DOCUMENT RESUME

ED 362 540 TM 020 563

AUTHOR Glidden, Peter L.; Fry, Erin K.
TITLE Developing a Framework for a Fine-Grained Computational Theory of Algebra Learning.
SPONS AGENCY Illinois Univ., Urbana. Bureau of Educational Research.
PUB DATE Apr 93
NOTE 10p.; Paper presented at the Annual Meeting of the American Educational Research Association (Atlanta, GA, April 12-16, 1993).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Algebra; Audiotape Recordings; *Cognitive Processes; Cognitive Psychology; *Computation; Educational Theories; Grade 8; Junior High Schools; *Junior High School Students; *Learning Strategies; *Mathematics Instruction; Middle Schools; Problem Solving; Protocol Analysis; Semantics; Syntax; Urban Schools; Visual Perception
IDENTIFIERS Middle School Students; *Theory Development

ABSTRACT

Theoretical tools from cognitive science were used to put together a framework for a fine-grained theory of how students learn elementary algebra in the classroom. Developing the framework included assessing the shortcomings of current models, evaluating whether unified theories of cognition can be adapted for learning mathematics in the classroom, and discussing the extension of these theories. Protocols were collected from an eighth-grade class using the University of Chicago School Mathematics Project algebra text in an urban middle school. Audiotapes were made of student classroom conversations in working groups. Audiotapes from three students are quoted. Analyzing the data provides a framework for a fine-grained theory. Data suggest that in the initial stages of learning algebra, visual clues play a more important role than does syntactic or semantic understanding. Directly related is the importance of examples. Data also suggest that the more novel the problem, the more students are likely to rely on visual clues. Students in the early stages of learning tend to do mathematics problems based on how the symbols are arranged on the page, rather than syntactically deconstructing the expression or the underlying semantics. It is suggested that although a production system architecture is applicable to doing and learning mathematics in the classroom, students do not compose productions as quickly and easily as suggested by some theories. (Contains 25 references.) (SLD)

DEVELOPING A FRAMEWORK FOR A FINE-GRAINED COMPUTATIONAL THEORY OF ALGEBRA LEARNING

Peter L. Glidden & Erin K. Fry
University of Illinois at Urbana-Champaign
Paper presented at 1993 AERA, Session 47.04

Developing A Framework for a Fine-Grained Computational Theory of Algebra Learning

Peter L. Glidden and Erin K. Fry
University of Illinois at Urbana-Champaign
Paper Presented at 1993 AERA, Session 47.04

Recently there have been several calls for developing better theories of mathematics learning (e.g., Davis, 1984, 1986; Fischbein, 1990; Larkin, 1989). Giving impetus to these calls are recent advances in Cognitive Science. Ironically, although Cognitive Science has proposed several process models of doing mathematics and has progressed to the point where unified theories of cognition have been proposed (e.g., Anderson, 1983; Newell, 1990), theoretical progress in mathematics education has been slow. The objective of this study is to apply theoretical tools from Cognitive Science to develop a framework for a fine-grained theory of how students learn elementary algebra in the classroom. Developing this framework includes: (a) addressing the shortcomings of current process models, (b) evaluating whether unified theories of cognition can be adapted for learning mathematics in the classroom, and (c) discussing how these theories might be extended.

Theoretical Context

Perhaps the most widely known process model in mathematics is the **Buggy** work began by Brown and Burton and continued by VanLehn (Brown & Burton, 1978; Brown & VanLehn, 1980; Burton, 1982; VanLehn, 1982, 1986, 1987, 1990). This research posited that student errors in multidigit subtraction were not random, but instead resulted from the regular application of incorrect rules. This line of research culminated in VanLehn's **Sierra**--a coarse-grained, generative theory of inducing procedures one step at a time from examples. Extending the idea of regular errors to the domain of algebra, Matz (1980, 1982) and Sleeman (1982, 1984, 1986) developed models of algebra learning. Sleeman developed a computer program that modeled student algebra errors in much the same way that **Buggy** did. Matz's theory posited that students' errors result from incorrectly extrapolating known rules to new problems.

VanLehn's **Sierra** raises some questions that direct further theoretical development. The first is separating procedural and conceptual knowledge. VanLehn examines performance on multidigit subtraction because it is ". . . a virtually meaningless procedure. Most elementary school students have only a dim conception of the underlying semantics of subtraction." (VanLehn, 1982, p. 15.). This may be true for many students, but as Thompson found (1989), when students are taught subtraction by regrouping manipulatives, they learn more of the underlying semantics and consequently exhibit fewer of the types of subtraction errors studied by VanLehn. The second question raised by VanLehn's work is his dismissal of any meaningful role for the teacher. Although teachers differ widely in their effectiveness, a theory of mathematics learning in the classroom somehow *must* include them.

Extending these process models can facilitated by building on recently developed unified theories of cognition (e.g., ACT* by Anderson, 1983; Soar by Newell and his colleagues [Newell, 1990]). Although these theories have important differences (e.g., storage schemes for conceptual knowledge, resolution procedures for competing productions, and procedures for learning), they share several common features (most importantly, a production-system architecture) upon which the process models above can be extended.

Finally, Davis's (1984) theory of mathematical cognition provides a precedent and context for using theoretical constructs from Cognitive Science in mathematics education. Although Davis's theory is schemata-based, in contrast to the production systems used in **Sierra**, ACT*, and Soar, it remains important as the most current and complete cognitive theory of doing mathematics in mathematics education. Davis based his theory on then current constructs from Cognitive Science. Thus, the present study extends Davis's work by adapting now current constructs from Cognitive Science to further the development of cognitive theories of mathematical cognition.

Methodology and Data

The methodology used in this study was pioneered by Newell and Simon (1972): protocols (including the text and student written work) are analyzed and the underlying cognitive processes were abstracted from them. These protocols were collected from an eighth-grade class using the University of Chicago School Mathematics Project (UCSMP) *Algebra* text (1990). This class was mixed race and SES, and it was taught at a middle school in a small midwestern city. A typical class period consisted of the teacher explaining the day's lesson to the class for about 15 minutes, followed by the students reading the text for about 10 minutes, and then the students working on exercises in small groups for the remaining 15 minutes. The second author audio-taped each class that was taught (i.e., she skipped test days).

This methodology was selected for two related reasons: (a) to reduce the "expectation effect" described by Schoenfeld (1985) and (b) the desire to investigate how students did and learned mathematics in as natural a setting as possible. In a clinical setting, Schoenfeld (1985) found that students tried to act "mathematical" to please the investigator. To reduce this "expectation effect," the second author sat in on the class for several sessions before she began audio-taping it and, for the most part, she worked with the same small group of two students for the duration of the study. It was believed that by participating in the daily routine of the class, she would gain acceptance from the subjects and not be perceived as an authority figure. Only occasionally did she ask probing questions and she did not ask students to "think out loud" because it was assumed that students would talk about the mathematics as they worked collaboratively.

Despite the similarity in methodology, there are two differences between the methodology in this study and that pioneered by Newell and Simon. The first is the difference in task domain: Newell and Simon were investigating performance in a knowledge-lean domain (solving cryptarithms) and the current domain is performance in a knowledge-rich domain (learning elementary algebra). The second is that, in contrast to Newell and Simon, these subjects were not asked to "think out loud." Consequently, it is likely that the subjects quietly brought new resources to bear on a problem. These differences suggest that the data collected represent only part of the underlying cognitive processes.

Results

Analyzing the data provides a framework for developing a fine-grained computational theory of algebra learning. In order to tell a coherent story, this framework is presented under the two themes of doing and learning mathematics. For each of these two themes, the framework is presented, followed by data to support it, concluding with the theoretical contribution.

Doing Mathematics

The data suggest that in the initial stages of learning algebra, visual clues play a more important role than syntactic or semantic understanding. Directly related to this is the importance of examples. If preceding examples present a consistent syntactic structure, students often apply that solution to a syntactically different problem. For example, the class was to simplify the following four expressions (UCSMP, 1990, p. 67.) (The numbers to the left are the numbers of the problems in the text.):

23. $x + b + 0 + -x + -b + -c$

24. $-(-(-4))$

25. $(17 + -35) + (36 + -18)$

26. $-7t^2 + 22t^2$

After simplifying the first three expressions correctly, both subjects wanted to solve the last by "canceling out the t^2's" giving -7 + 22. Their solution method, which is one of the misgeneralizations catalogued by Matz (1982), is best interpreted as a response to visual clues because if they understood the underlying syntactic or semantic structure, it is unlikely that they would not have used his solution method.

While learning about special numbers in equations (i.e., solution sets to equations like $0x = 4$ and $13t = 0$), they were asked to describe the solutions to the following two problems (UCSMP, 1990, p. 180.):

2. $7y = 0$

3. $0 \cdot w = 14$.

One subject, Mike, believed the solution to the first is all real numbers. Presumably, he sees the letter in the second position and concludes that the solution to that problem, like the answer to the first example, is all reals. In answering the next question, Mike offers all real numbers as possible solutions. Here again, the data suggest that Mike is using visual clues to guide solving the problem.

Later, while solving review problems, they correctly found the area of a rectangle 4.5ft by 3ft and a rectangle $\frac{1}{2}$m by $\frac{2}{3}$m. When asked to find the area of a rectangle $3x$ by $2x$, they suggested $6x$. When pressed by the investigator, they insisted that it was $6x$, not $6x^2$. They misinterpreted the x in the last problem as a symbol for units rather than as a variable. Again, the students are responding to visual clues rather than syntactical or semantic understanding.

A recurring difficulty was finding the reciprocals of numbers not written as fractions. This difficulty is ascribed to the lack of visual clues. For example, at different times our subjects had difficulty finding the reciprocals of -2, 0.7, and 4.2, even though they knew that reciprocal meant "opposite" (and waved their hands to indicate "flipping over"). When probed, the students could find the reciprocals $\frac{2}{3}$ and $\frac{1}{3}$, and extend that by suggesting $\frac{1}{2}$ (not $\frac{-1}{2}$!) as the reciprocal of -2. Similarly, they were led to find the reciprocals of 0.7 and 4.2. The data suggest that students initially find reciprocals by flipping over displayed fractions. If the fraction is implied, rather than explicit, they fail because there is no fraction to flip over, they can not respond to an absent visual clue.

The data also suggest that the more novel the problem, the more students are likely to rely on visual clues. While learning to solve equations of the form $ax + b = c$, all the examples presented had the variable term first. From these examples, Mike extrapolated the rule that to solve such a problem, one subtracts the *second* term (instead of the constant). Consequently, Mike tried to solve $5 + 3x = 9$, by subtracting $3x$ from both sides!

This sequence of vignettes suggests that early in gaining procedural competence, learners rely more the visual appearance of an algebraic expression than on the underlying syntax or semantics. They do mathematics based on how symbols are arranged on the page, rather than syntactically deconstructing the expression or understanding the underlying semantics. One could say, in fact, that they are trying to solve the problem without even *trying* to understand it.

This finding raises fundamental questions about students' initial understandings of new mathematics and, if a production rule architecture is used, how productions are matched to data (the pattern matching problem). Furthermore, it suggests that VanLehn's work offers more to mathematics education than Thompson (1989) posited. As discussed above, VanLehn (1990) studied arithmetic competence as symbol manipulation, while Thompson (1989) used manipulatives to promote semantic understanding. From Thompson's work, it is impossible to determine how students *initially* understood written subtraction. These vignettes suggest that they may initially have used visual clues, *then* made the connections to syntax and semantic understanding. In terms of developing a fine-grained theoretical framework, these data also suggest that VanLehn was correct to use the visual positioning of the digits in his theory of learning multidigit subtraction, but Thompson's work suggests that later these can be integrated with syntactic and semantic understanding. How to do this in a production system is an open question.

Learning Mathematics

With respect to learning mathematics, the data suggest that although a production system architecture is applicable to doing and learning mathematics in the classroom, students do not compose productions as quickly and easily as suggested by some theories. There are four aspects to this: (a) students are slow to compose productions, (b) their composed productions are tenuous, (c) they do not generalize very well, and (d) they use goals in plausible, but unanticipated ways.

Although these are recurring themes throughout most of the data, the first two aspects are best illustrated by examining the protocol taken from one class. In a series of related episodes, these data suggest how students compose productions in learning algebra. They are solving the following problem (UCSMP, 1990, p. 89.):

> *14. Hank has $5.27. A solar calculator costs $14. Let n be how much more he needs. a. What addition equation could be solved to find out how much more money he needs.*
> *b. Solve this equation.*

(*Dramatis personae*: EF = Erin Fry, M = Mike, C= Chris)

M: Hank has 5 dollars 27 cents. A solar calculator costs 14 dollars. Let *n* be how much more he needs.

C: What addition equation, so it's 527 plus N equals 14.

...

C: Solve the equation.

M: Need a calculator?

C: So, uh, we subtract the 527 on both sides. So

EF: Where did you learn how to do that: scratch off like that? Like subtract from both sides?

C: In, uh, grade school.

EF: Really? You guys did that?

...

C: Why? Is there something wrong with that?

EF: No, I just wondered where, I mean if it was in the reading the last day or if you, uh

C: That's the way my teacher showed us how.

...

This conversation suggests that students do use and apply production rules of some sort in doing mathematics. This said, it the following episode shows that students do not compose production rules quickly and easily. In solving the problems below (UCSMP, 1990, p. 89.), students do not recognize the composed production used above and they proceed to recompose it.

*In 17 and 81, **a.** find the solution. **b.** Check your answer.*

17. $3\frac{1}{4} + x = 10\frac{1}{2}$ ***18.*** $15.2 = f + 2.15$

C: OK, 17. Find the solution. Check your answer. So 3 and a fourth plus X equals ten and a half. So that would be 13 fourths. Is that right?

M: 13 fourths. 13 and 3 quarters. No, wait.

C: 13 fourths is an improper fraction.

M: Yeah. Yeah.

C: Then the answer should be 21 over 2.

M: How do you get that?

C: The improper fraction.

M: Oh.

C: Just, I'm changing it to an improper fraction so it's easier to add.

[Notice here that both Mike and Chris have production rules about simplifying mixed numbers]

M: Yeah.

C: Subtract 13 fourths

M: Well, you got to, don't you have to have a common denominator first?

C: Yeah, that's, you're right.

M: Well, it's 4 anyway.

C: So it's 21 times 2 is

M: 42 over 4. It's 13 over 4 and 42 over 4.

C: So 42 minus 13 is 29 over 4

M: 29 fourths or 6 and no

C: 7 and a fourth.

M: 7 and a fourth. *Oh, I see how they got, wait. All they did is they subtracted those two didn't they? Didn't they just subtract, they just* . . . [emphasis added]

C: Yeah.

M: Minus those two.

C: Yeah, I'll check it just in case. I think it's right. 17? 7 and a fourth, yeah. All right do we have B?

M: I just put, yeah, seven and a fourth.

C: Oh, I know how to do it. (pause) 18.

M: 2.15 plus F equal 15.2. That's easy. All you do is [subtract]

Mike and Chris have recomposed the very same production they used to solve Hank's calculator problem, yet they do not recognize it. One possible explanation is that they have more familiar with money than with "algebra." In any event, production compilation is a tenuous process. Despite their freshly recomposed production, Mike and Chris have difficulty applying it to the very next problem:

19. *Solve for d: a + d = c.*

M: A plus D equals C.
C: Solve for D. I don't understand what they mean by solve.
EF: In all those other ones you did it said solve for X or for C or whatever. How did you do that?
C: Oh, yeah, because it says right here, it says to solve for C
...
C: It says to solve for D you have to make D alone on one side. So, oh, that's easy. It's, uh, what is it? Is it subtract A on both sides?
M: Yeah, ...D equals C plus negative A.
C: You subtract A on . . .
M: Exactly
C: C minus A.
M: C plus negative A.
C: Doesn't matter; C minus A. It doesn't matter . . .

There is an explanation for their initial failure to solve the above problem. Up to this point, they had seen few examples involving letters alone, so they could not apply their composed production rule. Thus, one explanation is that their "pattern matcher" failed, that is, perhaps the condition part of their composed production, was flawed. This explanation also is consistent with their performance, two days later, when solving the following problem (UCSMP, 1990, p. 99):

14. *Solve -423 = -234 + y,* Mike and Chris had the following exchange:

M: 14. Negative 423 equals (pause) Oh, that's easy. (pause) Where is your negative?
C: Negative 657.
M: Minus
EF: Does that check?
M: What?
EF: See if that would check when you look at it.
M: It's negative 189.
C: All right, you're right, you're right. ...
M: Y equals negative 189.
C: I don't know what I was doing.

Thus, even though they composed a production rule, it was not complete. These examples illustrate two important results. Students do not correctly create a new production rule, generalize, or "chunk" in just one step as some theories posit--instead, they require repeated exposure. It also suggests that learning new mathematics is more complex than previously thought. In particular, it suggests that we need further elaboration of the relationship between verbal input and the creation of production rules (encoding). Finally, as others have posited (e.g., Rissland, 1978, 1985; VanLehn, 1990), examples play a critical role in learning because of their role in building generalized production rules.

Furthermore, for the most part, students are slow to make the correct generalizations, particularly for novel problems. This was evident in the data above, so only a few more examples are needed to illustrate this. Although the teacher gave an example of simplifying a fractional expression that contained two variables $(\frac{2}{a} + \frac{3}{b})$, Chris could not simplify an expression that contained only one variable, $(\frac{3}{4} + \frac{-2}{x})$. Mike and Chris could solve equalities, but not inequalities; they could solve triangle inequalities, but not triangle inequalities involving terms with x^2's in them; they could simplify $-p + p$, but they could not simplify $-p^3 + p^3$! All of these examples show the difficulty students have in making correct generalizations.

The last part of the framework that emerges from these data is the role of goals in guiding problem solving. There are two aspects to this: they expect the solution to require a specific number of steps and were puzzled if it took more or fewer steps, and the lack of goal verbalization in guiding their solution strategies.

The following two examples should suffice to illustrate the role expected length played in guiding their solutions. When asked to simplify $\frac{30}{a} + \frac{10}{a}$, they immediately responded, "$\frac{40}{a}$, NOT!" When asked to solve $z + (2 + -z) + 6 = z$, they got $2 + 0 + 6 = z$, therefore $8 = z$. Again, they could not believe this was the correct answer. This suggests that for them, doing a mathematics problem meant completing a certain number of steps, rather than reaching a desired conclusion.

The biggest surprise of the study was that the subjects' did not routinely articulate their goals as they solved problems. All cognitive models use some sort of goal structure to control attention while solving problems. To be sure, Mike and Chris did use goals even if they did not verbalize them. The rewrote mixed numbers as fractions, they modeled their solutions on worked examples, and so on. Yet they rarely talked about them. This result, in conjunction with the roles visual clues and expected length play, suggests the following teaching experiment: teach students to verbalize their goals while solving algebra problems and have them recapitulate them afterwards to see if they make fewer errors.

Conclusion

This study shows that developing a fine-grained theory of mathematics learning in the classroom is quite a different proposition from developing coarse-grained theories. Furthermore, developing such a learning theory is hampered by the lack of cognitive models of performance in a knowledge-rich domain. Nevertheless, this study adds to the growing work on process models of mathematics learning. Limitations of current process models and unified theories of cognition have been discussed as well as possible enhancements. A fine-grained theory of learning algebra in the classroom requires more detailed theories of generalization, particularly focussing on the role visual clues and examples play. As many researchers posit (e.g., Davis, 1984, 1986; Fischbein, 1990; Larkin, 1989), progress in mathematics education depends on developing better explicit theories of how people learn mathematics.

Acknowledgments

I want to thank Anne Smith for allowing us to visit her class, her students Mike and Chris, and Gerry Swift for transcribing the audio-tapes. This research was supported by the Bureau of Educational Research at the University of Illinois at Urbana-Champaign, but any opinions or conclusions are the authors' alone.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge: Harvard University Press.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2*, 155-192.

Brown, J. S. & VanLehn, K. (1980). Repair theory: A generative theory of bugs in procedural skills. *Cognitive Science, 4*, 379-426.

Burton, R. R. (1982). Diagnosing bugs in a simple procedural skill. In D. Sleeman and J. S. Brown (Eds.), *Intelligent tutoring systems.* London: Academic Press.

Davis, R. B. (1984). *Learning mathematics: The cognitive science approach to mathematics education.* London: Croom and Helm.

Davis, R. B. (1986). Conceptual and procedural knowledge in mathematics: A summary analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

Fischbein, E. (1990). Introduction. In P. Nesher & J. Kilpatrick (Eds.), *Mathematics and Cognition: A research synthesis by the international group for the psychology of mathematics education.*

Larkin, J. (1989). Eight reasons for explicit theories in mathematics education. In S. Wagner & C. Kieran (Eds.), *Research issues in the learning and teaching of algebra.* Reston, VA: Lawrence Erlbaum & National Council of Teachers of Mathematics.

Matz, M. (1980). Towards a computational theory of algebraic competence. *Journal of Mathematical Behavior, 3*(1), pp. 93-166.

Matz, M. (1982). Towards a process model for high school algebra errors. In D. Sleeman and J. S. Brown (Eds.), *Intelligent tutoring systems.* London: Academic Press.

Newell, A. (1990). *Unified theories of cognition.* Cambridge: Harvard University.

Newell, A. & Simon, H. A. (1972). Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Rissland, E. L. (1978). Understanding understanding mathematics. *Cognitive Science, 2*, 361-383.

Rissland, E. L. (1985). Artificial intelligence and the learning of mathematics: A tutorial sampling. In E. A. Silver (Ed.), *Teaching and learning mathematical problem solving: Multiple research perspectives.* Hillsdale, NJ: Lawrence Erlbaum.

Schoenfeld, A. H. (1985). Making sense of "out loud" problem-solving protocols. *The Journal of Mathematical Behavior*, 4, 171-191.

Silver, E. A. (1986). Using conceptual and procedural knowledge: A focus on relationships. In J. Hiebert (Ed.), *Conceptual and procedural Knowledge: The case of mathematics* pp. 133 - 179. Hillsdale, NJ: Lawrence Erlbaum.

Sleeman, D. (1982). Assessing aspects of competence in algebra. In D. Sleeman and J. S. Brown (Eds.), *Intelligent tutoring systems.* London: Academic Press.

Sleeman, D. (1984). An attempt to understand students' understanding of basic algebra. *Cognitive Science, 8*, 387-412.

Sleeman, D. (1986). Introductory algebra: A case study of student misconceptions. *The Journal of Mathematical Behavior, 5*, 25-52.

Thompson, P. W. (1989). Artificial intelligence, advanced technology, and learning and teaching algebra. In S. Wagner & C. Kieran, (Eds.), Research issues in the learning and teaching of algebra. Reston, VA: Lawrence Erlbaum & National Council of Teachers of Mathematics.

University of Chicago School Mathematics Project. (1990). *Algebra.* Glenville, IL: Scott Forseman.

VanLehn, K. (1982). Bugs are not enough: Empirical studies of bugs, impasses and repairs in procedural skills. *The Journal of Mathematical Behavior, 3*, 3-77.

VanLehn, K. (1986). Arithmetic procedures are induced from examples. In J. Hiebert (Ed.), *Conceptual and procedural Knowledge: The case of mathematics.* Hillsdale, NJ: Lawrence Erlbaum.

VanLehn, K. (1987). Learning one subprocedure per lesson. *Artificial Intelligence*, 31, 1-40.

VanLehn, K. (1990). *Mind bugs: The origins of procedural misconceptions.* Cambridge, MA: MIT Press.