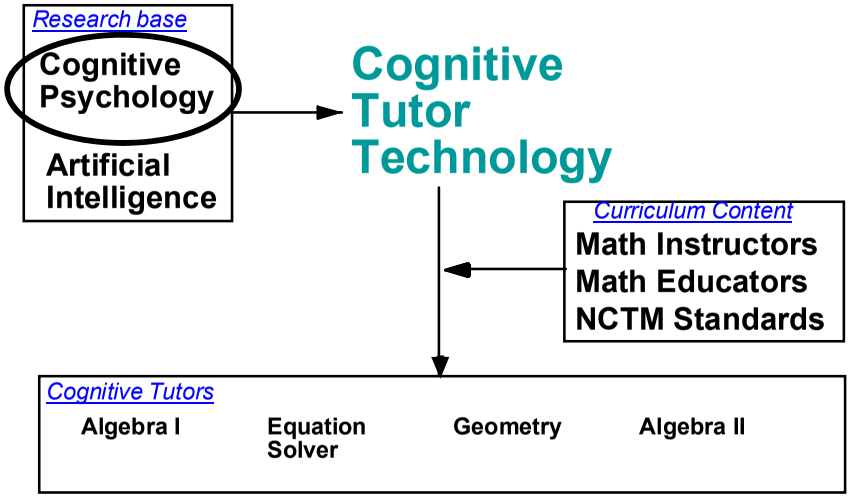
Cognitive Task Analysis: Think Alouds and Difficulty Factors Assessment

Ken Koedinger
HCI & Psychology
CMU Director of Pittsburgh Science of Learning Center

Overview

- Motivate Cognitive Task Analysis
- CTA Method 1: Think Alouds
- CTA Method 2: Difficulty Factors Assessment

Tutor Research & Development Process



Which problem is hardest for beginning algebra students?

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

Algebra Student Results: Story Problems are Easier!

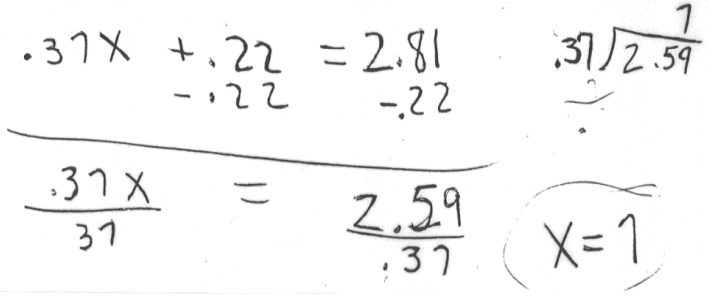
Koedinger & Nathan (2004). The real story behind story problems: Effects of representations on quantitative reasoning. In *International Journal of the Learning Sciences*.

Practical & Theoretical Implications of Surprising Results

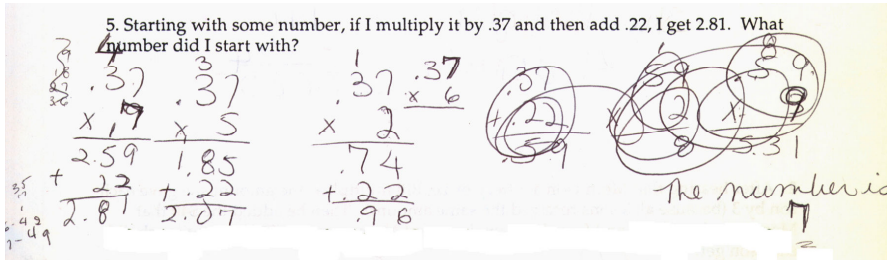
- Guided Cognitive Tutor Algebra design
  - Success due in part to smoothly bridging from students' existing common sense
- Inspired basic cognitive modeling work to explain these results
  - Coded student solutions for alternative strategies and for errors
  - What knowledge components could account for these?

Formal, Translate & Solve Strategy

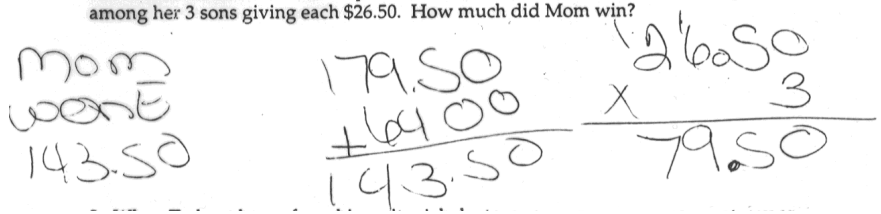
8. After buying donuts at Wholey Donuts, Laura multiplies the number of donuts she bought by their price of $0.37 per donut. Then she adds the $0.22 charge for the box they came in and gets $2.81. How many donuts did she buy?



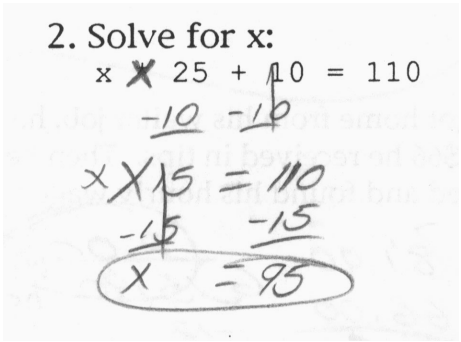
More Common: Informal Strategies

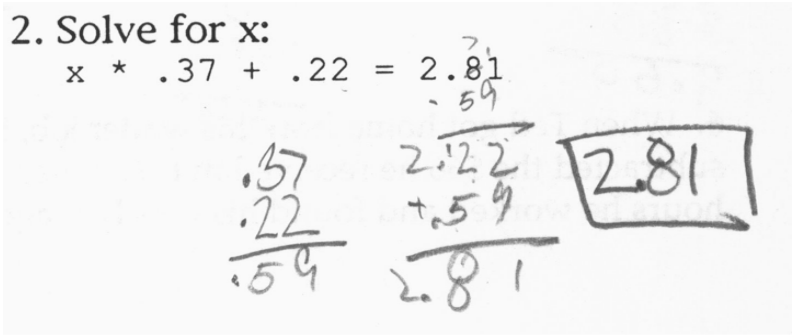


2. After hearing that Mom won a lottery prize, Bill took the amount she won and subtracted the $64 that Mom kept for herself. Then he divided the remaining money among her 3 sons giving each $26.50. How much did Mom win?



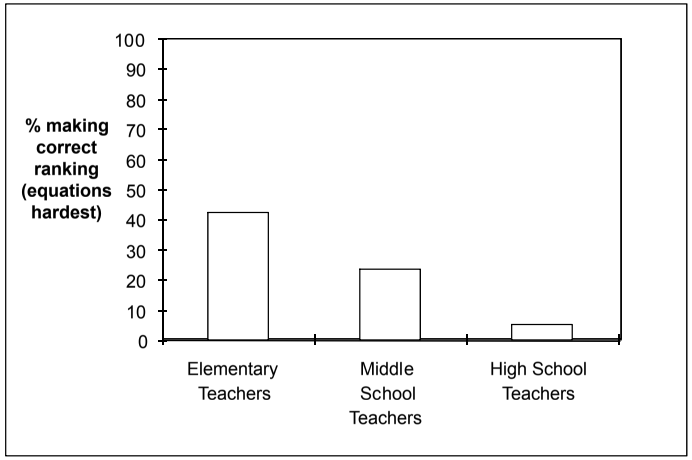
Algebra equations are like a foreign language -- takes extensive experience to acquire





Expert Blind Spot

Algebra teachers worst at recognizing algebra student difficulties



Nathan, M. J. & Koedinger, K.R. (2000). Teachers' and researchers' beliefs of early algebra development. *Journal of Mathematics Education Research*, *31* (2), 168-190.

Mantras for Technology Design

- To avoid expert blind spot, remember:

  "The Student Is Not Like Me"

- Version of the general HCI Mantra:

  "The User is Not Like Me"

- Use Cognitive & HCI methods to find out what students & users are really like
- That is, do *Cognitive Task Analysis*

Tutor Research & Development Process

1. Client & problem identification
2. Identify the target task & "interface"
3. *Perform Cognitive Task Analysis (CTA)*
4. Create Cognitive Model & Tutor
   a. Enhance interface based on CTA
   b. Create Cognitive Model based on CTA
   c. Build a curriculum based on CTA
5. Pilot & Parametric Studies
6. Classroom Use & Dissemination

Overview

- Motivate Cognitive Task Analysis
- *CTA Method 1: Think Alouds*
- CTA Method 2: Difficulty Factors Assessment

Kinds of Cognitive Task Analysis

- 2 Kinds of Approaches
  - Empirical: Based on observation, data
  - Analytical: Based on theory, modeling
- 2 Kinds of Goals
  - Descriptive: How students actually solve problems => what they already know
  - Prescriptive: How students should solve problems => what they need to learn
- 4 Combinations ...

Kinds of Cognitive Task Analysis

	Empirical	Theoretical
Descriptive	Think-aloud of novice. Difficulty Factors Assessment.	Cognitive modeling of errors, informal strategies.
Prescriptive	Think-aloud of expert. DFA	Cognitive modeling of success, normative strategies.

Steps In Task Analysis

- What are instructional objectives?
  - Standards, existing tests, signature tasks
- Has someone done the work for you? Don't reinvent the wheel. Do a literature review!
  - "8 weeks of analysis saves an hour in the library"
- Specify space of tasks
- Do either or both:
  - Theoretical task analysis: Use theory to create a process model that is sufficient to deal with space of tasks
  - Empirical task analysis: Do Think-Aloud, Difficulty Factors Assessment, ...

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 Usability Studies

- "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

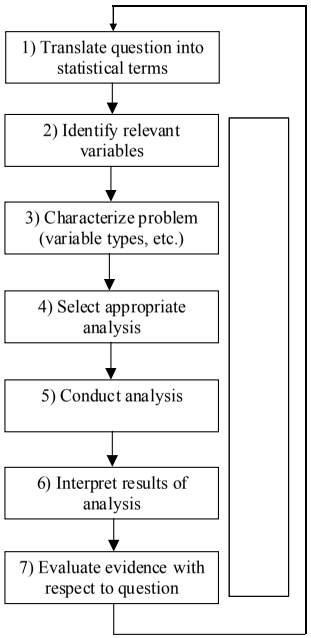
- Task: Exploratory Data Analysis
  - Given problem description and data set
  - Inspect data to generate summaries & conclusions
  - Evaluate the level of support for conclusions
- Example Problem

  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!

Task Analysis of Major Goals in Statistical Analysis

- This is an "analytic prescriptive" form of CTA
- ACT-R emphasizes "goal-factored" knowledge elements
- 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

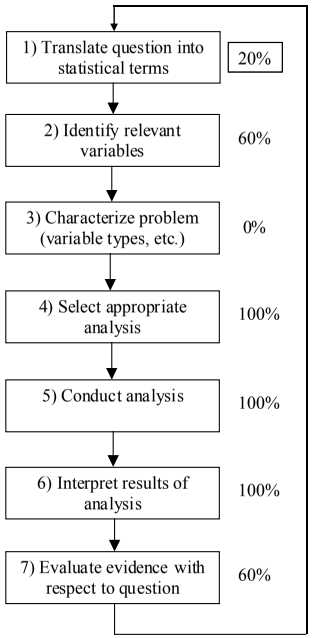
L#	Participants words & actions	Annotation
1	Oh, okay.	
2	So we need to, he wants to know whether there is a	Goal 1
3	difference in the amount of prize money, the amount of	
4	money won by players in the two tours.	
5	So, I think this is the prize money, uh, money contains the	Goal 2
6	prize money won 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…	

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
- Minor interpretation error
  - Line 13: student mentions the "average" when in fact boxplots display the median not the mean
- *Note: These observations should be indicated in the annotation column of the transcript* (I left them off given limited space)

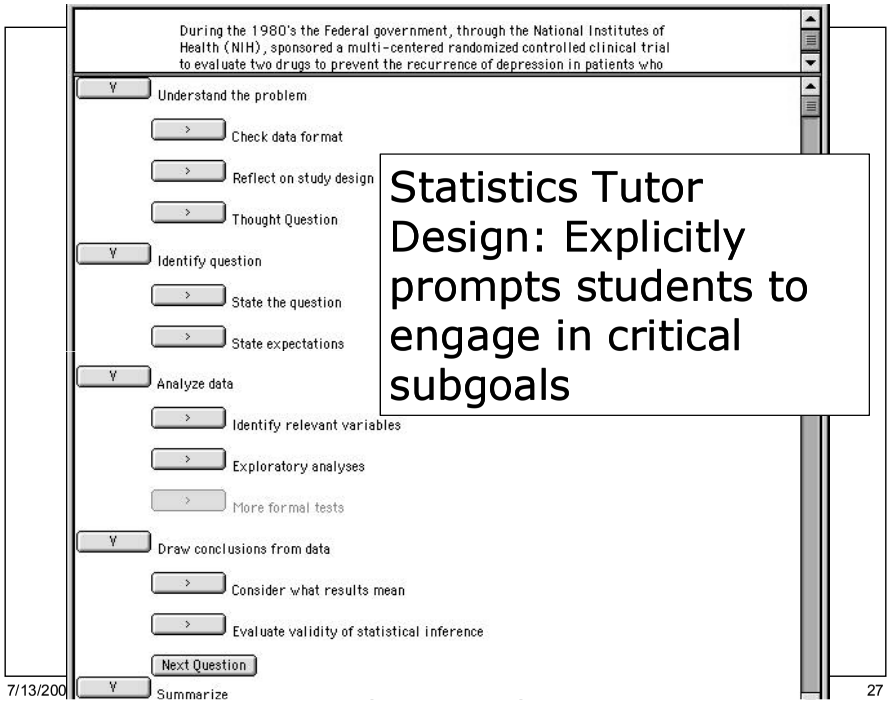
Comparing Think Aloud Results with Task Analysis

- 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

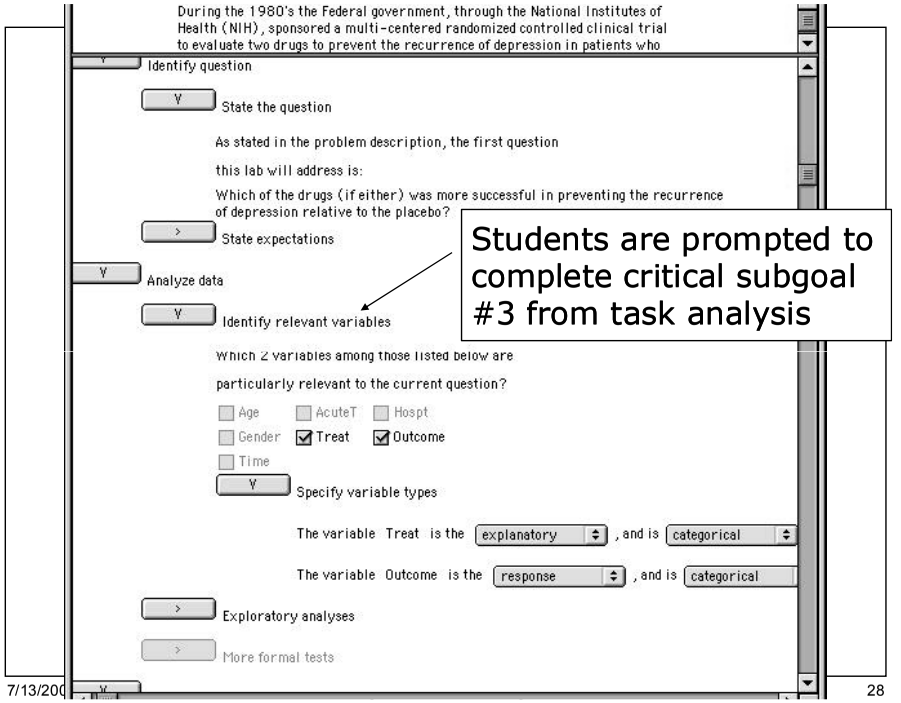


Inspiration for Production Rules

- Missing production (to set goal 3):
  Characterize problem
  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):
  Select any data representation
  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 task analysis

Think Aloud Summary

- 4 Kinds of Cognitive Task Analysis
  - Descrip vs. Prescrip; Empirical vs. Analytic
- Empirical CTA Methods
  - Think aloud & difficulty factors assessment
- Think aloud
  - Get subjects to talk while solving, do not have them explaining
  - Prescrip: What do experts know -- identify hidden thinking skills
  - Descrip: What is difficult for novices

Pros & Cons of Think Aloud

- Pros or advantages
  - Rich qualitative data
  - Get a great sense of student thinking processes
    - Students verbalizations may indicate goals, plans, strategies, or misconceptions
- Cons or disadvantages
  - Labor intensive: collect data individually, transcribing, analyzing
  - Subjective judgments to code verbal protocols
  - Usually does not provide data on learning changes over time

Overview

- Motivate Cognitive Task Analysis
- CTA Method 1: Think Alouds
- *CTA Method 2: Difficulty Factors Assessment*

Need for a Knowledge Decomposition Methodology

- Good instruction targets the edge of students' knowledge, what is "just-learnable"
- Need a method for decomposing a topic 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?
- Knowledge decomposition guides design of:
  - problem solving activities, tutor interface, cognitive model, hints and bug messages, problem sequence

Knowledge Decomposition through Difficulty Factors Assessment (DFA)

- Goal: Identify what is "just learnable" for students at different levels of competence
- The DFA methodology:
  1. Identify possible problem difficulty factors
     - Use think aloud or analytic task analysis
  2. Create test items & forms; Administer
  3. Analyze results:
     a. Main effects and interactions
     b. Strategies and errors
  4. Create a cognitive model
  5. Create a "developmental model", that is, the order in which productions are acquired

Example above was a DFA

- Difficulty factor illustrated above:
  - Presentation type: Story, Word, vs. Equation
- Other factors in that study:
  - Unknown position: Result-unknown vs. start-unknown
  - Number type: Whole vs. decimal numbers

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

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
- Employ other Cognitive Task Analysis techniques
  - Prescriptive analytic: Try to write production rules (in English) to solve task
  - Descriptive empirical: Think aloud study with novices

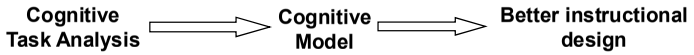
Advantages of Think Aloud (TA) (relative disadvantages of DFA)

- Get more rich qualitative data from TA
  - Written responses on DFAs can be sparse, sometimes we see only the answer
  - Students verbalizations during TA may better indicate goals, plans, strategies, or misconceptions
- Can see order of steps in TA
  - Written responses in DFA do not indicate order (see guess-and-test example)

Advantages of Difficulty Factors Assessment

- In contrast to other CTA approaches like
  - Think alouds, interviews, cognitive modeling
  - Which are often labor intensive, subjective, and reveal little about learning & development
- Difficulty Factors Assessment is typically
  - less labor intensive
  - more objective
  - indicates levels of learning & development
- Educational Data Mining track:
  - Learning Factors Analysis: a related discovery technique
  - Can be applied to existing data from tests or intelligent tutors

Cognitive Task Analysis Summary



- 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

END

EXTRA SLIDES

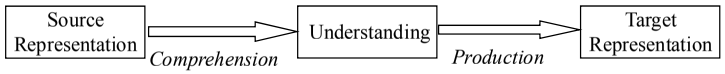
Think Aloud Activity you might try with another team

- 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

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.]

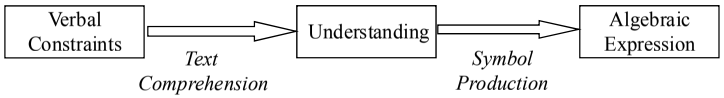
Symbolization Task



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., "… remaindor …" -> "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.

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.

Create new problems by adding or deleting difficulty factors

P3 Comprehension Hint
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.
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.]

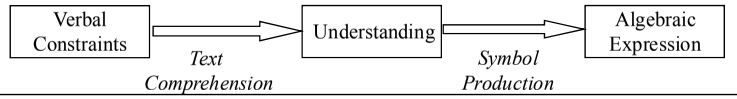
Hint 1: The amount Sue spends on all sisters is equal to the $72 she earned minus the "m" dollars she gives to Mom.

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

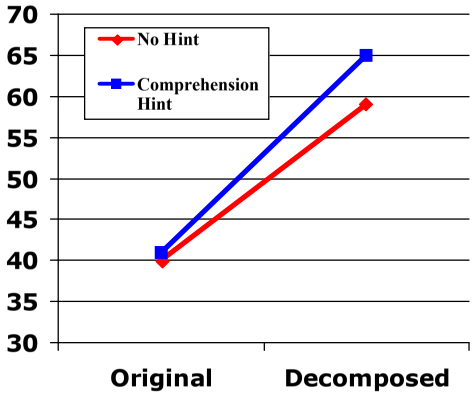
Sue made $72 washing cars. She decided to spend "m" dollars on a present for her mom. How much does she have left?

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?



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?

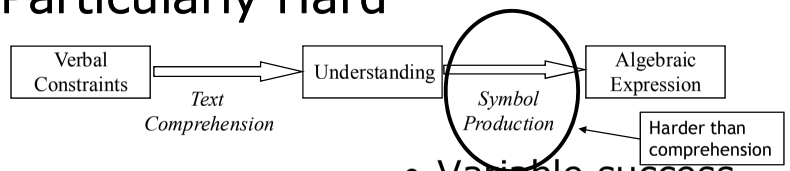
DECOMPOSED VERSION

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 "w = x - 74" into "y = w / 4". That is, express y in terms of x only
  - Apparently unrelated substitution exercises may improve story problem symbolization!