

AD-A204 495

**LEARNING MATHEMATICS FROM
EXAMPLES AND BY DOING**

Technical Report AIP-4

Xinming Zhu and Herbert A. Simon

Chinese Academy of Sciences
&
Carnegie-Mellon University

**The Artificial Intelligence
and Psychology Project**

Departments of
Computer Science and Psychology
Carnegie Mellon University

Learning Research and Development Center
University of Pittsburgh

DTIC
ELECTE
JAN 3 1989
S H D

89 1 03 133

(4)

**LEARNING MATHEMATICS FROM
EXAMPLES AND BY DOING**

Technical Report AIP-4

Xinming Zhu and Herbert A. Simon

Chinese Academy of Sciences
&
Carnegie-Mellon University

29 September 1987

This research was supported by the Computer Sciences Division, Office of Naval Research and DARPA under Contract Number N00014-86-K-0678. Reproduction in whole or in part is permitted for purposes of the United States Government. Approved for public release; distribution unlimited.

DTIC
S **ELECTE** **D**
JAN 3 1988
R H

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY			3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; Distribution unlimited		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE			5. MONITORING ORGANIZATION REPORT NUMBER(S)		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) AIP - 4			7a. NAME OF MONITORING ORGANIZATION Computer Sciences Division Office of Naval Research (Code 1153)		
6a. NAME OF PERFORMING ORGANIZATION Carnegie-Mellon University		6b. OFFICE SYMBOL (if applicable)	7b. ADDRESS (City, State, and ZIP Code) 800 N. Quincy Street Arlington, Virginia 22-17-5000		
6c. ADDRESS (City, State, and ZIP Code) Department of Psychology Pittsburgh, Pennsylvania 15213			9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-86-K-0678		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Same as Monitoring Organization		8b. OFFICE SYMBOL (if applicable)	10. SOURCE OF FUNDING NUMBERS p400005ub201/7-4-86		
8c. ADDRESS (City, State, and ZIP Code)		PROGRAM ELEMENT NO N/A	PROJECT NO N/A	TASK NO N/A	WORK UNIT ACCESSION NO N/A
11. TITLE (Include Security Classification) Learning Mathematics from Examples and by Doing					
12. PERSONAL AUTHOR(S) X. Zhu and H. A. Simon					
13a. TYPE OF REPORT Technical		13b. TIME COVERED FROM 86Sept15 to 91Sept14		14. DATE OF REPORT (Year, Month, Day) 87 September 29	
15. PAGE COUNT 42					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Learning, Algebra, Induction, learning from examples		
19. ABSTRACT (Continue on reverse if necessary and identify by block number)					
Human learning takes place in a wide variety of situations and almost surely employs many different processes. During the past few years, considerable attention has been paid to the ways in which people can learn from procedures by examining worked-out examples and by solving problems -- learning from examples and learning by doing. In this paper we wish to discuss the process of learning from examples, and to discuss a number of experiments with high-school students employing this learning paradigm. The experiments, in addition to comparing this learning strategy with more traditional ones, provided us with information about the processing actually used by students who employed the new strategies. By "traditional strategies" we mean procedures in which the teacher and textbook play an active role in presenting and explaining the material to be learned. (KR) ←					
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION		
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Alan L. Meyrowitz			22b. TELEPHONE (Include Area Code) (202) 696-4302		22c. OFFICE SYMBOL N00014

Learning Mathematics from Examples and by Doing

Xinming Zhu and Herbert A. Simon
Chinese Academy of Sciences
and Carnegie-Mellon University

Human learning takes place in a wide variety of situations and almost surely employs many different processes. During the past few years, considerable attention has been paid to the ways in which people can learn procedures by examining worked-out examples and by solving problems -- learning from examples and learning by doing (e.g., Anzai, 1978; Neves, 1978; Anzai & Simon, 1979; Neves & Anderson, 1981; Sweller and Cooper, 1985). Moreover, it has been shown how such learning can be simulated using computer programs known as *adaptive production systems*, and these computer programs now provide a model for the same processes in human learning.

In this paper we wish to discuss the process of learning from examples, and to discuss a number of experiments with high-school students employing this learning paradigm. The experiments, in addition to comparing this learning strategy with more traditional ones, provided us with information about the processes actually used by students who employed the new strategies. By "traditional strategies" we mean procedures in which the teacher and textbook play an active role in presenting and explaining the material to be learned.

Our interest in the prospects of learning from examples was sparked by incidents like this. A student who was late for class missed the lecture of the teacher but at the end of the class looked at the problems worked by another student. When the tardy student was tested, we were surprised to see that he worked the test problems correctly. Apparently he had learned by studying the worked-out examples. How generalizable is the result? How efficient is the process?

Some Historical Comments

Of course, the idea of learning from examples is not new. Worked-out examples constitute an important element in textbook presentations of new concepts and procedures. These examples can be shown to contain enough information about the procedures to permit diligent students to learn them without additional instruction, a fact that many good students appear to know and exploit. Presumably, the students infer from the examples the essential procedures, then internalize these procedures so that they can apply them to new problems.

Closely related to learning from examples is the alternative process of learning by doing -- that is, learning by solving problems. If feedback is provided on whether an answer is correct, and if the problem solver persists until the problem is solved, then the solution path (however halting and inefficient the search for it may have been) provides a worked-out example. This example can be used as a basis for finding the correct algorithm in the same manner that a worked-out example provided by the teacher can be used.

In the literature of education, the ideas of learning from examples and learning by doing also have a long history. They have their roots in ideas about child-centered education associated with the names of Pestalozzi, Froebel, and Hebart, and in the role assigned to the learner's experiences in John Dewey's prescriptions for progressive education. A number of followers of Dewey and Piaget have been strong advocates of learning by discovery, and there has been a significant amount of research on the conditions under which discovery learning is, or is not likely to be effective and efficient (Shulman & Keislar, 1966). Discovery learning is, of course nearly synonymous with learning by doing.

Psychological research on concept attainment (Bruner, Goodnow, & Austin, 1956), and computer modeling of concept attainment processes (Hovland & Hunt,

1960; Hunt, 1962) have always used a learning-from-examples paradigm, which is perhaps not surprising, since the interest lay in inductive discovery rather than in learning *per se*. In a different domain, Siklóssy (1968) showed that the vocabulary and grammar of languages could be learned by a computer program from examples of sentences paired with semantic descriptions ("pictures") of the sentence meanings.

Psychologists interested in creative processes have also long observed the common element of discovery that is central both to creativity and to learning by doing. In what is perhaps the earliest treatment of creativity from an information processing point of view (Newell, Shaw, & Simon, 1962, first issued as a technical report in 1958), a computer learning program was described that learned by doing, and the importance of hindsight (analysis of problem solutions after discovery) was underscored.

The learning programs we have mentioned have two important characteristics in common: (1) they consist in a gradual accumulation of selective principles that modify the sequence in which possible solutions will be examined in the problem space; (2) the selective principles are obtained by hindsight -- by analysis of the program's successes in its previous problem solving.

This early reference to learning by doing, and the discussion that follows it, foreshadow very closely the form of learning that we will see exemplified in the experiments reported in this paper.

Research on learning within an information processing paradigm received a new impetus a decade and a half ago with Waterman's (1970, 1975) invention of adaptive production systems as a programming model of the learning mechanism. An adaptive production system is simply a computer program that is capable of modifying itself adaptively by constructing new instructions (productions) and adding them to its memory. The productions provide a detailed model of what must be learned in order to perform the task, and the adaptive productive system provides an

For	
I	☒
d	☐
on	☐
on/	

Availability Codes

Dist	Avail and/or Special
A-1	



explanation of how the productions that make up this model can be acquired successively.

Adaptive production systems provided an operational method for combining task analysis (e.g., Gagné, 1965; Gagné & Briggs, 1974; Resnick, 1976) with a specific learning theory, hence of constructing an orderly framework within which materials for learning from examples and learning by doing could be designed, and instructional materials could be sequenced.

Neves (1978) was the first to apply these ideas to school learning, with an adaptive production system that was able to learn to solve linear algebraic equations. Considerable research activity has followed in the subsequent years (e.g., Anzai, 1978; Anzai & Simon, 1979; Neves & Anderson, 1981), and these learning theories are now being used by Anderson (1987) and others in the design of tutoring systems. There have also been new beginnings of experimentation with learning from examples and learning by doing (e.g., Sweller & Cooper, 1985). Our own empirical research on this topic began in 1983.

Scope of this Paper

We have carried out a number of experiments to test the feasibility and efficiency of learning from examples and learning by doing. In the first experiment we shall describe we took verbal protocols from some of the subjects to learn more about the processes they were using. In the remaining experiments we took protocols only for the geometry task, but measured the speed and extent of learning in comparison with the learning of students who were taught the same material conventionally.

This paper, focuses primarily on the protocol material and what it discloses about the processes that students use to learn from worked-out examples and by solving problems. Our main interest was in determining whether, and with what

degree of understanding, school children can learn school subjects by these methods. Since we shall be concerned mainly with the nature and feasibility of the methods, we will present only briefly some preliminary findings on their efficiency in comparison with traditional methods, and will postpone to another occasion a detailed report on outcomes. In addition to elucidating the learning processes that appear in the protocols, we will pay a good deal of attention to the question of whether the methods of learning from examples and by doing encourage rote learning or whether they promote, instead, learning with understanding.

Instructional Materials

Experiments on learning from examples and by doing were carried out with learning materials for a number of tasks contained in the mathematics curriculum of 5th, 7th, and 8th grade students. The tasks included simplifying fractions, factoring quadratic expressions, manipulating terms with exponents, and solving several geometry problems.

The Training Materials. The preparation of training materials was preceded by a careful and detailed analysis of each task. On the basis of this analysis, a production system was constructed, capable of performing the task, to represent the skills that students would acquire in mastering the task. It is assumed that a person able to perform the task possesses a set of productions very similar to those in the model

The training materials were then designed to motivate learning the productions successively. That is, the examples and problems were sequenced so that the initial ones could be handled with a small subset of the productions, while subsequent problems required additional productions for their solution. Thus, in accordance with the usual principles for shaping behavior, learners could attend to one or a few

aspects of the problem situation at a time.

A Production System. The first section of the Appendix sets forth a production system for factoring quadratic expressions of the form: $x^2 + ax + b$, where a and b are (positive or negative) integers. This is the production system we used to construct our examples and problems for teaching this skill. The procedure it describes begins by finding all pairs of positive integers that factor the constant term, b , of the quadratic. Then it divides into cases, according to the signs of the linear and constant terms. Depending on these signs, a pair of factors is selected whose sum (if the sign of the constant term is positive) or difference (if the sign of the constant term is negative) is equal in absolute magnitude to the coefficient of the linear term of the quadratic. Signs are now assigned to the factors, depending on the arrangement of signs in the quadratic. If the constant term is positive, then both factors are positive (negative) as the coefficient of the linear term is positive (negative), respectively. If the constant term is negative, then the larger factor is positive (negative) as the coefficient of the linear term is positive (negative), respectively and the smaller factor has a sign opposite to that of the larger factor.

Fortunately for the feasibility of students' learning to factor, these rules are easier to understand in the context of examples and problems than they are when thus stated in prose, or when put in the production system format of the Appendix. Moreover, the principle underlying the factoring of quadratics can be explained, as we shall see, in a unified way, without invoking four special cases that depend upon the signs of the coefficients. Nevertheless, factoring is usually taught in terms of the four cases, and we think this is probably an effective procedure for focusing the students' attention on one process at a time. First, they learn to search for a pair of factors of the constant term that adds up to the coefficient of the linear term. When they have acquired this skill, their attention is called to the signs of the linear

and constant terms of the quadratic, and the effects of these signs on the selection of the correct factor pair and the assignment of signs to the factors.

The details of the particular production system we describe in the Appendix are probably not important, for the task could be accomplished about equally well with a number of alternative systems. In fact, the protocol evidence will show that the production systems acquired by our subjects are not identical, but are variants of this basic scheme.

The alternative to factoring by recognizing distinct cases for different combinations of the signs of terms in the quadratic is to factor "algebraically," that is, using signed numbers throughout. Then the rule for factoring a quadratic reduces to a single case: Find all pairs of factors (regarding $+a$ and $-a$ as distinct factors) of the (signed) constant term of the quadratic; then select the pair whose algebraic sum equals the (signed) coefficient of the linear term, and write the factorization as $(X+c)(X+d)$, where c and d are the (signed) factors.

$$X^2 + aX + b = (X + c)(X + d)$$

$$\text{where } cxd=b, c+d=a.$$

We shall see that there is evidence that many of our subjects learned the more general principle, even if they used the productions for the four specific cases in their actual problem solving.

Prerequisite Knowledge. To learn to factor quadratic expressions, students must have some prerequisite knowledge (See Appendix. 2. Review Knowledge.) In particular, they must know what factoring means. Before reaching quadratics in the curriculum they have already studied the factorization of integers and of monomials in algebraic expressions. We convey additional information about quadratics to them by defining factoring a quadratic expression into two linear expressions as the inverse of the operation of multiplying the two linear expressions to form the quadratic. Since

the students already have the skill of multiplying algebraic expressions, this definition of the task of factoring should be understandable to them and can also provide them with a means for checking (by multiplying back) whether they have factored correctly. Hence, in learning either from examples or by doing, they can obtain feedback on the correctness of their answers. This feedback provides essential information for the learning process.

From the definition of factoring as the inverse of multiplication, and from the worked-out examples of multiplication in the training materials, students can derive the "algebraic" algorithm for factoring that we presented above. For they can observe that the constant term in a quadratic is the product of the factors, while the coefficient of the linear term is the (algebraic) sum of the factors. To facilitate their noticing these relations, we include in the review knowledge several problems of factoring integers and noting the sums of the factors. We also carry out the multiplication in steps, so that the summation of the factors to form the coefficient of the linear term of the quadratic is made explicit.

Awareness of these relations can provide guidance to the subjects' search for solutions in the learning-by-doing condition as well as in the learning-from-examples condition. As we have seen, the laws of signs in factoring, which we have represented by separate productions, are also implicit in the definition of factoring. In principle, students should be able to infer the laws of signs from the general definition, hence to find a rationale for each of the productions that embody these laws, and many of our subjects did accomplish this.

In presenting the prerequisite knowledge prior to the learning trials we go a little further than simply defining factoring. We also give the students practice in finding all the pairs of integral factors of a positive integer, and practice in finding the sum of each pair. In presenting the product of two linear expressions, we could

also have called attention explicitly to the fact that the linear term of the quadratic is the algebraic sum of the two constant terms of the linear expressions, while the constant term of the quadratic is the algebraic product of the two constant terms of those expressions. In fact, it did not prove necessary to provide this additional review knowledge in order for students to acquire the requisite skills from the examples and problems.

Experiment 1

Method

The task was to learn to factor quadratic expressions in algebra, of the form $x^2 + ax + b$, where a and b are positive or negative integers, and the factors are also integers. In these experiments we consider only expressions in which the coefficient of the quadratic term is $+1$.

Two groups of subjects were used in the experiment: a group of 20 who provided verbal protocols while they were run individually, and two classes, totalling 98 students, who were run in the classroom setting without taking verbal protocols. The first group (*protocol subjects*) provided us with rich data about the processes they used. Half of them worked in a learning from examples condition, and half in a learning by doing condition. (One of the learning-by-doing protocols was lost, so our protocol analysis for this condition contains only nine subjects.)

The second group (*classroom subjects*) provided us with a larger sample of data on learning speeds and levels of learning, all of them working in the learning-from-examples condition. In this initial experiment, no comparison was made with a control group using conventional learning methods. Our initial goal was to determine whether learning by example or by doing was feasible for a standard algebra skill, and to gain an understanding of the learning process.

The subjects were approximately 13 years old, enrolled in the first algebra course in a middle school in Beijing, China. The school is near the middle of the range of Beijing schools in terms of student ability. In the standard algebra curriculum, two class sessions are usually devoted to teaching the factorization of quadratics, and homework is assigned for each evening.

The procedures were as follows (See Appendix):

1. Take pretest
2. Review knowledge prerequisite for the task
3. Take second pretest (five problems)
4. Learn from examples or by doing
5. Take final test (identical with second pretest)

Results

Classroom Subjects. We begin with a brief report on the performance of the subjects who learned in the classroom without providing protocols. Since 4 of the classroom subjects could solve the pretest problems, their performance was excluded from the data we analysed. Of the remaining 94 subjects, 78 (83 per cent) solved all the test problems correctly after reviewing the prerequisite knowledge and working through a few examples and problems. Eight (8.5 per cent) of the subjects solved only three of the five test problems correctly and 8 subjects (8.5 per cent) solved two or less problems. These results were achieved within the time of a single class session, demonstrating that learning this skill from examples is wholly feasible well within the time allowed in the curriculum.

Protocol Subjects. The protocol subjects were run individually, and asked to talk aloud while they were doing the work. Their protocols were tape-recorded and

transcribed, and for the purposes of this paper have been translated literally into English. Ten of the protocol subjects were provided with materials that included worked-out examples, the other 10 only with a carefully arranged sequence of problems to be solved (Item 4 on list above). None of the subjects could solve any of the pretest problems (Items 1 and 3) or build appropriate procedures. After about 25 minutes spent in working through examples and doing problems, all twenty experimental subjects could solve the test problems (Item 5) and 15 of the 20 answered correctly the questions about the procedures. In the following paragraphs, we will examine their protocols in more detail.

The Protocols: Part I. From the protocols, we obtain a rather clear picture of how the subjects learn from the examples. We will begin by discussing the protocols of two subjects on Part I of the materials: E1, who was learning from examples (E condition); and D7, who was learning by solving problems (D condition). Their protocols on Part I are shown in Tables 1 and 2, respectively.

Protocol of E1 on Part I
Learning from Examples

[For the problems, see the Appendix, 4A, Part I.]

[Subject looking at first exercise]

1. According to the example,
 2. Exercise 1, $X^2 + 11X + 18$ is equal to $X \dots$
 3. [Looking at example 1] $2+3$ is equal to the coefficient 5,
 4. 2×3 is equal to the constant 6.
 5. [Looking at example 2] This example is $1+6=7$,
 6. $1 \times 6 = 6$.
 7. Exercise 1 [$X^2 + 11X + 18$] is that $X+3$ multiplies $X+6$.
 8. Exercise 2 [$X^2 + 9X + 18$] is that $X+6$ multiplies $X+3$.
 9. Exercise 3 [$X^2 + 19X + 18$] is that $(X+9)(X+2)$.
- [Reviewing exercise 1]
10. That is wrong.
 11. $3+6$ is not equal to 11.
 12. Exercise 1 is $X+2$ multiplying $X+9$,
 13. $2+9=11$
 14. $2 \times 9=18$
 15. This one [exercise 2] $3+6=9$,
 16. $3 \times 6=18$.
 17. This one [exercise 3] is also wrong.
 18. It should be $X+1$ multiplies $X+18$,
 19. $1+18=19$
 20. $1 \times 18=18$.

Protocol of D7 on Part I
Learning by Doing

[For problems, see the Appendix, 4B, Part I]

[Subject is doing the first exercise]

1. [Exercise 1] $X^2 + 5X + 6 = (X+) (X+)$
2. equals $(X+11)X+30$.
3. According to multiplication of polynomials, $(X+11)(X+30) \dots$
4. their product $\dots$
5. That is wrong.
6. It should be $(X+2)$ multiplied by $(X+3)$.
7. Since $X \times X$ equals X^2
8. $2 \times X = 2X$
9. $3 \times X = 3X$
10. $2X+3X=5X$
11. 2 multiplied by 3 equals 6.
12. [Experimenter] What rule have you found?
13. I found a rule.
14. The numbers in these brackets...
[the constants of the two factors]
15. added to each other, multiplied by X , we'll get this number
[the coefficient of the linear term].

16. These two figures multiplied by each other will equal the number at the end [the constant].

E1's Learning. E1 begins by looking at the first exercise (2), then looks back (3-6) at the first two of the five examples provided, and checks that the sum of the numerical factors equals the coefficient of the linear term of the quadratic, and the product of the factors equals the constant term. The subject then proceeds to solve Exercises 1-3 (7-9), but pays attention only to the second of these two conditions, thereby obtaining wrong answers. In Lines 10-11, he checks the first condition for Exercise 1, and sees it is not satisfied. He now (12-14) finds a pair of factors that satisfies both conditions. Then he goes back (15-20) and corrects Exercises 2 and 3. By the time he has finished, he has essentially built the production rule for positive coefficients.

The examples provided enable E1 to review the conditions, learned previously from multiplication of linear terms, that the constant term in the quadratic equals the product of the constants in the linear terms (the factors), while the coefficient of the linear term in the quadratic equals the sums of these factors. However, possibly due to limits on attention span, he constructs erroneous solutions that only take into account the former of the two requirements. His understanding that factoring is the inverse of multiplication does allow him to check his results and to discover that they are wrong.

From the protocol we see clearly how E1 uses previously acquired knowledge both to generate his solutions and to test them. We also see the difficulty he encounters in using all the knowledge he has. He initially generates only one pair of factors that satisfies the product condition without considering that it may not satisfy the sum condition. Only after discovering his error does he generate additional pairs and test each one for the sum condition, thereby debugging his incomplete procedure.

D7's Learning. D7, provided only with exercises to work, remembers (1-2) that

sums and products have something to do with the process, but applies the rule backwards (2), adding and multiplying the two coefficients in the polynomial, and getting a wrong answer for the first problem. Apparently checking the answer (3-5) by the inverse operation of multiplying, he then (6) arrives at the right answer, which he checks by multiplication (7-11). He is now able to state the production rule (14-16).

D7's protocol shows that he was able to learn to factor without being provided with any worked-out examples, simply by applying his knowledge (1) that factoring is the inverse of multiplication, and (2) that he must make use, somehow, of sums and products of factors. Multiplying out the erroneous factors that represent his attempted solution of the first problem, he rediscovers what rules he should apply.

Progression in Learning. These two protocols are not unrepresentative of the behavior of the subjects in both conditions at the beginning of the session. An efficient procedure would be to factor the constant term of the quadratic first, then find a pair of factors that sum to the coefficient of the linear term. However, initially, most subjects (5/10 in E condition, 7/9 in D condition) look first for two numbers that add to the coefficient of the linear term. This preference may reflect the fact that, on scanning the quadratic from left to right, the subjects first notice the linear term, derived from the factor sum, then the constant term, derived from the product. By the end of the session, all the D subjects and all but one of the E subjects are using the more efficient procedure of satisfying the product condition before the sum condition.

Often, subjects do not initially check whether the numbers they have found to satisfy one of the conditions also satisfy the second condition. (4/10 E subjects and 5/10 D subjects fail to make such a test on the early exercises.) As we have seen, E1 makes this error twice on the first three exercises, then notices his mistakes and

corrects them. The error is usually noticed when the subject checks his or her answer by multiplying the factors (E1, Lines 10-12).

If the subject does not check the answer, the failure to satisfy the second condition may go unnoticed. Thus, the protocol of Subject D4 on the second training problem of Part I reads:

" $X^2 + 7X + 6$, equals...[pause] That's 3, 4. By adding them we get the coefficient of the second term. So I chose 3 and 4. $3 + 4 = 7$."

In this case the error is not corrected. Such errors, where one of the conditions is violated, are rare in the latter parts of the session.

In the earlier parts of the session, subjects frequently check their answers explicitly by multiplying out the two factors to obtain the original quadratic expression. Later on, as they become skilled and accurate, this check, if it is made, is sometimes not verbalized. Two examples will illustrate this difference among subjects when they are solving some of the later problems. On Exercise 7 of Part III, Subject D3 says only:

" $X^2 + 7X + 10 = (X + 5)(X + 2)$ "

The protocol of Subject E3 on the same problem reads:

" $X^2 + 7X + 10$. 10 can be factored into 2 times 5. 2 plus 5 equals 7, which is the linear term's coefficient. So this problem equals $(X + 2)$ times $(X + 5)$."

Through all the thirteen exercises of Part III, the protocols of both subjects were essentially identical with their protocols for Exercise 7 (with the appropriate numbers inserted in the grammatical "slots"). There is no reason to suppose the one subject had a better understanding of the process than the other: both were able to state the rule quite correctly and clearly. One was simply more taciturn than the other, or had, perhaps, automated the solution process more completely.

All subjects appeared to recognize that the constant term of the quadratic

could be factored in more than one way, but they did not always find the pairs of factors systematically, and in a number of cases took some time to realize that N can be factored into $1 \times N$. When they sought to satisfy the sum condition first, they often proceeded rather systematically on early problems. Later, on most problems, subjects mentioned only the pair of numbers that satisfied both conditions; the selection process was not verbalized except when they got stuck. A rare exception is represented by D2 on Exercise 5, Part III:

[$X^2 + 11X - 12$] "Equals $X \dots$ Factor 12 into 3×4 , $2 \times 6 = 12$. We don't get $11X$. Factor it into 1 and 12. X plus... X minus 1 times X plus 12."

On that same problem, D5 merely said:

" $X^2 + 11X - 12$. $X - 1$ times $X + 12$."

Differences Between E and D Conditions. From what has been said thus far, there do not appear to be large differences between the subjects in the two conditions, although the subjects in the D condition had to infer the factorization process without the help of worked-out examples. What differences appear are most clearly visible in the problems of Part I. The subjects who are learning from examples typically examine several of the examples with some care, and quite often refer back to them while working the first exercises. A clear case is Subject E8:

"Example 1: $X^2 + 5X + 6 = (X+2)(X+3)$. This is factoring of the constant, 6, into the linear term, $5X$. It is, factor 6 into 2×3 . Then the sum of the two factors is just the linear term: $2+3=5$, $2 \times 3=6$, just right. Consequently, plus X with this number, and then multiply... So it equals $(X+2)(X+3)$."

"Example 2: The same as the one above. Factor the constant, 6. The product of the two factors must equal this. The sum of the two factors must equal the linear term's coefficient, 7. It's just good to factor 6 into 1×6 . So $1+6=7$ equals the coefficient of the linear term. Finally, multiply the two factored polynomials. The answer is $(X+1)(X+6)$."

"Exercise 1: This is the same as the example. Factor 18 into two numbers, letting their sum equal 11. 18 can be factored into 2×9 . $2+9=11$, which just equals the coefficient of the linear term, 11. So also factor X . Equals $(X+2)(X+9)$."

Notice (and this is common among the subjects in the E condition) that E8 does not work mechanically from the examples, but has clearly in mind the principle that factorization of a quadratic is the inverse of multiplying two linear polynomials. This is especially evident from his comments on Example 2. For the D subjects, understanding this principle is essential to solving any of the problems. So, Subject D3 says, on the Training Problem 4 of Set 1:

" $X^2 + 7X + 12$. There are supposed to be two numbers in the parentheses. Their sum is 7, their product is 12. Equals $(X+3)(X+4)$."

D subjects who have not digested the principle from the definition of factoring quadratics have considerable difficulty with the initial training problems. Consider Subject D2:

"Training 1: $X^2 + 5X + 6 = (X + \quad)(X + \quad)$. Plus 5. Hmm, wrong. Factor 5 into 2... X. Factor 6 into 2×3 . $3X$.

"Training 2: $X^2 + 7X + 6 = (X + \quad)(X + \quad)$. Factor 6 into... $X^2 \cdot X^2$... plus 2. X, hmm, wrong. Need one more X, $2X$. Since $2X$ equals X^2 , then $3X$. $7X = 2X + \dots$ Need one more X. $X^2 + 3X + 4X \dots$ I don't know how to solve this problem."

He solves Problem 3, falters on Problem 4, but thenceforth seems to understand the principle and solves the problems successfully.

Knowledge of Rule. By the end of Part I, almost all the subjects (9/10 E subjects and 8/9 D subjects) can, in response to the experimenter's question, state fairly accurately the rule they are following in factoring, but often in the form of an example, rather than a generalization. Thus D4, in answer to the question "How do you factor?"

"Use multiplication. 2×3 . If satisfied, see the linear term. See if $2 + 3$ equals the coefficient of the linear term. If yes, that's right. Thus it satisfies the problem. We got the coefficient of the linear term by addition."

In reply to the same question, E7 says:

"The sum of the two factors is the coefficient of the linear term. The product of two factors is the constant. All are the same."

Some answers refer explicitly to factoring as the inverse of multiplication. E4 says:

"Factoring equals multiplication formation -- its inversion. Factoring is to get numbers by modification of the multiplication equations; to go back to get the initial ones before utilization of the multiplication process, i.e., before the calculation."

Some subjects in the E condition perhaps lean too heavily on the examples.

Thus, E6, during Part II:

"For the problems in these two parts, read the examples first, then do the exercises. If I cannot solve the problems, I compare the exercises with the examples carefully."

It is not at all certain that this kind of direct comparison of exercises with examples will lead to understanding of the principles. Most of the time, fortunately, subjects seem to try to extract the principles from the examples before they go on to the exercises. Some thought needs to be given to designing the learning materials so that the latter strategy, rather than the former, will be encouraged.

The Protocols: Part II. On Part II, subjects compare the first and second examples, and note that a negative sign in the linear term produces negative signs in both factors. Many subjects are soon able to verbalize the rule (E10):

"Well, mainly one has to see the linear term. If its coefficient has a plus sign, then the factors of the constant have plus signs... But if there is a minus sign in front of the linear term, then the two factors both have minus signs."

Some protocols provide evidence that the reason for the factor's being negative is understood. Thus, E5:

"The second one [the linear term] has a negative sign. So the latter two terms [the constants of the factors] must have negative signs...because minus, minus... These two both have negative signs."

In the D condition, the production for Part II is learned almost in the same way -- by noting that the first and second problems differ only in the sign of the linear term. A little trial and error shows that if the linear term is negative, the signs of both factors should then be negative. When asked what rule he had found, one subject (D7) replied:

"For example, $X^2 - 9X + 18$. $9X$ is negative, 18 is positive. In the parentheses they are all (X _). (X_3)(X_6). The signs in parentheses should all be minus "

The subjects understand the reason for the negative factors, even when examples are not provided. D5, for example, is quite explicit, saying:

"According to the formula which I mentioned a moment ago, since the sum of these two numbers, 1 and 6, is the coefficient of the linear term, and the coefficient of the linear term is negative, so these two numbers should be negative. Negative times negative is positive."

In Part II, a few subjects give evidence that they are using the algebraic algorithm, which does not require them to distinguish different cases for the different arrangements of signs in the quadratic. This can be seen in E3's summary of the rules for Part II:

"I have found from here that one might factor the constant into these numbers. It might be factored into four pairs of numbers, 2 times 9, 3 times 6. Three pairs of numbers, 6 might be factored into 2 times 3. Factor the constant into a pair of numbers. If the sum of these two numbers equals the linear term's coefficient, these two numbers factored from the constant would be good." [Experimenter: "What rule have you found from this problem?"] E3: "It has been found from this problem that 6 can be factored into minus 2 times minus 3. The sum of minus 2 and minus 3 just equals the linear term's coefficient."

In general, the problems in Part II were not difficult for subjects in either condition. Where difficulties were encountered, they generally had to do with finding the correct pair of factors, rather than with assigning the correct positive or negative signs to them. Nine of the E subjects and seven of the D subjects were able to verbalize rules for handling the problems of Part II.

The Protocols: Part III. On the problems of Part III, we again observe some trial and error. Subjects become aware that the assignment of signs to the factors depends upon the pattern of signs in the quadratic, and gradually formulate the correct rules. Checking by multiplying the factors of the proposed answer reveals errors and allows the search to be successful.

Typically, the fact that the factors must have opposite signs is learned before it is understood which factor should have the positive sign and which the negative. A characteristic first attempt at a (incorrect) rule is (D7):

"If the problem is $(x^2 + \quad)$, then in the first parentheses it will be $(x \quad)$; in the second parentheses it will be $(x + \quad)$."

After making and correcting several mistakes, using this rule, the subject corrects the rule:

"As far as the polynomial with successive subtractions $[x^2 \quad]$, its negative factor is larger than the positive factor. If the sign after x^2 is "+", and the next sign is "-" $[x^2 + \quad]$, then also one factor is positive, the other is negative; but its positive factor is larger than the negative factor."

Table 3 gives the protocol of D6 on the first two exercises of Part III. In the first exercise, D6 attends to the constant term of the quadratic (1-3), but fails to notice the sign of the linear term (4-5). He now notices that Exercise 2 is identical to Exercise 1, except for the sign of the linear term (7-9). This observation leads to comparing the solutions of the two problems (10-12) and the recognition (13-15) that his answer to Exercise 1 is wrong. Now the two answers are corrected (16-19), and the conditions for the coefficients of the linear terms are restated (20-27).

Protocol of D6 on Part III

1. [Working exercise 1 in Part III] $x^2 + 5x - 6 = (x + 1)(x - 6)$
2. So it should be minus.
3. It should be $(x + 1)(x - 6)$.
4. Because -6 plus 1 equals 5
5. so it can match $5x$.
6. Positive multiplied by negative is negative too.

7. [Working exercise 2] $x^2 - 5x - 6 = (x - 1)(x + 6)$
8. $-5x$ is negative
9. 6 is also negative
10. It should be $(x-1)(x+6)$.
11. Oh, it is right,
12. because $-5x$ is negative
13. It is wrong,
14. it is positive.
15. I'm wrong in problem 1, $x^2 + 5x - 6 = (x - 1)(x + 6)$.
16. It should be $(x-1)(x+6)$.
17. Problem 2, $x^2 - 5x - 6 = (x - 1)(x + 6)$
18. It should be reversed.
19. It should be $(x+1)(x-6)$,
20. because -6 plus 1 is equal to -5 ,
21. while this is $+5$.
22. $x^2 + 5x - 6 = (x - 1)(x + 6)$
23. so it can't match.
24. In the second problem, $x^2 - 5x - 6 = (x - 1)(x + 6)$
25. $+1$ plus -6 is equal to -5
26. It can match the number $-5x$
27. and $+1x-6$ also equals -6 [the constant].

The fact that the factors must be written down in some order causes some subjects to attend to this feature rather than to the relative sizes of the two factors. and consequently to arrive at a faulty generalization. Here is an example of such a faulty production (E10):

"From the above, it is clear that in the polynomial there is only one minus sign, and it is in front of the constant, then the front one of its two factors should be with a minus. And if there are minus signs in front of the first-order term and in front of the constant, then the factor which has the minus sign is the second one."

This error might be avoided if the pretraining included examples showing that when the order of factors to be multiplied is reversed the same quadratic polynomial is produced.

Comparison of protocols of the same subject from Parts I and III shows that subjects gradually chunk the procedures they are learning into larger integrated units. So, in Part III, subjects typically verbalize their choice of signs, but verbalize only the result of their choice of factors (E10):

"In Problem 4, because the polynomial has two minus signs, therefore the minus sign should be in the second factor. So $(x+3)(x-4)$."

In this protocol, as in many of the other Part III protocols, there is no evidence to show how the subject searched the factor pairs to find the one that also satisfied the sum condition. Some subjects have also automated the assignment of signs in the same manner, verbalizing only the final answer.

The subjects found the rules for Part III problems substantially harder to verbalize than the rules for the problems of Parts I and II. Only 5 of the 9 D subjects and 6 of the 10 E subjects were able to provide reasonably accurate rules (sometimes in the form of examples) for the problems of Part III, although in the post-test, all subjects showed that they could solve these kinds of problems correctly.

E3 provides a clear statement of the rule:

"For these problems, some have positive linear terms and negative constants. Thus, one factor is positive and another one is negative. If the linear item is positive, the larger factor is positive and the smaller one is negative. If the linear term is negative, the larger factor is negative and the smaller one positive."

Discussion

The test results and protocols indicate that as long as the examples are appropriate and the problems are well arranged, most students master the skill of factoring quadratics quite well within a short time by working through examples and problems, after being given knowledge that factoring is the inverse of multiplication. At the end, all of these middle school students had acquired about five productions, could use these productions to solve additional problems, and in most cases could verbalize them fairly accurately. If not elegantly.

The Learning Process

The process that subjects in the E condition use to acquire a production is approximately this: first they compare the polynomial with the factored expression to see what change has taken place. Based on their previous knowledge of multiplication of

polynomials, and by trying to multiply $(X+c)(X+d)$, they notice that c and d add to the coefficient of X in the quadratic, and that c times d equals the constant. From this knowledge, they are able to compose the first production.

Similarly, from the examples in Part II, they notice that a change in the sign of the linear term changes the sign of the factors. This enables them to state the conditions and action of the next production. The examples in Part III lead to further discrimination of the four distinct cases on the basis of the pattern of signs in the quadratic.

Because the subjects in the D condition are provided with no worked-out examples, they must use their knowledge that factoring is the inverse of multiplication to solve the initial problems. The protocol of D7, shown above, gives a typical instance of earlier faltering, but rapid progress as soon as one or two problems are solved.

What Understanding is Achieved? Since the product of the learning experience is a set of rules that most students can verbalize, should we conclude that the students have merely memorized these rules? We believe this is not a correct conclusion, but that, instead, they understand the process -- the learning has been meaningful. There are a number of reasons for this interpretation. To discuss them, we must say what we mean by "understanding."

We test whether someone understands knowledge by determining whether he or she can use it in appropriate ways. Knowledge can be understood shallowly or at great depth, the depth being measured by the range of tasks that can be performed with its help. Thus, one test of understanding of factoring is to ask someone to solve factorization problems. Another test is to ask for the relation of factoring to multiplication. A third test is to ask for a derivation of the sum and product rules for the factors from the definition of factorization as the inverse of multiplication. A

fourth test is to ask for a derivation of the pattern of signs in the factors from the pattern of signs in the quadratic. A fifth test is to ask for a verbal statement of the procedures for factorization.

Even greater depths of understanding can be probed. Tasks can be proposed to see if the skill of factoring quadratics of the form " $X^2 + aX + b$ " can be transferred to quadratics of the form " $X^2 + aXY + bY^2$ " or " $aX^2 + bX + c$." Or, in another direction, students might be asked to explain the relation between the factors of " $X^2 + aX + b$ " and the solution of the equation " $X^2 + aX + b = 0$." As one measure that the semantics of the factorization have been understood, students could be shown a rectangle with dimensions $(X+a)$ and $(X+b)$, and asked to identify the subareas of the rectangle that correspond to the terms of the quadratic: X^2 , aX , bX , and ac . At the level of a university course in abstract algebra, other questions, much deeper than these, can be posed.

In terms of this notion of understanding, it is clear that the subjects in this experiment had not merely acquired rote learning about factoring quadratics, but had acquired a considerable depth of understanding -- some, of course, a greater depth than others.

The first evidence of their understanding is that they can do more than merely recite the rules: they can and do also apply them skilfully by making the appropriate perceptual discriminations and then carrying out the appropriate actions. They recognize when each rule is relevant, and then use it.

Second, the students demonstrate understanding by using the basic meaning of factorization to check their trial results by multiplying back. They have seen how the corresponding patterns of signs in the quadratic and its factors derive from the nature of the multiplication process for polynomials. They have also seen how the sum and product of the factors appears in the resulting quadratic. The students in

the D condition have even demonstrated that they can acquire these skills by derivation from the basic definition, without the help of examples.

For these reasons, we must conclude that the students have learned a good deal about the semantic meaning of factorization, as well as the skill of doing it, and that it is from this knowledge that they generate their versions of the verbal rules that govern it. Since the rules were never given to them explicitly, they have certainly not simply memorized them and recited them by rote.

For some of the students (but probably only a few) understanding was limited largely to being able to factor. Most of them, however, as our account has shown, also understood why the rules worked as they did. We did not probe deeper levels of understanding -- how broadly the new skill would transfer, whether they could understand a geometric representation of the factoring of a quadratic, or whether, which is highly doubtful, they had acquired any of the understanding that would be expected in a course on abstract algebra. They did appear to understand most of the things that first-year algebra textbooks undertake to teach about factoring quadratics.

In sum, Experiment 1 demonstrated that students can learn how to factor quadratics by studying examples and/or by working problems, and that they can do it relatively quickly and efficiently if care is taken to arrange the examples and problems so that they do not make too many errors of induction or require too much trial and error search.

Additional Experiments

The success of our initial experiment, completed in 1983, encouraged us to extend the inquiry in several directions. First, we wished to replicate the experiment on factorization of quadratics, but including comparisons with control groups of subjects who were taught by lecture in the normal way. Second, we wished to generalize the result to other tasks. Since our main purpose in this paper is to discuss the processes of learning from examples rather than the efficiency of this procedure under classroom conditions, we will give only a brief summary of the findings of these experiments, leaving a detailed analysis to a later report.

With respect to replication, we obtained very similar findings on the factorization task in three other schools, and results that were favorable in comparison with control groups, as we shall see in a moment. As for new tasks, we constructed and tested materials for a task of manipulating exponents, a geometry task, a lesson in ratios and proportions, and three lessons in fractions. For all tasks, the students in the experimental condition were usually more successful, and usually took less time to learn, than the students in the control condition.

Replications of the Factorization Task

In School A, with 32 students in each condition, the average post-test score of the experimental students was 93.13% correct, and of the control students, 75.50% correct. The difference was significant at the .001 level, by t-test. In school B, with 39 students in the experimental condition and 38 in the control condition, the average scores were 97.23% and 95.08%, respectively. Here the difference is not significant. In School A, the control students took slightly (but not significantly) longer to learn than the experimental students. In School B, there was no difference in learning times. (The average times in School A were about 40 minutes, and in

School B about 30 minutes.)

The Exponent and Geometry Tasks

In the experiments with exponents, students learned the three formulas: $a^m \times a^n = a^{m+n}$; $(a^m)^n = a^{m \times n}$; $(ab)^m = a^m \times b^m$. On problems involving the first formula, 32 of 36 students (89%) scored 80% or higher; on the second formula, all 36 students scored 80% or higher; on the third formula, 21 of 35 students scored 80% or higher, and all 35 scored 60% or higher.

These results were replicated in 1985, with comparisons with control subjects. The experimental subjects scored higher than the control subjects on all three laws of exponents (significantly higher at the .05 and .001 levels on the second and third laws), and required only 180 minutes for learning, as compared with 270 minutes for the control subjects. The average scores of the experimental subjects for the three laws of exponents were 93.1, 95.4, and 90.2, and for tasks involving several of the laws together, 88 (as compared with 73.1 for the control subjects).

The geometry task involved proving properties of triangles, parallelograms, and their components. The experiment was run in 1984 in four different schools, improvements being introduced into the training materials on the basis of the experience gained in the first two schools. The average scores in the four schools were 76.9%, 77.9%, 89.1% and 83.5%, respectively. In the last two schools experimental groups had slightly higher (but not significantly higher) scores than the control groups (89.1% versus 85.2% in School M, and 83.5% versus 83.2% in School N). Significantly more students acquired the skill in the experimental groups (94.6% versus 88% in School M, and 90.6% versus 85% in School N). Moreover, the average learning times required by the experimental groups were substantially shorter than the control groups (50 minutes versus 75 minutes in School M, and 50 minutes versus 70 minutes in School N).

In a 1985 replication of the geometry experiment in one school, the average score of the experimental group was 89%, and of the control group, 83%, a non-significant difference at the .05 level. Both groups required the same class learning time (50 minutes), but the control group spent an additional 20 minutes doing homework.

Ratios and Fractions

The results in 1985 with lessons on ratios and proportions and fractions were similar to those of the other experiments. The experimental groups performed as well as, or slightly better than, the control groups on all tasks, and required substantially less time to learn the tasks of manipulating fractions.

Results of Retests

In comparing different instructional methods, the relative durability of the skills attained by the one method or the other is of the greatest importance. Retention also bears upon the question of whether the learning was rote or meaningful, for Katona (1940) has shown that material learned with understanding is better retained than material learned by rote.

Retention after one year was tested for two of the tasks discussed above: factorization and geometry. A class that learned the factorization of quadratics in 1984, with an average score of 97.2 obtained an average score of 96.0 when retested in 1985. The control class, which had an average score of 95.0 in 1984, had an average score in 1985 of 92.1. Thus, retention was good for both groups but perhaps slightly better for the experimental group.

With respect to geometry, experimental and control classes that scored 89.1 and 85.2, respectively, in 1984, scored 96.3 and 86.7, respectively, in 1985 -- showing some advantage for the experimental subjects.

All of these additional experiments further demonstrate the feasibility of teaching students through worked-out examples and carefully designed sequences of problems. While the differences in success between these methods and traditional methods of instruction were generally modest in magnitude, they were positive, and they were generally achieved with a saving -- sometimes substantial -- in time.

The geometry experiment showed the importance of designing the instructional materials carefully. By redesigning the materials, we were able to enhance the learning performance of the students, and in particular, to enable most of the weaker students to acquire the ability to solve the problems.

As we pointed out earlier, our analysis of the protocols of students in the first experiment gives us strong reason to believe that the students are learning in a meaningful fashion, and not simply by rote. This conclusion is borne out by the high levels of retention of the skills over one year.

Experiment With a Full Course

Within the past two years, materials have been completed for teaching, by the learning-from-examples method, the standard curriculum in the Chinese middle schools of two years of algebra and one of geometry; and these materials have been used, following the method, with a Beijing middle-school class judged to be of "average" ability. In two years, most students in the class have completed the entire three-year curriculum. We will be able to say more about these results when we have analysed the scores of these students on the standardized mathematics tests given to children all over China, which our students took in June 1987.

Even if the evidence becomes clear that a time saving of one-third has actually been achieved, the findings will need to be interpreted with caution. It is one thing to conclude that children can learn from examples, and can do so efficiently. It is quite another matter to conclude that the time saving was caused by the method of

Instruction used.

This is especially so since other experiments in algebra instruction in Chinese schools, conducted by Professor Zhongheng Lu and his colleagues in the Institute of Psychology, Chinese Academy of Sciences, in collaboration with middle-school teachers, have shown comparable savings. Professor Lu's experiments, begun in 1965 as an investigation of the effectiveness of American methods of programmed instruction, soon took a distinctly new turn. Interrupted by the Cultural Revolution, they were resumed about 1975, and have continued on a considerable scale since. The methods used and the findings are reported in detail in some 121 papers reprinted (in Chinese) in Lu (1983, 1984, 1986).

In the curriculum designed by Professor Lu, which is called "guided self-study," students are provided with a very complete textbook, which contains thorough explanations of each topic, followed by examples, exercises, and tests. Students study the textbook and work the problems during the class period, with a minimum of lecturing from the classroom teacher.

Thus, Professor Lu's approach is different in important respects from ours. What the two methods have in common is that the primary responsibility for learning is placed on the student, who spends his time actively reading and solving problems, and not passively, attending to an instructor. Of course, there is also the possibility that the effectiveness of both methods stems from a "Hawthorne effect," the students being motivated by the knowledge that they are the subjects of special attention and are using novel methods of learning.

All of these possibilities are quite compatible. There is ample evidence that students learn only when they are attending to the task. The motivation for appropriate attention may come from many sources. One of the important ones, evidently, is the availability of materials that allow students to adopt active strategies

in their learning, so that a maximum of time is devoted to relevant activities. We make no claim that learning from examples is the sole route to this goal, although the evidence thus far suggests that it is a very effective one.

Conclusions

Our experiments provide substantial evidence for the possibility of teaching a number of different skills in mathematics by presenting students with carefully chosen sequences of worked-out examples and problems, and without lectures or other direct instruction. In learning by these methods, the students were at least as successful as, and sometimes more successful than, students learning by conventional methods, and in most cases they learned in a shorter time.

Although the students were usually able to state the rules they had learned, they had not simply memorized these rules, for they were able to recognize when they were applicable to a problem and then to apply them. Moreover, protocol evidence shows that they understood the semantic meanings of the rules in terms of their derivation from the fundamental definition of factorization. Hence we may say that the learning that took place was learning with understanding and not merely rote learning.

Materials for teaching three years of middle-school mathematics by the learning-from-examples method have now been developed, and have been tested, thus far very successfully, with a middle-school class in Beijing. As next steps in the study of learning from examples and its practical applicability to school instruction experiments will need to be carried out on a wider range of school subjects, and further tests will need to be made of retention and transfer of the skills acquired.

Since the processes of learning from examples are well adapted to computer aided instruction, the method has considerable potential for wide use. Whether

computers are used, or simple paper-and-pencil materials, the methods described here give promise of freeing considerable teacher time for working with individual students who are having special difficulties.

Appendix: Materials for Experiment 1

Production System

As a basis for preparing the training materials, a detailed analysis was made of the factorization task, and a production system was constructed that was powerful enough to perform the task successfully. This production system is assumed to represent (approximately) the set of skills that students must acquire in order to do the task. It served as a guide to the design of the teaching materials. It consists of seven productions:

1. If the goal is to factor an expression like $(X^2 + aX + b)$
 ---->
 X^2 is factored as $X.X$.
2. If the goal is to factor a quadratic expression
 ---->
 find all pairs of positive integers whose product equals the absolute value of the constant term of the expression.
3. If the constant term and the coefficient of the linear term are both positive
 ---->
 select a pair of factors whose sum equals the absolute value of the linear term's coefficient.
4. If the constant term is positive and the coefficient of the linear term negative
 ---->
 select a pair of factors whose sum is equal to the absolute value of the linear term's coefficient, and take their negatives.
5. If the constant term is negative and the coefficient of the linear term positive
 ---->
 select a pair of factors whose difference is equal to the linear term's coefficient, and set the smaller factor negative.
6. If both the constant term and the coefficient of the linear term are negative

linear ---->
 select a pair of factors whose difference is equal to the
 term's coefficient, and set the larger factor negative.

7. If {c,d} is the pair of factors that has been selected

 ---->
 write $X^2 + aX + b = (X + c)(X + d)$.

1. First Pretest

FILL IN THE BLANKS (The correct answers are shown in angle brackets, < >. Multiple choice alternatives shown to the students are given in parentheses, ().)

- 1.If: The goal is to factor a 3-termed quadratic polynomial,
 $x^2 + 5x + 6$, with the coefficient of x^2 equal to 1,
- Then: x^2 factors as <X.X>
The constant term is the <product> (product, sum, quotient)
of the <two> factors.
And the coefficient of the linear term is the <sum> (sum,
product, quotient) of these <two> factors.
- 2.If: The constant term is positive,
- Then: The two factors of the constant term have <the same>
(different, the same, positive, negative) signs.
- 3.If: The constant term is positive, and the linear term
is also positive,
- Then: The two factors of the constant term are <positive>
(different, the same, positive, negative).
- 4.If: The constant term is positive, the linear term is negative,
- Then: The two factors of the constant term are <negative>
(different, the same, positive, negative).
- 5.If: The constant term is negative, the linear term is positive
(e.g., $x^2 + 4x - 12$),
- Then: The two factors of the constant term are <opposite> in sign,
and the positive factor is <larger> in size than the
negative one.
- 6.If: The constant term is negative, the linear term is negative
(e.g., $x^2 - 4x - 12$),
- Then: The two factors of the constant term are <opposite> in sign,
and the negative factor is <larger> in size than the
positive one.

2. Review Knowledge

Subjects in both conditions studied the following materials:

FILL IN THE BLANKS

Example: 8 can be expressed as the product of the two numbers, 2×4 , and the sum of these two numbers is 6.

Problems: 8 can also be expressed as the product of two numbers, 1×8 , and the sum of these two numbers is .

6 can be expressed as the product of two numbers, , and the sum of these two numbers is .

6 can be expressed as the product of two numbers, , and the sum of these two numbers is .

Example: In $x^2 + 5x + 6$, x^2 is the quadratic term, $5x$ is the linear term, 6 is the constant term.

Problems: In $x^2 + 8x + 12$, x^2 is , $5x$ is , 12 is .

In $x^2 + 13x + 12$, 12 is , x^2 is , 13x is .

USING THE RULES OF REPEATED MULTIPLICATION, WORK THE PROBLEMS BELOW:

Example: $(x+2)(x+4) = x^2 + 4x + 2x + 8 = x^2 + 6x + 8$.

Problem: $(x+3)(x+5) = x^2 + + + 15 = x^2 + + 15$.

INVERSE PROCESS: $x^2 + 6x + 8 = (x + 2)(x + 4)$

$x^2 + 6x + 8$ can be expressed as the product of $(x+2)$ and $(x+4)$.
The components are called the factors of the quadratic.

3. Second Pretest

FACTOR THE FOLLOWING:

(1) $x^2 + 7x - 18 =$

(2) $x^2 - 2x - 8 =$

(3) $x^2 + x - 20 =$

(4) $x^2 - 5x - 36 =$

(5) $x^2 - 9x + 14 =$

4A. Learning from Examples**Part I**

Examples: (1) $x^2 + 5x + 6 = (x + 2)(x + 3)$

(2) $x^2 + 7x + 6 = (x + 1)(x + 6)$

(3) $x^2 + 8x + 12 = (x + 2)(x + 6)$

(4) $x^2 + 7x + 12 = (x + 3)(x + 4)$

(5) $x^2 + 13x + 12 = (x + 1)(x + 12)$

Exercises: (1) $x^2 + 11x + 18 = (\quad) (\quad)$

(2) $x^2 + 9x + 18 = (\quad) (\quad)$

(3) $x^2 + 19x + 18 = (\quad) (\quad)$

Part II

Examples: (1) $x^2 + 5x + 6 = (x + 2)(x + 3)$

(2) $x^2 - 5x + 6 = (x - 2)(x - 3)$

(3) $x^2 + 7x + 6 = (x + 1)(x + 6)$

(4) $x^2 - 7x + 6 = (x - 1)(x - 6)$

Exercises: (1) $x^2 + 9x + 18 = (\quad) (\quad)$

(2) $x^2 - 9x + 18 = (\quad) (\quad)$

(3) $x^2 - 11x + 18 = (\quad) (\quad)$

(4) $x^2 + 11x + 15 = (\quad) (\quad)$

Part III

Examples: (1) $x^2 + 5x - 6 = (x - 1)(x + 6)$

(2) $x^2 - 5x - 6 = (x + 1)(x - 6)$

(3) $x^2 + x - 6 = (x - 2)(x + 3)$

(4) $x^2 - x - 6 = (x + 2)(x - 3)$

- Exercises: (1) $x^2 + 4x - 12 = (\quad) (\quad)$
 (2) $x^2 - 4x - 12 = (\quad) (\quad)$
 (3) $x^2 + x - 12 = (\quad) (\quad)$
 (4) $x^2 - x - 12 = (\quad) (\quad)$
 (5) $x^2 + 11x - 12 = (\quad) (\quad)$
 (6) $x^2 - 11x - 12 = (\quad) (\quad)$
 (7) $x^2 + 7x + 10 = (\quad) (\quad)$
 (8) $x^2 - 6x + 8 = (\quad) (\quad)$
 (9) $x^2 + 7x - 18 = (\quad) (\quad)$
 (10) $x^2 - 2x - 8 = (\quad) (\quad)$
 (11) $x^2 + x - 20 = (\quad) (\quad)$
 (12) $x^2 - 5x - 36 = (\quad) (\quad)$
 (13) $x^2 - 9x + 14 = (\quad) (\quad)$

4B. Learning by Doing

Part I

Problems: [Identical with Examples for Learning-from-Examples]
 except that right-hand sides of equations are shown as:
 $= (x + \underline{\quad})(x + \underline{\quad})$]

Exercises: [Identical with Exercises for Learning-from-Examples]

Part II

Problems: [As before, except that right-hand sides are shown as:

- (1) $= (x - \underline{\quad} 2)(x - \underline{\quad} 3)$, (2) $= (x - \underline{\quad} 2)(x - \underline{\quad} 3)$,
 (3) $= (x - \underline{\quad} 1)(x - \underline{\quad} 6)$, (4) $= (x - \underline{\quad} 1)(x - \underline{\quad} 6)$.]

Exercises: [As before.]

Part III

Problems: [As before, except that right-hand sides are shown as:

- (1) $= (x - \underline{\quad} 1)(x - \underline{\quad} 6)$, (2) $= (x - \underline{\quad} 1)(x - \underline{\quad} 6)$,
 (3) $= (x - \underline{\quad} 2)(x - \underline{\quad} 3)$, (4) $= (x - \underline{\quad} 2)(x - \underline{\quad} 3)$.]

Exercises: [As before.]

References

- Anderson, J.R. (1987), Production systems, learning, and tutoring; in Klahr, D., Langley, P. & Neches, R., *Production System Models of Learning and Development*, Cambridge, MA: MIT Press.
- Anzai, Y. (1978) Learning strategies by computer. *Proceedings of the Second National Conference of the Canadian Society for Computational Studies of Intelligence*
- Anzai, Y. and Simon, H. A. (1979) The theory of learning by doing, *Psychological Review* 86:124-140.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A Study of Thinking*, New York: Wiley, 1956.
- Gagné, R. M. (1965). *The Conditions of Learning*, New York: Holt, Rinehart & Winston.
- Gagné, R. M. & Briggs, L. J. (1974). *Principles of Instructional Design*, New York: Holt, Rinehart & Winston.
- Hovland, C. I. & Hunt, E. B. (1960). The computer simulation of concept attainment. *Behavioral Science*, 5: 265-267.
- Hunt, E. B. (1962). *Concept Learning*, New York: Wiley.
- Katona, G. (1940) *Organizing and Memorizing* New York: Columbia University Press.
- Lu, Zhongheng (ed.) (1983, 1984, 1986). *Zhongxue Shuxue Zixue Fudao Jiaoxue Shiyan Wenxuan* (Selected Works on Educational Experiments With Guided Self-Study of Middle-School Mathematics), Vols. 1, 2, 3. Beijing, PRC. (Reprints of papers (in Chinese) from Chinese psychological journals, reporting the results of the experiments.)
- Neves, D. M. (1978) A computer program that learns algebraic procedures by examining examples and working problems in a textbook, *Proceedings of the Second National Conference of the Canadian Society for Computational Studies of Intelligence* 191-195.
- Neves, D. M., and Anderson, J. R. (1981) Knowledge compilation: mechanisms for the automatization of cognitive skills. In J. R. Anderson (ed.), *Cognitive Skills and Their Acquisition* [Hillsdale, N.J.: Lawrence Erlbaum Associates], 57-84.
- Newell, A., Shaw, J. C., & Simon, H. A. (1962) The processes of creative thinking. Reprinted in H. A. Simon, *Models of Thought*, New Haven: Yale University Press, 1979, Ch. 4.1.

- Resnick, L. B. (1976) Task analysis in instructional design: Some cases from mathematics, in D. Klahr (Ed.), *Cognition and Instruction*, Hillsdale, N. J.; Erlbaum.
- Shulman L. S., & Kieslar, E. R. (eds.) (1966). *Learning by Discovery*, Chicago: Rand-McNally.
- Siklóssy, L. (1968). Natural language learning by computer. Dissertation, Carnegie-Mellon University, Pittsburgh. Revised version reprinted in Simon, H. A. & Siklóssy, L. (eds.), *Representation and Meaning*, Englewood Cliffs, N.J.: Prentice-Hall, 1972, Chapt. 6.
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2. 59-89.
- Waterman, D. A. (1970), Generalization learning techniques for automating the learning of heuristics. *Artificial Intelligence* 1:121-170.
- Waterman, D. A. (1975), Adaptive production systems. *Proceedings of the Fourth International Joint Conference on Artificial Intelligence*, pp. 296-303.