hour. One night he made $66 in tips and earned a total of $81.90. How many hours did Ted work?

Word Problem

Starting with some number, if I multiply it by 6 and then add 66, I get 81.90. What number did I start with?

Equation

x * 6 + 66 = 81.90



Koedinger & Nathan (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *The Journal of the Learning Sciences*.

Overview

- Cognitive Task Analysis
  - What is it? Why do it?
- CTA methods
  - ***Difficulty Factors Assessment***
  - Think Aloud
- Hands-on exercise

Cognitive Task Analysis (CTA) purposes & uses

- For a topic you want to teach, CTA helps to decompose it into knowledge components
  - What components are learners' missing?
  - What order do they acquire these components?
  - Which components are particularly hard to acquire?
  - What "hidden skills" must be acquired?
- Good instruction targets what students don't know & builds on what they do know

CTA using Difficulty Factors Assessment (DFA)

1. Design & administer DFA
   a. Hypothesize task difficulty factors
   b. Design tasks to vary factors
   c. Administer as paper quiz or on-line
2. Analyze results
3. Create models
4. Redesign instruction

1a. Hypothesize task difficulty factors

- Rational analysis
  - Example: Story problems are solved by translating to equations
    (this analysis turned out to be inaccurate)
- Use course experience
  - Start with test items or homework tasks
  - Modify by reducing or increasing difficulty
- Read relevant literature
- Think aloud

1b. Design tasks to vary factors, create forms

- Difficulty factors involved
  - Presentation type
    - Story, Word, vs. Equation
  - Unknown position
    - Result-unknown vs. start-unknown
  - Number type
    - Whole vs. decimal numbers
- Multiple quiz forms

Difficulty Factor Space and Latin Square Sampling

TABLE 3
Difficulty Factor Space and Distribution of Problems on Forms

			Integer				*Decimal*			
Presentation	*Unknown Positition*	*Final Arithmetic*	*Donut* $4 \times 25 + 10 = 110$	*Lottery* $20 \times 3 + 40 = 100$	*Waiter* $4 \times 6 + 66 = 90$	*Basketball* $3 \times 5 + 34 = 49$	*Donut* $7 \times 0.37 + .22 = 2.81$	*Lottery* $26.50 \times 3 + 64 = 143.50$	*Waiter* $2.65 \times 6 + 66 = 81.90$	*Basketball* $10.84 \times 4 + 25 = 68.36$
Story	Result	×, +					1.1			
Story	Result	–, /							1.3	
Story	Start	×, +								1.4
Story	Start	–, /						1.2		
Word equation	Result	×, +								
Word equation	Result	–, /		1.6						
Word equation	Start	×, +								
Word equation	Start	–, /	1.5							
Equation	Result	×, +				1.8				
Equation	Result	–, /								
Equation	Start	×, +			1.7					
Equation	Start	–, /								

Note. The terms *integer* and *decimal* are the Number Type. *Donut, lottery, waiter,* and *basketball* are the Cover Story. The equations are Base Equations.

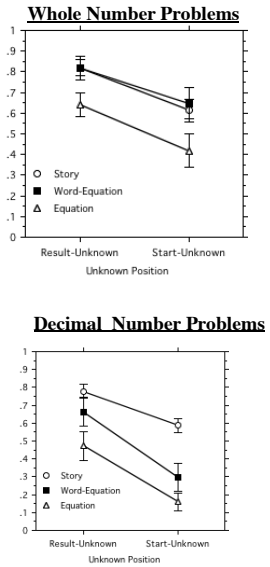
Story x Pres x Unkn-pos x Num-type x Major-Op
4 x 3 x 2 x 2 x 2 = 96 different problems
8 problems/form, 12 forms

1c. Administer as paper quiz or on-line

- This story problem DFA & others have been administered as paper quizzes
  - Koedinger, K. R., Alibali, M. W., & Nathan, M. M. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. Cognitive Science, 32(2), 366-397.
  - Baker, R. S. J. d., Corbett, A.T. & Koedinger, K. R. (2007). The difficulty factors approach to the design of lessons in intelligent tutor curricula. *International Journal of Artificial Intelligence in Education, 17(4)*, 341-369.
  - Rittle-Johnson, B. & Koedinger, K. R. (2005). Designing knowledge scaffolds to support mathematical problem solving. Cognition and Instruction. 23(3), 313-349
- Some implemented on-line
  - Koedinger, K.R. & McLaughlin, E.A. (2010). Seeing language learning inside the math: Cognitive analysis yields transfer. In S. Ohlsson & R. Catrambone (Eds.), Proceedings of the 32nd Annual Conference of the Cognitive Science Society. (pp. 471-476.) Austin, TX: Cognitive Science Society.

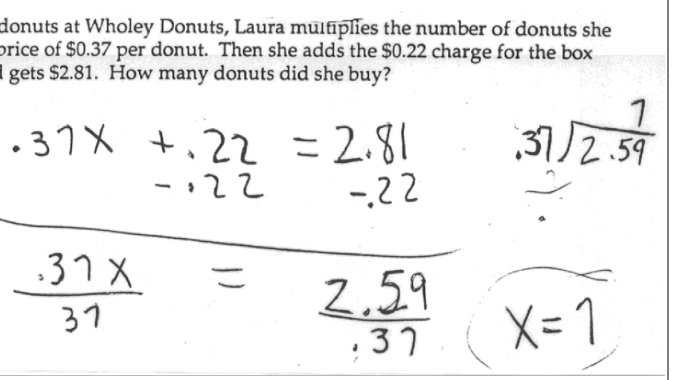
2. Analyze results

- Quantitative
  - Main effects & interactions
- Qualitative
  - Strategies & errors
  - …

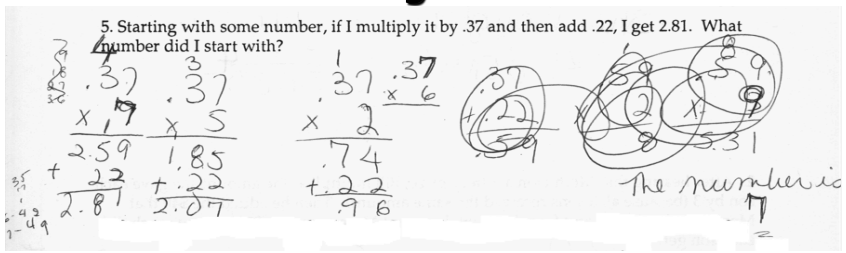


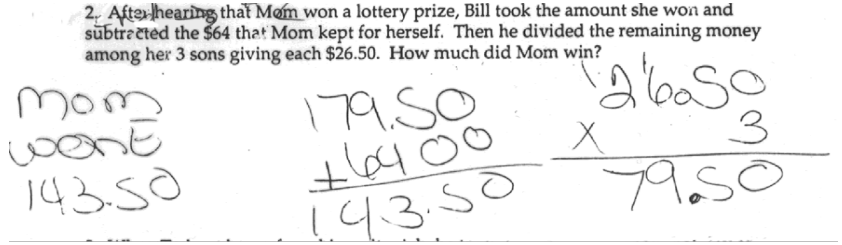
Formal, Translate & Solve Strategy

8. After buying donuts at Wholey Donuts, Laura multiplies the number of donuts she bought by their price of $0.37 per donut. Then she adds the $0.22 charge for the box they came in and gets $2.81. How many donuts did she buy?

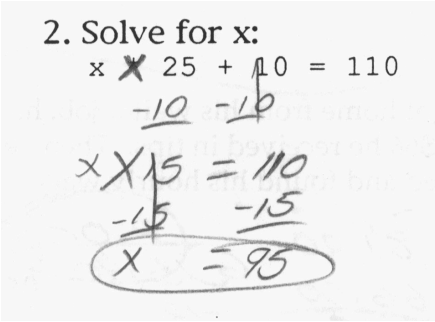


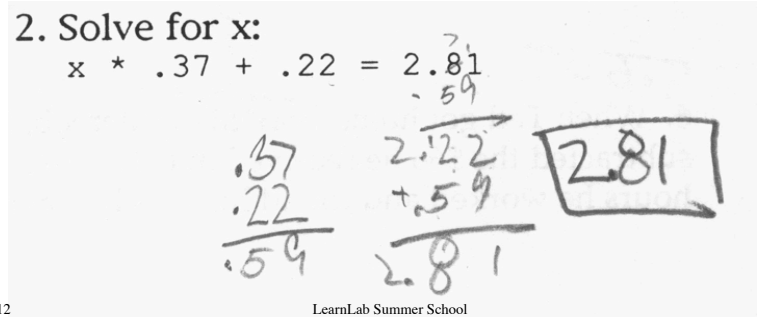
More Common: Informal Strategies





Algebra equations are like a foreign language -- takes extensive experience to acquire





3. Create models

- Cognitive model
  - Write production rules (in English is OK)
  - Create a task x KC matrix (a "Q matrix")
- Learning progression: sequence of tasks
  - Koedinger, K. R., Alibali, M. W., & Nathan, M. M. (2008). Trade-offs between grounded and abstract representations: Evidence from algebra problem solving. Cognitive Science, 32(2), 366-397.

Example Production Rule

English version

FOCUS-ON-LAST-RELATION

IF the goal is to solve a problem
  and there is a quantitative relationship with an operation
  involving an unknown input
  and a known input quantity
  that produces a known output quantity
THEN
  focus on that quantitative relationship

ACT-R 3.0 version

```
(p  Focus-on-Last-Relation
=Goal> isa Problem-Goal
      focus NIL
      context =Equation
=REL> isa *Relation
      equation =Equation
      arg2     =Arg2
      result   =Res3
=Arg2> isa *Quantity
       status Known
=Res3> isa *Quantity
       status Known
==>
=Goal>  focus =REL)
```

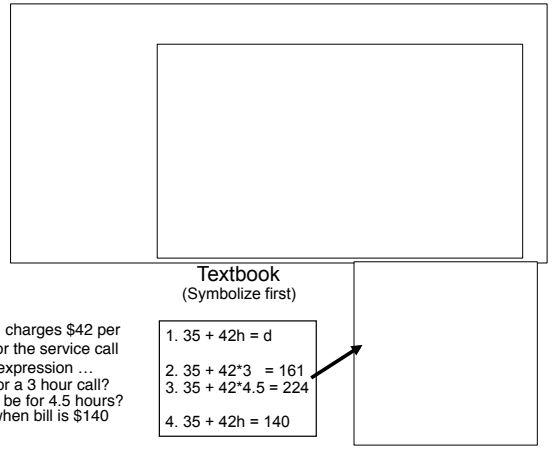
4. Redesign instruction

See next …

Using Cognitive Task Analysis to design better algebra instruction

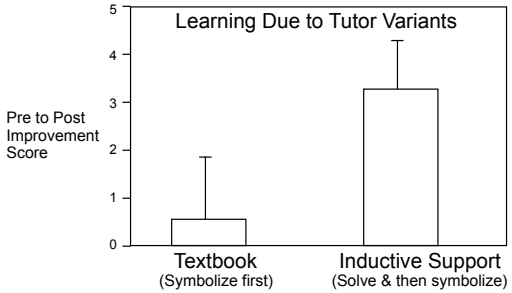
- With CTA, we discovered something about student prior knowledge
  - Can design instruction to build on it?
- Inductive support strategy
  - Help students generalize abstract math from their own intuitive concrete solutions
  - Similar to "progressive formalization" or "concreteness-fading" (Golstone & Son, 05)
- Does it work?
  - Test idea with an *in vivo experiment*

Textbook vs. Cognitively-Based Design



Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments*.

Parametric Study: Textbook vs. Cognitively-Based Design



…Plumbing Co. charges $42 per hour plus $35 for the service call
1. …. write an expression …
2. How much for a 3 hour call?
3. What will bill be for 4.5 hours?
4. Find hours when bill is $140

1. 35 + 42h = d
2. 35 + 42*3 = 161
3. 35 + 42*4.5 = 224
4. 35 + 42h = 140

2. 35 + 42*3 = 161
3. 35 + 42*4.5 = 224
1. 35 + 42h = d
4. 35 + 42h = 140

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of a cognitive tutor for algebra symbolization. *Interactive Learning Environments*.

Overview

- Cognitive Task Analysis
  - What is it? Why do it?
- CTA methods
  - Difficulty Factors Assessment
  - ***Think Aloud***
- Hands-on exercise

What *is* a Think-Aloud Study?

Basically, ask a users to "think aloud" as they work...

...on a task you want to study

...while you observe & audio or videotape

...either in context (school) or in lab

...possibly using paper/storyboard/interface you are interested in improving

The Roots of Think-Aloud Protocols

- Allen Newell and Herb Simon created the technique in 1970s
  - Applied in '72 book: "Human Problem Solving"
- Anders Ericsson & Herb Simon's book
  - "Protocol Analysis: Verbal Reports as Data" 1984, 1993
  - Explained & validated technique

The Cognitive Psychology Theory behind Think-Aloud Protocols

- People can easily verbalize the *linguistic* contents of Working Memory (WM)
- People cannot directly verbalize:
  - The processes performed on the contents of WM
    - Procedural knowledge, which drives what we *do*, is outside our conscious awareness, it is "tacit", "implicit" knowledge.
    - People articulate better external states & some internal goals, not good at articulating operations & reasons for choice
  - *Non-linguistic* contents of WM, like visual images
- People can attempt to verbalize procedural or non-linguistic knowledge, however, doing so:
  - May alter the thinking process (for better or worse)
  - May interfere with the task at hand, slowing performance

How to Collect Data in a Think-Aloud Study

(Gomoll, 1990, is a good guide)

1. Set up observation
   - write tasks
   - recruit students
2. Describe general purpose of observation
3. Tell student that it's OK to quit at any time
4. Explain how to "think aloud"
   - give a demonstration
   - *give an unrelated practice task, e.g., add digits*
5. Explain that you will not provide help
6. Describe tasks
7. Ask for questions before you start; then begin observation
   - *say "please keep talking" if the participant falls silent for 5 seconds or more*
   - *be sensitive to a severe desire to quit*
8. Conclude the observation

Example: Think Alouds in Statistics Tutor Development

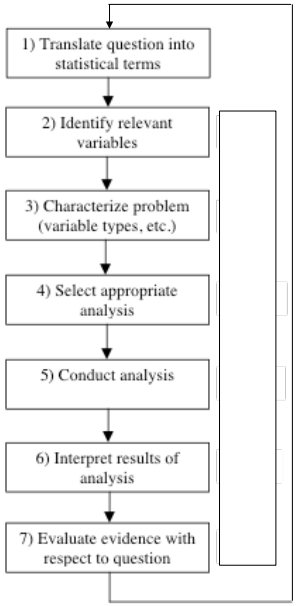
- Task: Exploratory Data Analysis
  - Given problem description and data set
  - Inspect data to generate summaries & conclusions
  - Evaluate the level of support for conclusions
- Example Problem

  In men's golf, professional players compete in either the regular tour (if they're under 51 years old) or in the senior tour (if they are 51 or older). Your friend wants to know if there is a difference in the amount of prize money won by the players in the 2 tours. This friend has recorded the prize money of the top 30 players in each tour. The variable *money* contains the money won by each of the players last year. The variable *tour* indicates which tour the player competed in, 1=regular, 2=senior. The variable *rank* indicates player rank, 1=top in the tour.

Thanks to Marsha Lovett!

Rational Task Analysis of Major Goals

- Inspired by ACT-R theory
- Break down task:
  - 7 major goals
  - Each goal has involves multiple steps or subgoals to perform
  - Key productions react to major goals & set subgoals



Sample Transcript

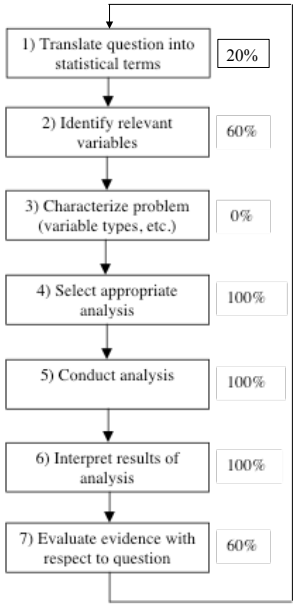
L#	Participants words & actions	Annotation
1	Oh, okay.	
2	So we need to, he wants to know whether there is a	Goal 1
3	difference in the amount of prize money, the amount of	
4	money won by players in the two tours.	
5	So, I think this is the prize money, uh, money contains the	Goal 2
6	prize money won by each of these players.	
7	Tour indicates which tour the player competes in.	
8	Well, you don't really need rank, in order to solve this, right?	
9	Cause like, well, I don't know.	
10	Um... I'm gonna do a boxplot... ...	Goal 4
11	[Subject uses statistics package to make a boxplot]	Goal 5
12	oh, cool (laugh)- I did it.	
13	All right, uh, so just looking at the average.	Goal 6
14	It looks like the people in the senior tour get less money.	
15	Um, and there's a lot less variation in the amount of money	
16	that, like all the prizes.	
17	A couple little outliers in each which means like, I don't	
18	know, like some people won, like a lot of money at a time…	

Key observations about this verbal report

- No evidence for goal 3 "characterize the problem"
  - Line 10: student simply jumps to selecting a data representation (goal 4) without thinking about why.
- No evidence for goal 7 "evaluate evidence"

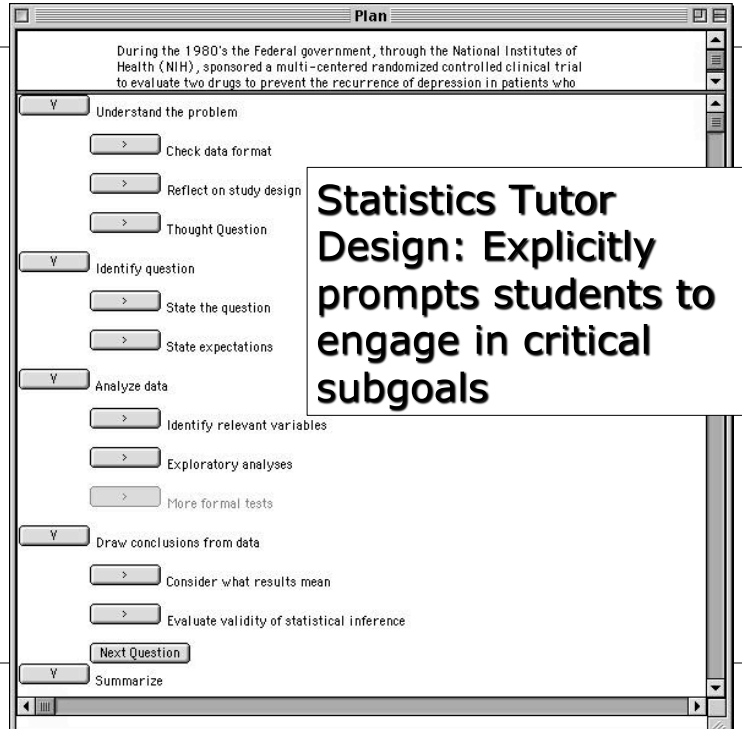
Comparing Think Aloud Results with Rational Task Analysis

- Percentages to the right of each step represent the percentage of students in the think-aloud study who showed explicit evidence of engaging in that step.
- Step 3 is totally absent!
  - A tutor can help students to do & remember to do step 3

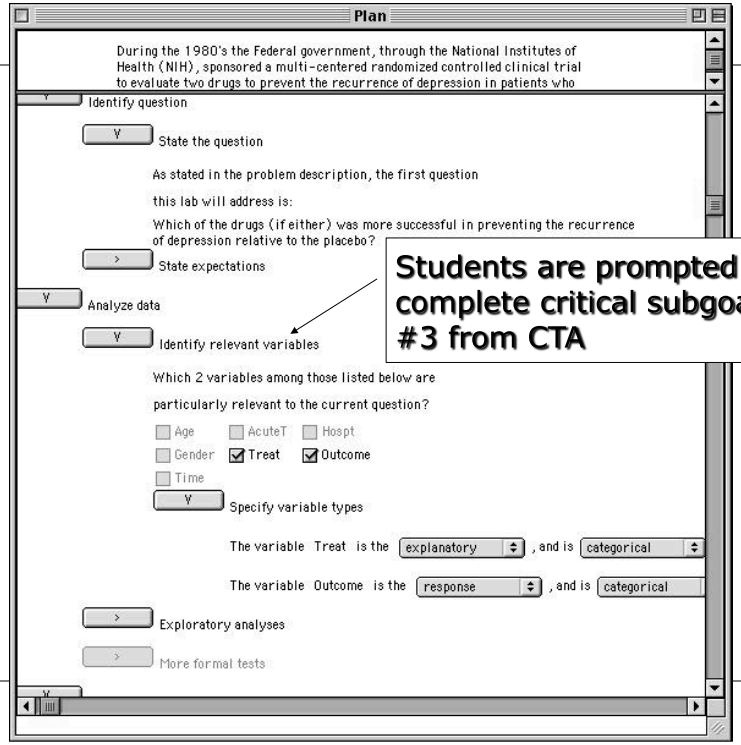


Inspiration for Production Rules (Knowledge Components)

- Missing production (to set goal 3):
  <u>Characterize problem</u>
  If goal is to do an exploratory data analysis
  & relevant variables have been identified
  then
  set a subgoal to identify variable types
- Buggy production (skipping from goal 2 to 4):
  <u>Select any data representation</u>
  If goal is to do an exploratory data analysis
  & relevant variables have been identified
  then
  set a subgoal to conduct an analysis by picking any data representation



Statistics Tutor Design: Explicitly prompts students to engage in critical subgoals



Students are prompted to complete critical subgoal #3 from CTA

Overview

- Cognitive Task Analysis
  - What is it? Why do it?
- CTA methods
  - Difficulty Factors Assessment
  - Think Aloud
- ***Hands-on exercise***

Cognitive Task Analysis Exercise

- Use Think Aloud to design a Difficulty Factors Assessment
- Find someone next to you to work with
  - I will give two problems
  - Take turns giving a think-aloud solving these next two problems

Try this ...

- One person think aloud while solving this problem. You *can* use paper. Other person is experimenter.

  Experimenter: Remember to say "keep talking" whenever participant is silent

- Ready ...
- What is 5 ÷ 3/4 = ?

Now this ...

- Switch roles:
  - Other person think aloud
  - What's written on paper is part of TA
  - Did the experimenter say "keep talking"?
- Ready …
- If 5 yards of ribbon are cut into pieces that are each 3/4 yard long to make bows, how many bows can be made?

Think about student thinking …

- Which will be easier?
- Why?
- Strategy & error analysis:
  - What strategies will students use?
  - Will there be differences in strategy selection between problem types?
  - What errors might account for observed differences?

How could you design a DFA to test your hypotheses?

- Can you put these two problems on the same quiz form?
  - Why not? What can you do instead?
- What other factors might be involved?
  - Size of the numbers--big nums discourage informal strategy
  - "Tempting" nums like 6 ÷ 3/5
  - Order: context first vs. context second

"Latin Square" Design

	5 ÷ 3/4 = ?	7 ÷ 2/3 = ?
No Context	Form 1	*Form 2*
Context	*Form 2*	Form 1

- Don't give problems with same answer on same form
- Can give problems with both values of a difficulty factor
- Example above
  - Students using either Form 1 or Form 2 will get both a No-Context & a Context problem
  - But, two forms "counter balance" the number types

Cognitive Task Analysis Summary

Cognitive Task Analysis ⇨ **Cognitive Model** ⇨ **Better instructional design**

- A cognitive model of student reasoning & learning *in a specific domain* guides instructional design
- Do Cognitive Task Analysis (CTA) to develop a cognitive model
  - Rational CTA: Articulate knowledge components in English (or in a computer simulation like a production rule system)
  - Empirical CTA methods: Think Aloud, Difficulty Factors Assessment, educational data mining techniques …
- Think aloud: Rich data on student thinking processes
  - Best way to develop good intuitions about student thinking!
- Difficulty Factors Analysis
  - Quickly & systematically focus in on what's hard for learners
  - Can also do through educational data mining

END

Extra slides follow …

Think Aloud activity you might try with another team

- Team A members do Think Alouds with Team B members
  - Alternate experimenter & participant roles
  - Experiment presents your task
  - Participant performs task & thinks aloud
- First round:
  - A1 is experimenter, B1 is participant
  - A2 is participant, B2 is experimenter
- Second round -- switch roles
  - A1 is participant, B1 is experimenter
  - A2 is experimenter, B2 is participant

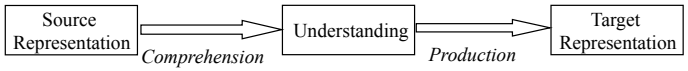
Strategies for identifying potentially interesting difficulty factors

- Ask yourself & teachers: What's most difficult for students to learn in this class?
- Add or reduce complexity in an existing test item
  - Add complexity: multiple operations, type & scale of numbers involved, distractors, abstract formalisms
  - Reduce complexity by drawing on prior knowledge
    - Place problem in familiar context
    - Use concrete instances instead of abstractions
    - Use a concrete pictorial representation

Extended Example of a Difficulty Factors Assessment design

- Heffernan, N. & Koedinger, K. R. (1997). The composition effect in symbolizing: The role of symbol production vs. text comprehension. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*. [Marr prize winner.]

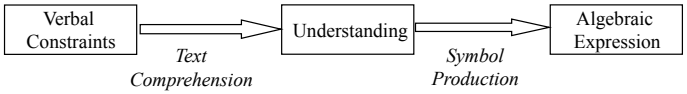
Symbolization Task



EXAMPLE PROBLEM

Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom and then use the remainder to buy presents for each of her 4 sisters. Write an expression for how much she can spend on each sister.

Example answer:
(72-m)/4



Rational Cognitive Task Analysis:
How Does One Symbolize?

- Comprehend
  - Figuring out the math operations involved (e.g., "… remainder …" -> "subtract")
- Produce symbols
  - "subtraction" -> "-"
  - Order of operations, getting paren's right
  - Being to able to write "embedded clauses",
    expr -> num op num
    expr -> expr op expr

Select Difficulty Factors to identify challenging cognitive processes

Potentially challenging cognitive process

1. Reading story problem
2. Avoiding shallow processing
3. Writing expressions with variables
4. Composing 2-op symbolic sentences

Associated difficulty factor manipulation

1. Comprehension hints vs. none
2. Distractor numbers vs. none
3. Variable vs. numbers
4. Decomposed (two 1-op) vs. composed (one 2-op)

Start with Core Problem.

Create new problems by adding or deleting difficulty factors

P0 Core Problem
Ann is in a rowboat in a lake. She is 800 yards from the dock. She then rows for "m" minutes back towards the dock. Ann rows at a speed of 40 yards per minute. Write an expression for Ann's distance from the dock.

P1 Decomposed Problem
A) Ann is in a rowboat in a lake. She is 800 yards from the dock. She then rows "y" yards back towards the dock. Write an expression for Ann's distance from the dock.
B) Ann is in a rowboat in a lake. She then rows for "m" minutes back towards the dock. Ann rows at a speed of 40 yards per minute. Write an expression for the distance Ann has rowed.

P2 Distractor Problem
Ann is in a rowboat in a lake **that is 2400 yards wide**. She is 800 yards from the dock. She then rows for "m" minutes back towards the dock. Ann rows at a speed of 40 yards per minute. Write an expression for Ann's distance from the dock.

P3 Comprehension Hint
[Core Problem]
Hint 1: Ann's distance from the dock is equal to the 800 yards she started out from the dock minus the distance she has rowed in "m" minutes.
Hint 2: The distance she has rowed in "m" minutes is equal to the 40 yards she rows per minute multiplied by the "m" minutes it takes her.

P4 No Variable Problem
Ann is in a rowboat in a lake. She is 800 yards from the dock. She then rows for **11** minutes back towards the dock. Ann rows at a speed of 40 yards per minute. Write an expression for Ann's distance from the dock.

Overall Results

Difficulty factor	Significant Effect?
• Comprehension hints	• No
• Distractor numbers	• Yes
• Variable vs. numbers	• No
• Decomposed (two 1-op) vs. composed (one 2-op)	• Yes

Focus on two of these factors: Comprehension & Decomposition

CORE PROBLEM

Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom and then use the remainder to buy presents for each of her 4 sisters. She will spend the same amount on each sister. How much she can spend on each sister?

COMPREHENSION HINT VERSION

[Core problem followed by these hints.]

Hint 1: The amount Sue spends on all sisters is equal to the $72 she earned minus the "m" dollars she gives to Mom.

Hint 2: The amount Sue spends on each sister is equal to the amount Sue spends on all sisters divided by 4 (the number of sisters she has).

DECOMPOSED VERSION

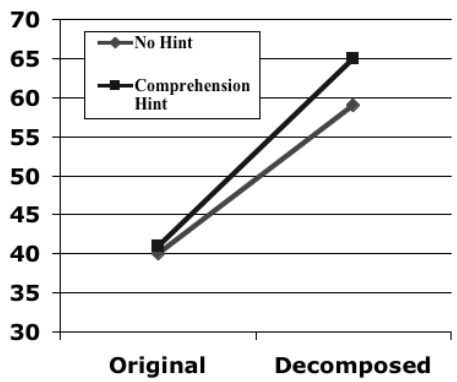
Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom. How much does she have left?

Sue has "x" dollars for presents for each of her 4 sisters. She will spend the same amount on each sister. How much she can spend on each sister?

Verbal Constraints → *Text Comprehension* → Understanding → *Symbol Production* → Algebraic Expression

Composition Effect =>

Symbol production not text comprehension



No comprehension hint effect: Students do not have much trouble comprehending problems, e.g., understanding "for each of" as "divides".

Composition effect: Students have trouble composing two operator algebraic sentences -- even when they understand both operations!

Error Analysis

CORE PROBLEM

Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom and then use the remainder to buy presents for each of her 4 sisters. She will spend the same amount on each sister. How much can she spend on each sister?

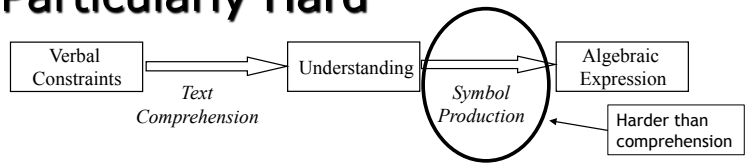
DECOMPOSED VERSION

Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom. How much does she have left?

Sue has "x" dollars for presents for each of her 4 sisters. She will spend the same amount on each sister. How much she can spend on each sister?

Correct Answer: (72 - m)/4	72 - m, x/4
Basic errors:	
Wrong operator: (72 - m) * 4	72+m
Argument order: 4 / (72 - m)	4/x
Composition errors:	
Invented notation: 72 - m = n / 4 =	4) $\overline{x}$
Missing parentheses: 72 - m/4	NA
Subexpression: 72 - m or m/4	NA

Producing Symbolic Sentences is Particularly Hard



- Decomposed success --> Students can *comprehend* of text
- Composed failure --> Cannot *produce* 2-op sentences:
  "(x - 72)/4"
  "800 - 40m"
- Variable success --> Producing is hard even without variable:
  "(96 - 72)/4"
  "800 - 40*3"

Example Production Rules

- *Works on decomposed problems:*
  If the goal is to symbolize quantity =Q,
  =Q is the result of applying operator =Op to =Num1 and =Num2
  =Op has symbol =Op-Sym
  Then write "=Num1 =Op-Sym =Num2"
- *Works on composed (w/o parens!)*
  If the goal is to symbolize quantity =Q,
  =Q is the result of applying operator =Op to expression =Expr1 and =Expr2
  =Op has symbol =Op-Sym
  Then write "=Expr1 =Op-Sym =Expr2"

This Analysis has Subtle Implications for Instruction

- Inductive support:
  Have students solve problems using small integers before writing symbols
- Create problems to isolate key difficulty
  - "Substitute x-74 for w in w / 4"
  - Apparently unrelated substitution exercises may improve story problem symbolization!