R842-09 ERIC REPORT RESUME

ED 010 184 1-30-67 24 (REV)
A STUDY OF PSYCHOLOGICAL PATTERNS IN LEARNING ELEMENTARY MATHEMATICS.
RESTLE, FRANK
GKF22258 INDIANA UNIV., BLOOMINGTON
CRP-5-339
BR-5-8059
- -66
EDRS PRICE MF-$0.18 HC-$4.04 101P.

*MATHEMATICS, LEARNING PROCESSES, *PSYCHOLOGICAL STUDIES, ELEMENTARY SCHOOL STUDENTS, *SECONDARY SCHOOL STUDENTS, *COLLEGE STUDENTS, *THOUGHT PROCESSES, PROBLEM SOLVING, BLOOMINGTON, INDIANA

THE STUDIES REPORTED WERE ATTEMPTS TO LEARN THE INNER ARRANGEMENT OF DOING MATHEMATICS. MATHEMATICS WAS VIEWED AS A MATTER OF SOLVING PROBLEMS BY A SEQUENCE OF STEPS PERFORMED AS AN ALGORITHM. THE PROCESSES STUDIED WERE--(1) COUNTING OR ENUMERATING SETS, (2) SIMPLIFYING COMPLEX LOGICAL STATEMENTS, (3) ADDING LARGE NUMBERS, AND (4) MULTIPLYING. WITHIN EACH EXPERIMENT, A VARIETY OF TASKS WAS CONSTRUCTED SO THAT DIFFERENT DISCRIMINATIONS, SLIGHTLY DIFFERENT STEPS OF ALGORITHM, OR SLIGHTLY DIFFERENT OVERALL PLANS COULD BE TESTED. STUDIES ON ENUMERATION TESTED 18, 20, AND 24 SCHOOL CHILDREN AND COLLEGE STUDENTS IN THREE EXPERIMENTS. SUBJECTS FOR THE OTHER PROCESSES STUDIED WERE 20 SCHOOL CHILDREN, 12 SCHOOL CHILDREN DIVIDED ACCORDING TO AGE, AND 50 COLLEGE STUDENTS. STUDIES DEMONSTRATED THAT PATTERNS OF PERFORMANCE MAY EMERGE UNDER STUDY OF THE ELEMENTS OF A MATHEMATICAL TASK. (JH)

ED010184

FINAL TECHNICAL REPORT

Project Title: A STUDY OF PSYCHOLOGICAL PATTERNS IN LEARNING ELEMENTARY MATHEMATICS

Small Contract Project No. 5-8059-2-12-1 (formerly S-339)

Submitted by: Frank Restle, Ph.D.
Professor of Psychology
Indiana University
Bloomington, Indiana

1966

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U.S. Department of Health, Education, and Welfare.

Problem

The glamorous part of mathematics is solving original problems, but the meat and potatos are the ordinary processes of calculating answers and simplifying complicated expressions. The two kinds of activities are complementary, at least for the mathematician, whose ideas give life to his calculations, but whose calculations are the bone and muscle of his results. For many children who will not be mathematicians, the ability to calculate accurately, rapidly, and to calculate the relevant answer, is of great value. It is true that very large and repetitive calculations are now turned over to electronic computers, but, (a) many practical calculations are too small to do on a computer, for the answer can be obtained before the problem is even programmed for the computer, and (b) programming a computer to calculate is itself a process rather like calculating.

Calculation has been little studied by psychologists, who have preferred to study problem-solving and thinking. The remarkable studies of Jean Piaget, Margaret Donaldson, and many others in "children's thinking" often have dealt with problems that involve very little calculation. If we wish to study the pure process of thinking, then the success of the process should not be obscured by mere slips in calculation. However, this obvious point has led psychologists away from any sort of calculation.

A near exception is the studies of processes of solving problems in intelligence tests, as performed by Herbert Simon and his colleagues. This group have written computer programs, with "heuristic" assumptions built in, to perform tasks such as continuing partial series of numbers, etc. Many of the problems are model calculations, but differ from ordinary mathematics in that (a) the necessary information for a complete solution is not given, (b) the task is not taught as such in school, and (c) there is no particular use for the skill tested. These three properties seem to be characteristic of problems that are most useful in intelligence tests.

The present research is based on the assumption that ordinary mathematics is both useful and interesting. In a sense, it is trivial to add a column of numbers. That is, no new results are likely to be obtained from a mathematical standpoint. However, the term "trivial" may obscure the true value of the operations. First, the answer to the sum of a column of numbers may be interesting. If the numbers are the amounts of checks written last month, then the sum is the total amount money spent through checks, and may be a most interesting practical piece of information. If the numbers are the numbers of errors made at various stages of practice in a learning experiment, then the sums are the values of the learning curve, and have scientific interest. It might be argued that adding a column of numbers results in a useful fact more often than any other known mathematical operation.

Second, adding a column of numbers is an example of a process of reducing a complex expression to a canonical, "simplified" form, and that process is both a subject of interesting investigation and is a tool in almost all other mathematical processes. The calculation,

$$
\begin{array}{r} 1347 \\ +\underline{3857} \\ 5204 \end{array}
$$

can be thought of as reducing a complex expression,

$$
\begin{array}{r} 1347 \\ +\underline{3857} \end{array}
$$

to its canonical or simplest form, 5204.

It happens that for every set of numbers to be added, the sum of every set of integers or rational numbers, there is a unique canonical form consisting of a single number. The existence of this unique form is of great theoretical significance, in that it establishes that such calculations can be carried out by algorithms (well defined sequences of steps, like a computer program).

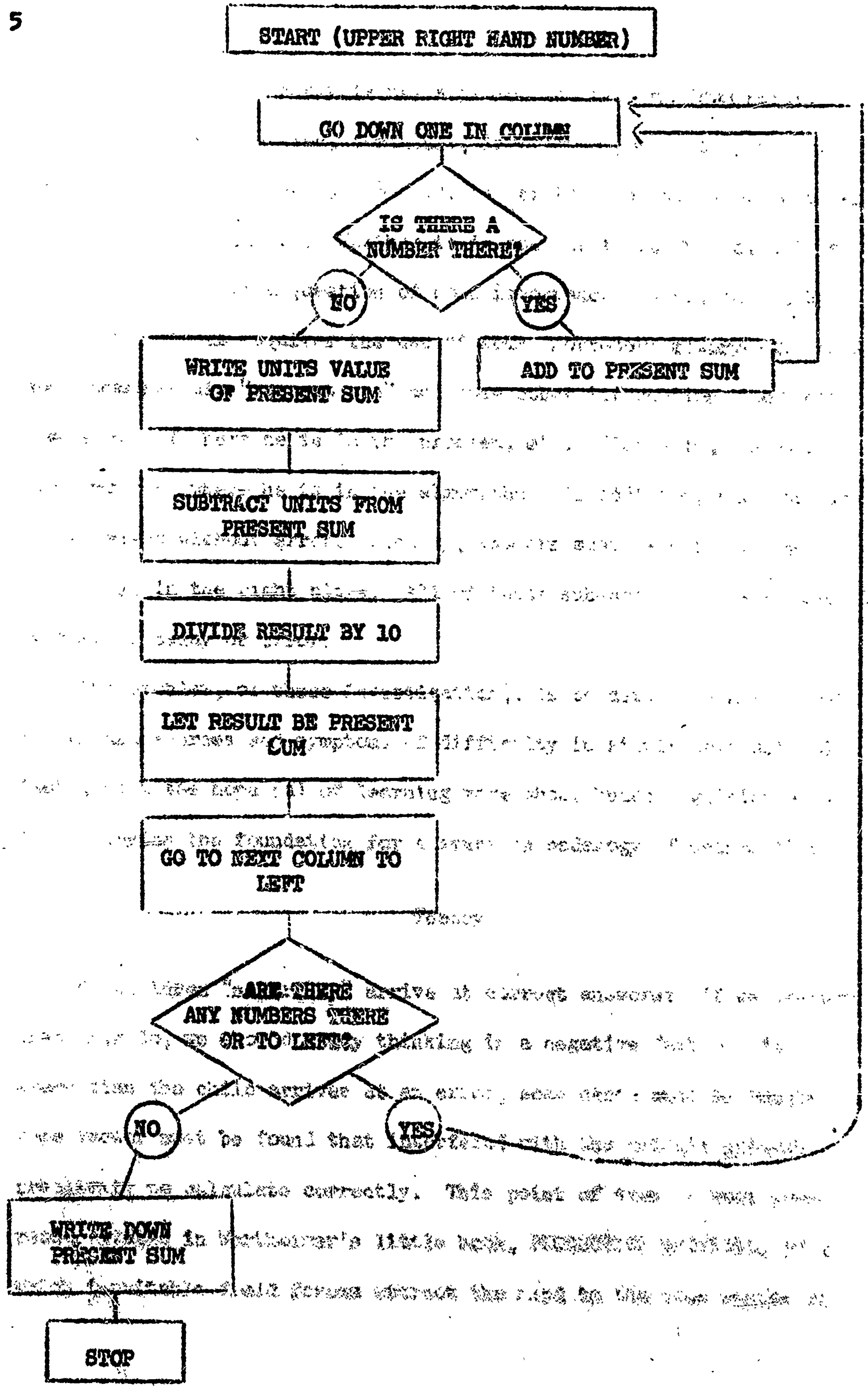
Polya has said that problems in mathematics take one of two forms; TO PROVE, or TO FIND. The theorems about canonical forms, prove that there is a unique sum to every set of numbers, etc., may be relatively complicated and advanced. The probess of finding the solution is, also, complicated though in the case of elementary arithmetic it consists of an algorithm. The existence of the algorithm and proofs of the uniqueness of the solution assure us that there will be no mathematical surprises when we add a column of numbers—a fact upon which banking, among other human activities, is based. However, we may not therefore dismiss summing as a mathematical activity.

Many mathematical processes are trivial in that the outcome is a "forgone conclusion", but this does not mean that the outcome is known before the process is carried out.

Although we have some understanding, in a general way, of the mathematical nature of adding a column of numbers, it is remarkable that we have so little understanding of the psychology of this process. Like other calculations, adding a column of numbers may be thought of as an algorithm, and this brings to mind two questions. First, what are the elementary steps in the process? Second, how are these steps combined into the whole serial process?

In adding numbers, we should not be too sure that we know what the elementary stages are. Following an ordinary method as the writer himself performs the task, consider adding by the usual method as if it were a computer program. The program is sketched below, though of course the information given below is insufficient to find where to write the answers, how to carry, etc.

Notice that each column of the numbers is made up of a chain or loop of sub-processes, including adding new numbers to a present sum and then checking to see if one is at the end of the column. Psychologically, adding 7 to a present-sum of 25 may be relatively complicated, and is not easily performed by a young child. At each change of column, a new check and the carrying operation must be performed. Thus, there are a number of "instructions" or steps, and a number of decisions to be made, within this algorithm. Furthermore, the algorithm, since it contains loops, represents a serial process in which the elementary processes are strung together within a framework. All in all, adding a column of numbers by the standard method is relatively complicated.


START (UPPER RIGHT HAND NUMBER)
GO DOWN ONE IN COLUMN
IS THERE A NUMBER THERE?
NO
YES
ADD TO PRESENT SUM
WRITE UNITS VALUE OF PRESENT SUM
SUBTRACT UNITS FROM PRESENT SUM
DIVIDE RESULT BY 10
LET RESULT BE PRESENT SUM
GO TO NEXT COLUMN TO LEFT
ARE THERE ANY NUMBERS THERE OR TO LEFT?
NO
YES
WRITE DOWN PRESENT SUM
STOP

A psychological analysis may approach the problem from several other directions. Adding, certainly, involves a good bit of memory; we merely remember that $5 + 8 = 13$. A fairly large store of permanent information is employed. Whether it is subject to confusion, delays in access, etc., is a question of some importance. Also, adding by the usual methods requires the use of some short-term memory--the child must remember his "present sum" and some other information, must not lose track of where he is in the problem, etc. Similarly, the child must remember where he is in the algorithm. In addition, numbers must be perceived without error. Finally, answers must be written down correctly, in the right place. All of these sub-activities are possible sources of delay or error.

The problem, of these investigations, is to determine more about the actual sources and symptoms of difficulty in simple arithmetical tasks, with the hope (a) of learning more about human cognition and (b) of laying the foundation for a sensible pedagogy of mathematics.

Theory

Do children "naturally" arrive at correct answers? If we believe that they do, we should study thinking in a negative fashion, for every time the child arrives at an error, some cause must be imputed, some factor must be found that interfered with the child's natural proclivity to calculate correctly. This point of view is most accurately stated in Wertheimer's little book, PRODUCTIVE THINKING, in which inevitable field forces attract the mind to the true source of the difficulty, then swirl about and combine to produce the solution. Only rigidly, ordinarily induced by bad teaching, can prevent this natural process from proceeding to its charming conclusion.

In opposition to this romantic position stands Piaget, with his assertion that the young child is by nature rigid and unable to compute, and only develops flexibility, the ability to be carried by the structure of the situation to the proper answer, by a series of new developments leading to groupments and operational thinking.

The present studies differ from Piaget's usual orientation in that attention centers on the particular tasks of arithmetic, rather than on the general abilities and levels of development studied by Piaget. It is natural, with the interests of the Geneva laboratory, that the tasks should be chosen not for intrinsic interest but for the information they purportedly throw on the thought processes of the child. Therefore, problems in "conservation" of volumes, in number of flowers in subsets compared with the whole set, etc., are characteristic topics of study. The child is asked what moves the trees on a breezy day, or is required to solve simple problems with a balance.

Such tasks are chosen, usually so as to minimize the amount of calculation required, because the investigator is interested in general levels of development, not in the mere mechanics of manipulating symbols. As a result, such investigations tell us almost nothing about how children calculate.

A theory derived from work like Piaget's will attribute solution of a given problem to the general ability, developed by the child, to perform operations with the required level of abstractness and the required flexibility of interrelationship. The analysis may include the idea that specific relationships, not only "levels of ability" must be present, but this merely leads us back into the educational and psychological history of the child.

The present theory has three parts. First, the assumptions and elementary consequences of "discrimination-learning" theory will be stated. Second, the general idea of algorithms and their application to elementary mathematics will be outlined. Finally, an integration of the two sets of ideas will be attempted, in the way of asking how the child will guide his algorithmic behavior by discriminations.

Discrimination-Learning Theory

Discrimination-Learning Theory is a set of experimental findings, connected by common underlying assumptions, that can serve as the basis for predicting behavior in more complex situations. In this, it is like classical-conditioning theory and Skinnerian operant-conditioning theory.

The conditioning theories center on the idea that learning results in the formation of new connections between stimuli and responses. If a CS is followed by a UCS, a conditioned response CR will appear, attached to the CS. If S is followed by R and th by a reinforcement, in Skinnerian theory, then the connection of R to S is strengthened.

Discrimination-Learning Theory is concerned, not with the formation but with the selection of appropriate responses to stimuli.

In its simplest form, the theory deals with a sequence of trials of training on each of which a stimulus situation is shown and a "correct answer" indicated. In animal training, the animal is required to make a choice response each trial, then is either rewarded or punished depending on whether he is correct or not. The correctness of a response is indicated only by the reward or punishment. This is trial-and-error learning.

The subject, faced with successive trials of such a problem, establishes a pattern of responding that may be called a "strategy". By a strategy is meant a pattern of responding that provides the subject with a response to each of the situations presented in the problem. That is, Discrimination Learning Theory does not assume that the subject acquires each part of his behavior pattern independently--on the contrary, it assumes that all segments of the behavior system are organized from the beginning.

If the subject meets with success, that is, if he attains a proportion of rewards over punishment that is high relative to his standard or adaptation-level, then he maintains the same strategy. If he fails, that is, if his level of rewards falls below the adaptation level, the subject changes strategy.

The subject is conceived to have available a fixed set of strategies S . Some of these strategies, (not necessarily a unique one), will serve to solve the problem and provide the subject with satisfaction and success. This is the set of "correct" strategies, symbolized by C . Other strategies are irrelevant to the problem at hand, and if followed lead the subject to unsatisfactory performance and "failure". It is understood that such irrelevant strategies, the set I , may lead to temporary or partial success.

If S is constituted only of the sets C and I , so that $C \cup I = S$, then we have a relatively simple experimental problem. The subject begins the problem by choosing some one strategy at random from S . If the strategy so chosen is in C , the subject begins at once performing correctly. If the strategy is in I, then eventually and before long the subject experiences failure, and then resamples from the set S , with replacement. (That is, the set S is not changed.) If this is also an irrelevant strategy, an element of the set I , then later the subject will again resample. He resamples over and over until he happens to hit upon a correct strategy, in the set C . Since that strategy leads to success, the subject does not change it, and his performance will stabilize at success.

In this simple situation there are several properties of the behavior that are of theoretical and practical interest.

1. Performance or "learning" is all-or-none.

Proof: If the subject has an irrelevant strategy he must resample, and that puts him back in the situation of sampling at which he started. Thus, until he hits a strategy in set C , the subject makes no partial progress.

2. The probability of suffering exactly n failures forms a geometric distribution and depends upon the proportion of correct strategies, c .

Proof: The probability of zero errors is the probability that the first sample is correct, which we call c . Suppose the subject instead first chooses an irrelevant strategy, with probability 1-c . Then he certainly suffers a failure. If he then samples and chooses a correct strategy, (probability c), the first failure will be the only one. Therefore, the probability of exactly one failure is (1-c)c . To suffer exactly n failures, the subject must choose irrelevant strategies the first n samples, with probability $(1-c)^n$, then choose a correct strategy so as to end the failures. Thus, in general, the probability that T , total failures, equals n , is given by

$$P(T=n) = (1-c)^n c$$

a geometric distribution. (It is called a geometric distribution because the sequence of numbers, c , (1-c)c , $(1-c)^2c$, etc., form a geometric progression). Furthermore, the parameter c of that distribution is the probability of sampling a correct strategy, given the subject resamples.

Corollaries: Elementary probability theory tells us that the mean total failures, given the above model, and the variance of total failures, are

$$[1] \quad E(T) = (1-c)/c$$

$$[2] \quad Var(T) = (1-c)/c^2$$

whence it can be seen that the standard deviation is

$$[3] \quad \sigma = \sqrt{1-c}/c$$

and is slightly larger than the mean. The data are highly variable. Furthermore, the geometric distribution is J-shaped, with its mode at zero failures.

3. The speed of learning depends upon the probability of choosing a correct strategy.

Proof: From point 2 above, especially Equation 1. Notice that as c increases the fraction $(1-c)/c$ decreases—hence as the probability of choosing a correct strategy, c, becomes larger, the expected total failures before solution gets smaller.

Remark: Mean total failures, $E(T)$, is a useful index of "difficulty" of the problem, and we shall think of "rapid learning" as "reduced mean errors."

4. Let any set A of strategies, a subset of S, have a measure $m(A)$. We construe $m(A)$ as the tendency for the subject, when he samples from a S containing A, to choose his sample from A.

Since $S = C \cup I$, and C has no elements in common with I by definition, it follows from elementary measure theory that

$$[4] \qquad m(S) = m(C) + m(I)$$

Now c, the probability of choosing a correct strategy, is defined as

$$[5] \qquad c = \frac{m(C)}{m(C) + m(I)}$$

5. Other things equal, the speed of learning can be increased by adding alternative correct strategies to the problem.

Proof: Suppose that the original set of correct strategies was called C_o, and the added alternative correct strategies are set C_a. The language of the statement implies that there is no overlap between C_o and C_a (that is, we shall think of the added strategies as only the new ones added.) In the new problem the set of correct strategies is

$$C_n = C_o \cup C_a$$

and since the two are disjoint,

$$[6] \quad m(C_n) = m(C_o) + m(C_a).$$

There is no reason to believe that adding new relevant strategies should change the set of irrelevant strategies, so we assume that the set I is unchanged; let I_o be the new set of irrelevant strategies;

$$I_n = I_o$$
$$= I.$$

The total sets of strategies are also different, of course;

$$S_o = C_o \cup I$$
$$S_n = C_n \cup I$$

Putting this together, we find that the proportions of correct strategies, old and new, are

$$c_o = \frac{m(C_o)}{m(C_o) + m(I)}$$

$$c_n = \frac{m(C_o) + m(C_a)}{m(C_a) + m(C_o) + m(I)}$$

and, provided that $m(C_a)$ is greater tahn zero, the result is that c_n is greater than c_o. As mentioned under Point 2, this means fewer failures, and as stated under Point 3, this in turn is what we mean by faster learning.

Discussion: An example from industry is the use of color coding. A mechanic must learn to discriminate between the transmission and fuel systems, but the two may be intermixed and similar in general appearance inside a truck motor. If all parts of the transmission are (in a training vehicle) painted yellow, and all parts of the fuel system are painted red, then these color cues are the basis for a number of new strategies. The result is that the trainee can more rapidly discriminate the two systems. As we see below, he may not be able to work on regular trucks afterwards, however.

The Proof of POint 5 is relatively complete and elaborate. Since the arguments are elementary throughout this section, they will be given in relatively sketchy form. It is to be understood that each particular point should be spelled out in complete detail.

6. Other things equal, the speed of learning can be increased by reducing distractions, if this has the effect of removing some irrelevant strategies from S.

Proof: Let the original set of irrelevant strategies be called I_o and the new set, after removing distraction, be I_n, and assume that $m(I_n) < m(I_o)$. Then, since other things are kept equal, $C_o = C_n = C$. Hence,

$$c_o = \frac{m(C)}{m(C) + m(I_o)}$$

and

$$c_n = \frac{m(C)}{m(C) + m(I_n)}$$

and c_n is greater than c_o.

7. Other things equal, the mean total errors $E(T)$ increases linearly with the number or measure of irrelevant strategies.

Proof: Let I_x be the variable set of irrelevant strategies. Then the proportion of relevant strategies is

$$[7] \qquad c_x = \frac{m(C)}{m(C) + m(I_x)}$$

From Eq. 1,

$$E(T_x) = \frac{1 - c_x}{c_x}$$

and substituting the results of Eq. 7,

$$E(T_x) = \frac{\frac{m(I_x)}{m(C) + m(I_x)}}{\frac{m(C)}{m(C) + m(I_x)}}$$

$$= \frac{m(I_x)}{m(C)}$$

Since m(C) is a constant, this means that the mean errors are proportional to the measure irrelevant strategies.

Discussion: One application of this idea is to the schoolroom; when the children are lively and moving about, the walls are covered with colorful and interesting posters, and the day is full of activities, it must be assumed that the children are full of irrelevant strategies with respect to any particular task they are to work on. The result is happy children with low learning rates.

Point 7. Other things equal, the rate of learning can be increased by emphasizing the relevant cues.

Discussion: By "emphasizing" is meant using some stimulus, such as a pointer, italics in print, color in a picture, etc., for the purpose of directing the learner's attention to relevant cues. This will increase the probability of using a given correct strategy, not by adding new correct strategies, but by changing the sampling probabilities.

Point 8. Other things equal, the rate of learning can be increased by giving the subject a useful record of strategies he has tried. In this way, hopefully, the subject will discard all irrelevant strategies tried, and after each failure he will have a higher probability of hitting a correct strategy.

Discussion: If the learner uses the record to eliminate strategies he is not sampling with replacement from a fixed set S . The notion of learning by systematic elimination thus transcends simple discrimination--learning theory.

Point 9. Transfer of training from one simple discrimination to another occurs all-or-none.

Proof: Call the tasks A and B . Suppose the subject has chosen a particular strategy s to solve problem A . When given the first trial of problem B , he presumably uses the same strategy s if he can. If s is also a successful strategy in problem B , then he transfers perfectly ("all").

Now suppose that strategy s does not apply in problem B at all, so the subject cannot use it. Then he must begin sampling on problem B just as if he had no earlier training ("none").

The only remaining possibility is that the subject can apply strategy s in problem B , but it leads to failure. After that failure the subject presumably must resample, and is again just like a subject with no previous training ("none").

Point 10. If a subject suffers so much as one failure in transfer, then his expected future performance is the same as if he had no previous training.

Proof. This is a simple consequence of the all-or-none transfer idea. A single failure is enough to tell us that the subject is not transferring "all". Hence the transfer must be none, from Point 9.

Point 11. The probability of transfer from task A to task B, given mastery of A, is equal to the conditional probability that a strategy, given it is chosen from C_A, is an element of $C_A \cap C_B$.

Discussion. This again is simply a restatement from the proof of Point 9. However, it has a useful mathematical point, as follows: Consider that the subject had no training that transferred to Task A; Task A is the starting point. Then the probability of transfer to Task B is given by

$$[9] \quad P(\text{Transfer } A \to B) = \frac{m(C_A \cap C_B)}{m(C_A)}$$

Application of Discrimination-Learning Theory to the Performance of Mathematics

Clearly, adding a column of numbers or counting the pennies on a desk are activities quite different from trial-and-error learning. However, mathematical tasks may require discriminations, in which case the difficulties observed may be related to some of the variables in Discrimination-Learning theory.

For example, consider a child counting pennies. He must, first, discriminate the pennies from whatever irrelevant stimuli may be about. This, in turn, will be easy if there are no other confusing objects about. If no other coins are present, then mere roundness can be used as the basis for discrimination, though if there are dimes in the pile additional cues of color and size must be employed. This shows how, by simplifying the situation, one can make more strategies relevant. Similarly, if all the coins to be counted are physically close together, then location can be the basis of a discrimination of the set to be counted. During the process of counting, the child must always be able to discriminate between the set of coins already counted, and those still left to be counted. This discrimination may be greatly aided by moving the coins from one pile to another, adding new strategies based on location.

Discrimination learning theory would also hold that this, and any other discriminations needed, would be easier if distractions are removed. In mathematics this may mean both improving general working conditions, and also study of the design of questions, many of which include irrelevant or unnecessary information, or suggest many inappropriate strategies.

The question of transfer-of-training is the most delicate. The all-or-none transfer theory arising from discrimination-learning theory has a direct application that can be illustrated by a possible pattern of failure of transfer.

A child becomes quite skillful at answering the following sort of "Word Problems". "Nancy is 12. John is 7. What is the difference in age between Nancy and John?" Our child is taught to do this by listening for the word "difference" and being told to subtract.

Another problem is given. "Nancy is 3 years older than Mary, and Mary is 2 years older than John. What is the difference in age between Nancy and John?" Using the strategy that solved the earlier problem, the child subtracts; 3 - 2 = 1 , and he offers the answer that Nancy is one year older than John.

In this situation the child has found a strategy that solved task A (any task in which the difference is found by subtracting) but now fails on the new problem, in which differences are given and must be added. In this case, the old strategy applied but leads to failure.

This possibility illustrates an important characteristic of discrimination-learning theory and its approach to learning. After a given problem is mastered, according to discrimination-learning theory, we still do not necessarily know what strategy the subject is using. All we know is that the strategy he is using, whatever it is, is good enough to solve this particular problem. Therefore, it may come as a surprise to the teacher that the child, having answered a whole sequence of "difference" problems correctly suddenly gives a "stupid" answer on a new problem. What this means is that the earlier behavior was mediated by a strategy not intended by the teacher.

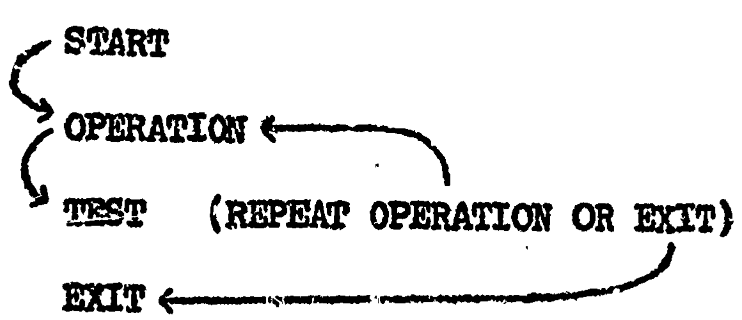
Algorithms

An algorithm is a recipe for mathematical or other activities, including methods of making all necessary decisions. If it is followed exactly it always terminates, and always yields the correct answer. Any set of problems that can be solved by algorithms are "trivial" in that no intelligence or mathematical insight, and no luck, are needed to solve them.

The theory and methods of algorithms are now considered, not trivial, but quite the contrary, for any problem that will yield to an algorithm is "computable" and can be handled by a digital computer.

The general theory of algorithms is extremely precise and formal, and not appropriate for this context. However, some obvious properties of algorithms may be mentioned to suggest how a psychological analysis may be performed.

An algorithm is a variation or elaboration of a system organized as follows:

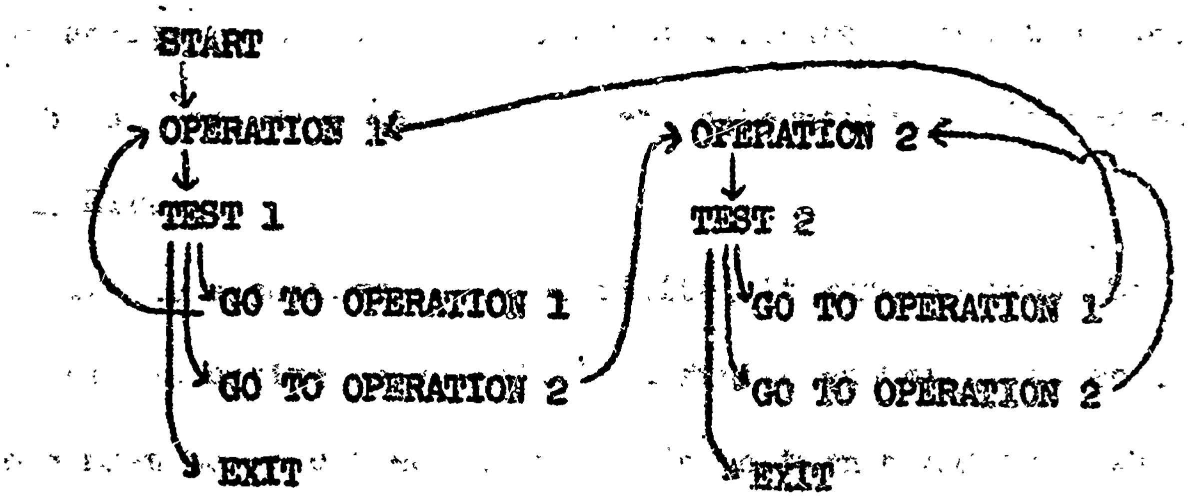


The pair of parts labelled OPERATION and TEST form a loop, in that the operation may be repeated over and over. A completely closed loop would be repeated infinitely often, but the TEST step provides a basis for stopping the algorithm.

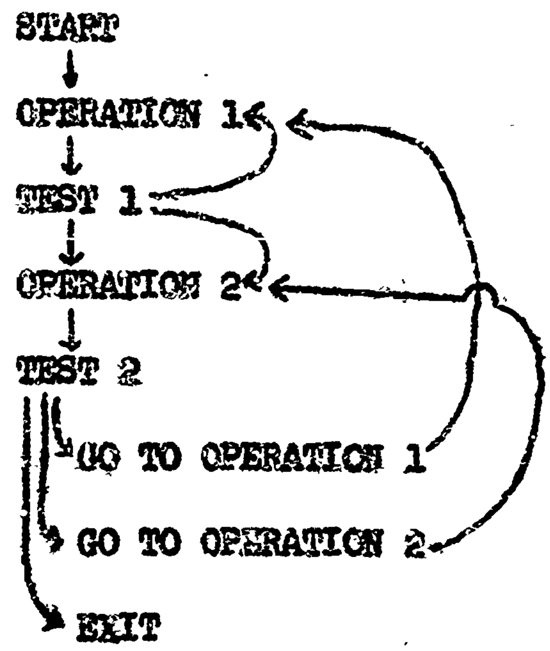
In psychological problems, a certain situation is entered at START. The step OPERATION modifies the situation, and the step TEST corresponds to a comparison of the new situation with some criterion. Depending on the relation of situation to this criterion, the algorithm may return to OPERATION and again change the situation, or may EXIT from the algorithm.

Notice that the above is similar to the concept of a TOTE unit (Test-operate-test-exit) put forward by Miller, Galanter, and Pribram in their book, PLANS AND THE STRUCTURE OF BEHAVIOR.

It is possible to develop more complex algorithms by using more than one operation, and then increasing the complexity of the TEST to permit any of several decisions. A slight complication is the following:



in which two elementary SOTE units are hooked together. In the pattern above the two units are hooked up in parallel. They may also be connected in series, as follows:



In this second arrangement, OPERATION 1 — TEST 1 are an optional preliminary to OPERATION 2 — TEST 2.

Basically, one must always proceed from START to an operation, and then after an operation that may be relatively complicated, one must emerge at a TEST. The TEST is a branching point at which there may be several alternatives. Every alternative is either an OPERATION or an EXIT.

If the algorithm is to be finite, it must be true that the system will always arrive at an EXIT in a finite number of stages. Therefore, there must be no closed loops. This in turn means that wherever the system goes it must always encounter tests that may turn it to EXIT. Furthermore, if the system does not exit, it must go to an OPERATION or series of operations that change the situation, and the change must be in the direction of whatever situation results in an EXIT.

This last requirement is very complicated to discuss and evaluate. Obviously, it is a general description of a computer program that does not hang up or go into an endless loop, and anyone who has programmed a computer knows that the logic of a good program may be extremely complex. In fact, it is often much easier to run a program and find out if it works, than to analyze it and determine its logical viability.

In psychological models of algorithms, a system may only be <u>stochastically</u> finite. By this prepossessing statement I mean only this; the system may be in a loop, returning to operations and not exiting, any number of times. However, the changes in the situation, or else the outcomes of the tests, are probabilistic. There is, on every sweep through the loop, a non-zero probability of exiting, and the probabilities are such that AS N INCREASES, THE PROBABILITY OF REMAINING IN THE LOOP FOR AT LEAST N TIME-UNITS APPROACHES ZERO. Thus, we cannot set a maximum number of trials for the algorithmic process, and say that for some M, the system exits by trial M. All we can say is that for any probability P however small, there exists a number of trials depending upon P, N(P), such that the probability of the system looping as many as N(P) trials is less than P.

A computer program with this characteristic is diagrammed below:

```
START
  |
  v
1    CALL A RANDOM DIGIT  N = 1,2,3,4,5,6,7,8,9, or 10 .
     TEST IF N = 1
         If YES, EXIT
         If NO, GO TO STATEMENT 1
```

Suppose that on each try, the probability that $N = 1$ is equal to p. Then with probability $1 - p$, the system loops. The probability that it loops N times is $(1-p)^N$. Thus if one wishes to be sure, to a probability $P = .01$, that the system will be finished by N trials, we merely choose an N so large that

$$(1-p)^N < .01 .$$

There always is such an N; any N larger than $\log .01 / \log (1-p)$ will do, since the criterion was equivalent to

$$N \log (1-p) < \log .01 .$$

(Note, please, that both logarithms are negative). The above inequality transforms to

$$N > \log (.01) / \log (1-p)$$

when both sides are multiplied by the negative number, $1/\log(1-p)$.

Teh characteristic of a stochastically-finite system is that the OPERATIONS need not progress in a particular "direction" toward the result that produces EXIT. Instead, the OPERATIONS may merely sample about, producing no net improvement, but with some finite probability of arriving at a satisfying situation in one, or a finite series, of steps.

Clearly, under these circumstances, one need only think of the OPERATION-TEST pairs as states of a Markov Chain. One of the states, EXIT, is an absorbing state, and there are no other ergodic subsets. Then if the systm has a finite number of states, it follows that the algorithm is stochastically finite.

The above remarks are somewhat informal, but serve to sketch in the connection between algorithms, or repeated steps of calculation, and the all-or-none learning process of discrimination learning theory. Each OPERATION-TEST pair can be thought of as an instance of discrimination learning, and on each pass through that State the subject has a possibility of solving the problem, in this case, going to an EXIT situation. If learning is all-or-none, then each of these sub-systems can be thought of as part of a Markov Chain, for the all-or-none property is identical, in its theoretical significance, with the basic assumptions of a simple Markov Chain.

Consider a child adding the column of numbers,

$$\begin{array}{r} 5 \\ 3 \\ \underline{+7} \end{array}$$

An algorithm for this process is as follows:

START: Find the top number N and read it into storage X .

OPERATION 1: Look below the number below X . Call it Y .

TEST: Is there a number there?

If Yes, Go to Operation 2.

If No, Go to Operation 3.

OPERATION 2: Add Y to X . Call this result X . Go to Operation 1.

OPERATION 3: Write down X as the answer. Go to EXIT.

Several comments about this example are in order. First, it is formally a proper algorithm provided the operations are defined, and provided there is a point in the problem at which the TEST will come out "No", that is, the child will arrive at the end of the column of numbers.

It might be objected that OPERATION 2 includes the concept of "adding two numbers", and it is circular to define a algorithm for adding numbers that includes, within it, the concept of addition. However, it may be assumed, at least as an hypothesis, that the psychological act of adding two numbers together is less than, and different from, the act of summing a whole column of more than two numbers.

A more complete analysis of the problem will analyze the OPERATIONS given above into parts. It may be that within a single operation one can find a sequence of steps and tests. The writer sometimes, when tired, is rather bad at arithmetic. Suppose he must add 4 to 37 . He may put out four fingers of one hand, and say to himself "37". Then he counts from "37", at each count folding in one of the extended fingers, and at each step checking that there is at least one finger out. When there are no fingers, the sum is simply the last number said. Thus, "ADD 37 + 4" becomes the following algorithm:

START

OPERATION 1: Put out number of fingers corresponding to smaller number.

OPERATION 2: Say the larger number, N .

OPERATION 3: Increment the larger number N to N + 1.

OPERATION 4: Fold in one finger.

TEST: ARE THERE ANY FINGERS STILL OUT?

If YES, go to OPERATION 3.

If NO, go to OPERATION 5.

OPERATION 5: Write down N . EXIT .

To understand arithmetical calculation, it is not sufficient to understand a child adding two numbers. By further analysis, that process may be broken down into its components, though we may find that the components, such as counting on the fingers, are no longer used, and the result is obtained "directly". However, we must also synthesize together the individual acts of adding into a total process, by which the young adult becomes able to add a whole column of 5-place digits by hand, if necessary. The purpose here in introducing the notion of algorithms is to permit the synthesis of the psychological process by which arithmetic calculations are carried out.

<u>Discriminative Control of Algorithms</u>

Notice that the power of algorithms is that they permit the reuse, over, and over, of the same OPERATION, and thus effect a great economy, in that very complex tasks, with hundreds of steps, can be accomplished after the mastery of only two or three OPERATIONS.

However, if we would return again and again to an OPERATION, we must have some way of terminating the process, otherwise each calculation would merely throw the performer into an endless whirl of aimless activity.

Furthermore, most complex calculations are not the repeat of exactly the same operation over and over, but instead require variations in performance depending upon the current state of the calculation.

Both in order to stop the process, and to give it the necessary flexibility, each step is controlled by a TEST.

Speaking in psychological terms, a TEST is a discrimination. When the algorithm specifies a test, the subject must inspect the situation (which no doubt has many irrelevant aspects along with those relevant to the algorithm), and select from it the essential stimulus characteristic. Then one of two or more responses must be chosen. It happens that the responses are OPERATIONS rather than simple choices, but it is still possible to consider the choice process itself by use of Discrimination Learning Theory.

In this way it is conceived that performance of an algorithm involves performance of several discrimination problems. When a child learns an algorithm, he must learn each of these discriminations, and his ability to perform the complex task without er , and without false EXITS (quitting), depends upon his precision in discrimination. The proportion and clarity of relevant cues, the prepotency of correct strategies, elimination of distractions, etc., should all have a role in mastery of algorithms, and difficulties may often be attributed to failure of specific TEST procedures and discriminations.

In some of our work, we have noticed a psychological process that conflicts, in some ways, with the above notion of an algorithm as the basis of calculation. We notice that the child does not, usually, merely find the beginning of a problem and then plug along, step by step, to the conclusion. If the problem is at all long, we notice that children become worried and confused, and although they may be on the "right track", we find that problems are often failed, not because a mistake has been made, but because the child thinks he may have made a mistake, or thinks he is not on the right track.

Consider, now, that the child periodically or continually runs an extra TEST, asking if his calculation is coming along in a satisfactory way. What could this mean? It might mean that the child often tests to see how much of the calculation he has performed, to try to predict how soon certain branches in the process will be reached, and to test that he has not gone into an endless "loop". These would be checks on the quality of performance. Experienced computer programmers recognize the use of "test" segments which evaluate the progress of the calculation, and serve to terminate a pathological deviation from the programmer's intention. The child, quite naturally, performs such checks during his calculations.

One way the progress of the algorithm can be tested, is by dividing the task into parts. This requires that the child have a general idea of the whole task, from beginning to end, and establish some landmarks so that he can tell what part of the process is present. In other words, the whole process must have some over-all perceptual organization, must form a GESTALT, and must be divided into parts that can be discriminated. Such organization is strongly reminiscent of Wertheimer's principles of grouping, illustrated in so many elementary Psychology text books by diagrams such as the following:

oo oo oo oo oo

in which the circles group into pairs, through proximity. A long task might be divided up, temporally and on the basis of characteristics of the situations, into segments, and this division might be a considerable help in control and test of the algorithm.

Basically, the child requires (1) the ability to comprehend the problem as a whole—this, at least, requires that the boundaries of the problem, at which it abuts other activities, be clearly discriminated. It is also essential that the child be able to discriminate if his procedures have led him into behavior that can have no bearing on the problem.

(2) The child requires some further discriminations between parts of the algorithm. Division of the visual figure into parts, oo oo oo oo, depends upon discrimination between the parts on the basis of their location. Another basis for grouping is the shape of the parts, as in the grouped figure ooxxooxxooxx. Such a structure is organized by discrimination between o and x.

In adding a column of 5-digit numbers, the child may separate the process into five parts, one for each column. This requires discrimination by position, and grouping together the numbers in the first column from the numbers in the second column, etc. The child's location within the program as a whole is measured from right to left, whereas his location within a part is measured from top to bottom. Using such cues, the child can always know where he is and how rapidly he is progressing, and can compare his present with his part location so as to see that "all is well". Without such a happy arrangement, the algorithm of addition may well be beyond the capacity of a normal child.

Clearly, this theory implies the existence of a hierarchy of discriminations--at the top, discrimination of types of problems, and at the bottom, discriminations of particular numbers, etc. Within the present theoretical structure, however, this hierarchy of discriminations and concepts is given special and particular meaning, and through use of Discrimination Learning Theory, particular quantitative predictions become possible.

Method

The studies reported are attempts to learn the inner arrangement of doing mathematics. In the view adoped, mathematics is (in part) a matter of solving problems. Many problems are solved by a sequence of steps performed as an algorithm. The processes studies were (a) counting or enumerating sets, (b) simplifying complex logical statements, (c) adding large numbers, and (d) multiplying.

In each study we used subjects who were able to do the component tasks quite well, and who had at least an adequate ability at the whole process. Our aim was to study the response time of these tasks, and attempt to determine what consumes the time. In addition, we planned to analyze any errors that might occur, attempting to determine their source and thereby further identify the sources of difficulty in the tasks.

The basic experimental design was very simple. A variety of tasks was constructed, within each experiment, so designed that different discriminations, slightly different steps of algorithm, or slightly different over-all plans would be appropriate. In addition, of course, the particular materials used and the answer were varied from problem to problem in the set. Then children or college students were tested on the whole set. Where possible, adequate warm-up and pretraining were given so that there was little trend over trials, and the order of different types of problems was counterbalanced, in an attempt to minimize systematic effects either learning or fatigue. Since each subject was tested on all conditions within an experiment, differences observed could not be attributed to differences in ability.

A test session would go through a complete set of materials, all of which require the same operation (e.g., enumeration, addition, multiplication, etc.) College students were tested in the fashion usually employed in psychological laboratories, being instructed and pretrained, then tested in one or more sessions of about 20 minutes. Children were treated with more care—a preliminary period was devoted to establishing rapport, and the child's attention was caught before each display. Each decision was made in an attempt to have a well-motivated subject, ready for the particular problem, perform at his best in rapid calculation.

There is a reason to question the use of speed as a measure of calculation performance, since of course speed is less important than accuracy, and sheer success. However, given ample time college students and many children can perform with high accuracy. What they do, often enough, is repeat a calculation over and "check" the answer. Repetitions are an important part of real-life calculations, but it seems to us that repetition is a separate process and should be studied by itself. By imposing a time pressure, we probably caused most subjects to do the problem once only. Since errors were infrequent, we need some other measure of the difficulty of a problem, and the time consumed to calculate a correct answer seems satisfactory. Granted that a subject can adjust his speed, trading quickness for accuracy, but he presumably does not make this adjustment separately on problems of slightly different structure. Therefore, whatever his compromise between speed and care may be, we should expect either more errors or more time, probably both, on difficult than on easy problems.

Experiment I

ABSTRACT

Mathematically, the process of enumeration is fundamental to arithmetic. Psychologically, it is a sensori-motor chain controlled at every stage by a shifting perceptual organization. Enumeration requires a chant (ONE, TWO, THREE, ...), a shifting indicator response (pointing), and a perceptual grouping of objects into those already counted and those still ahead. The arrangement of the objects has, theoretically, an important effect on the speed and accuracy of enumeration. Further analysis shows that the serial chain of behavior, required for counting a fairly large set of objects, must be divided into parts, and the objects grouped into corresponding subsets. Three experiments show the relationship between arrangement of objects and counting.

The Process of Enumeration

Mary Beckwith and Frank Restle

Indiana University [1]

At the foundation of applied arithmetic lies the ability to enumerate sets. A basic procedure in counting is to chant the numerals "ONE, TWO, THREE ...", at each word pointing to one of the objects in the set to be counted. The chain of indicator responses is often generated from a basic response like pointing, by a modification like moving to the right. Then, if I is the indicator response and T is its transformation to get to the next object, the responses are I, $T(I)$, $T[T(I)]$, ..., $T^n(I)$. The indicator responses must touch every object once and stop when all the objects have been counted. This requires perceptual control, a sliding discrimination between the set counted, (C), and the set as yet uncounted, (U). As objects are shifted from C to U it is seen that at any stage n, $C_{n+1} \supset C_n$, whereas $U_{n+1} \subset U_n$. The sequence of sets constitutes a <u>linear array</u>, (Restle, 1959a, 1961), is ordered, and provides a basis for simple measurement of distance. Finally, the end of the process of enumeration must be recognized, usually by the emptiness of U. We shall refer to the above process as "simple enumeration".

Young children may fail to produce the chant correctly, but older subjects presumably have difficulty mainly in the sensori-motor components of counting. Some of the factors in sensori-motor control of counting are seen when the experimenter varies the arrangement of the objects to be counted. In simple enumeration, if the objects lie in a straight line, then the transformation of the pointing response T is merely a move to (say) the right, and the discrimination between C and U can be made relative to the location of the finger. If the objects are in a circle, the transformation may again be a simple shift in the clockwise direction, but the discrimination between C and U becomes more difficult as the subject approaches the origin of his counting. This may lead to errors, and perhaps to long delays if the subject starts over. When the objects to be counted are all alike, the circular arrangement should be very difficult, though the straight line is not affected. If the objects differ among themselves, say in shape, then the circular arrangement should become easier. With a rectangular arrangement, again there is a simple path, and there should be no difficulty with the stopping rule. In simple enumeration, the rectangular arrangement has no great advantages over the straight line, and since it provides some optional paths, may even introduce delays. Finally, the objects may be scrambled - approximately equally spaced but in no particular pattern. This forces the subject to devise a path through the set; he must use a complex sequence of transformations of the indicator responses and, since no single cue distinguishes elements of C from U , the subject must construct and remember cues as he goes along.

Simple enumeration is not a complete psychological description of the process of enumeration, because it takes no account of the fact that assemblies of discrete objects are perceived in groups. Of Wertheimer's six principles (Woodworth and Schlosberg, 1954, p. 409) our experiments permit variation of proximity, similarity, and good continuation. The subjects presumably perceive the field of objects as organized into groups. A natural question is whether these groupings have any function in counting.

Enumerating as many as sixteen objects is a long serial task, and the subject may organize the segments into a convenient number of smaller subtasks. If so, then it is necessary to group objects conveniently to permit rapid and accurate counting; for example, if subjects need to group objects, the rectangular arrangement may be consistently better than the straight line.

History of the Problem

Although counting is basic to applied arithmetic, there is little literature of experimental studies dealing with counting. Many studies have dealt with the immediate apprehension of number, a process supposed not to be mediated by steps, but to occur instantly. Freeman (1912) set the span-of-apprehension of circles of light at five for adults and four for children over eight years; younger children were untestable. Fernberger (1921) used random arrangements of four to twelve dots, and found that the relative frequencies of correct responses made a continuous ogival function of number of dots shown. Correct judgments were made fifty percent of the time, (Fernberger's measure of visual discrimination), in the range from six to eleven dots. Hunter and Sigler (1950) varied intensity and duration of presentation of dots and found that the usual visual laws applied.

One approach to the span-of-apprehension idea is as detection of a number of targets (Schlosberg, 1948; Casperson and Schlosberg, 1950; Restle, Rae, and Kiesler, 1961). This approach deals mainly with perception of a number of dots under short exposures or poor illumination, but may undermine the concept of a "magic number".

Warren (1897) arranged white squares in a vertical line on a black background and measured the time taken to count them and concluded that four was the limit of perceptual counting. Bourdon (1908) also used the method of reaction time, and found a small increase in time as a function of number of dots to be counted. He judged that seven was the limit of the number that could be accurately ascertained at one glance. Saltzman and Garner, (1948) used concentric rings as targets and decided that the method of reaction time was useful for the study of perception of number.

Von Szeliski (1924) found that fields of up to six dots or figures could be perceived without eye movement. Jensen, Reese, and Reese (1950) used fields of up to 29 dots, and found that errors appeared first in eleven dot fields and become more frequent in larger fields. After six dots, the plot of reaction time against number of dots becomes more scattered and a more shallow, positively accelerated curve appears. From this, Jensen, Resse, and Reese concluded that the subjects subitize up to five or six dots.

Taves (1941) studied judgments or numerousness (thought of as a stimulus variable like loudness), and believed that two mechanisms were involved, one used for up to seven dots, the other for larger fields.

He found that judged numerousness was decreased if the dots were arranged into a single figure, for example, a circle. Kaufman, Lord, Reese, and Volkman (1949) also suggested two separate discriminatory mechanisms; one acting for up to six dots, the other for larger collections.

All of these studies concentrate upon rapid apprehension or judgment of number, and are cast in a theoretical framework drawn from perception. Much effort has gone into determining the number of things that can be apprehended at a single glance, even though civilized adults do not count at a glance nor do civilized children.

In summary, two main methods of studying enumeration are the tachistoscopic, in which a display is briefly displayed and judged, and the reaction-time experiment, in which the material is displayed continuously but the subject responds as quickly as possible. Both methods are perceptual in character, though the reaction-time method permits some more complex methods of enumeration to occur. The main question has been, what is the maximum number of dots that can be perceived at a glance. This number has been estimated from about four to seven, and clearly depends upon the subjects and the display used.

It is amazing that although everyone knows that objects are enumerated by counting, most studies of enumeration or the judgment or number have attempted to rule out counting and ensure that only the primitive method of guessing is employed. Classical psychological theory assumes that perception is a basic and simple process from which other, more complex cognitions are to be derived; furthermore, the task of experimental psychology is to analyze compound experiences into their elements. Following these tenets, the process of counting is reduced to its elements, "perception" of numbers, and no effort is expended in synthesizing the process of counting.

The experiments reported below have a different concept of enumeration. First, we assume that enumeration is a compound task, but that the true and interesting structure of the task is found not by analysis into bits, but by determining how the parts go together. Second, we assume that perception of small numbers may be a skill developed by adults, a sort of short-cut to counting, rather than an elementary mental event. Throughout, we shall use the reaction-time method. The subject is instructed to enumerate as quickly as possible.

Experiments I and II

Method

Subjects: The first of these experiments used eighteen students in an ungraded class in the summer of 1964, 10 boys and 8 girls, ranging in age from 7-3 to 9-10. The second experiment used 20 college students, ages 17-22, drawn from the elementary psychology course at Indiana University.

Apparatus and Procedure: The stimulus cards were 8 1/2 by 11 inch sheets of white paper on which objects to be counted were outlined in black ink. Each subject was run separately, sitting across a table from the experimenter, in an otherwise unused classroom, which was arranged to reduce external visual distraction.

Each card was presented flat on the table, first shielded with an aluminum panel on a handle which was raised at each trial to display the card while simultaneously a standard timer was engaged. Response and latency of response were recorded. Each child's attention was attracted before each trial. Time was spent gaining the attention and cooperation of each child, while somewhat more formal instructions were given to the college students.

<u>Experimental Design</u>: Each subject counted the objects on forty stimulus cards which were presented in a different random order for each subject. Eight of these were fillers inserted to make the repeated appearance of certain numbers less obvious.

Variables were

a) number of objects to be counted

b) arrangement of objects; line, circle, rectangle, scrambled

c) all objects the same shape or different shapes.

Because of the rectangular arrangements, numbers with two fairly equal factors were employed - 12, 15, 16, and 18. Filler cards had 5, 9, 10, 11, 13, 14, or 17 objects. Objects were spaced about 1/2 inch apart, with the straight line being horizontal and the rectangles being 4x3, 5x3, 4x4, and 6x3. Objects on a card were either all the same shape or divided equally among squares, circles, isosceles triangles, and tear-drop shapes. The circle's diameter was 1/2 inch, with all other figures being of approximately equal apparent size. Thus, four numbers by four shapes by same-or-different yields 32 displays in a simple factorial design, replicated once per subject.

Results: Most children employed pointing and chanting, though some counted silently and sometimes without pointing, while college students counted silently, occasionally pointing. All children counted all displays making 106 errors out of 576 counts = 18.4 percent error, with a mean time of .60 seconds per object. The college students averaged .30 seconds per object, and made only 11.6 percent errors.

Larger sets of objects took longer to count for both groups. Linear equations fit goth sets of data well, being $\bar{t} = 0.75(N-3)$ for the children and $\bar{t} = 0.35(N-2)$ for the college students, where $\bar{t}$ is a mean reaction time in seconds and N is the number of objects to be counted. Since no measurements were taken with small numbers of objects, we have no reason to believe the straight line may be extrapolated back without curving. (See Jensen, et. al., 1950)

Object arrangement influences both groups' response latencies, with the children counting the rectangular array most rapidly, then the line, the circle, and the scrambled arrangement, in that order. The proportions of errors for the children were: rectangle, .097; line, .104; circle, .215; and scramble, .326, thus showing the scrambled arrangement most difficult both in time and errors. For college students, there was little difference between line, circle, and scramble, but the rectangular array was much faster.

Discussion: The results seem to show a difference between children's and college students' use of spatial arrangement. Children may use spatial arrangement to get a simple ordering of the stimuli, while college students seem to make special use of the rectangular array.

presumably by using multiplication. The children may not multiply faster than they enumerate or they may not have yet learned to apply the concept of multiplication to this situation.

When they cannot multiply, college students and perhaps children may divide the set of objects, possibly by perceptual principles of grouping, into smaller subsets, enumerate the subsets, then add the numbers. Sets of up to five or six may be subitized by skilled adult subjects. Calculation speed should depend upon convenience of grouping, enumeration, and addition. Rectangular arrays enforce a convenient subdivision which permits several simultaneous enumerations.

Insert Tables 1 and 2 about here

Indeed, both groups consistently count the rectangular array rapidly, but for the college students the array of sixteen objects is especially fast. This array was 4x4 and previous informal observations have shown that older children, and presumably college students are faster at squaring than at other multiplications. Sets of sixteen are not counted unusually rapidly in other arrangements, or in any arrangement by young children, so there is reason to believe that the advantage lies in multiplication.

It may also be noticed that rectangular arrays are counted more rapidly if all the objects are the same, whereas with circular and linear arrangements, counting is faster with different objects. Rectangular arrangement enforces a clear grouping on the set; differentiation of figures only competes with this grouping. Spatial grouping is more difficult with straight lines and circles, and subjects may use dissimilarities among figures to provide the necessary groupings.

Experiment III

In the first two experiments, only position and shape of the objects were varied. In the third experiment, color variation is also used so as to establish alternative perceptual groupings.

The general design of this experiment is based upon earlier studies of additivity of cues in cue-learning (Bourne and Restle, 1959; Bower and Trabasso, 1964; Restle, 1955, 1959a, 1962; Trabasso, 1960, 1963). In such experiments several problems are constructed using two or more dimensions. One group may learn to classify drawings based on their shape, disregarding color. Another group may classify drawings using color, disregarding shape. A third group learns a classification in which color and shape are redundant, which is easier than the other two (additivity of cues). Several mathematical models of cue learning have given quantitative predictions of learning speed based on the number of redundant relevant cues, the number of irrelevant cues, and the weights and strengths of such cues.

<u>Method</u>

<u>Subjects</u>: 24 college students, aged between 17 and 20, from an introductory psychology class.

<u>Apparatus and Procedure</u>: These were identical with Experiment I.

Materials and Design: Thirty displays, each consisting of a collection of shapes cut from construction paper and pasted to 8 1/2 x 11 sheets of white paper. The shapes were four colors; brick red, Air Force blue, light blue-green, and light violet. They were four shapes; circle, square, triangle, and parrallelogram. Test items had 12, 15, 16, 17, and 19 figures, and five filler presentations. In Condition 1, Color and Shape, the four colors and shapes were segregated, so that one section of the figure was all blue squares, for example, another part red circle, etc. In Condition 2, Color, the four colors were segregated but the shapes were scattered randomly over the field. In Condition 3, Shape, the four shapes were segregated but the colors were scattered randomly over the field. In Condition 4, Ambiguous, both the four colors and the four shapes were segregated, but differently, so that the four circles might be half blue and half red, and of the three triangles, two might be red, and one violet, etc. In Condition 5, Random, both color and shape were scattered independently and randomly over the field.

Hypothesis: If rapid counting, in the college student, depends upon grouping the field of objects, then Condition 1, having a strong perceptual organization based on color and shape should be counted most rapidly. Conditions 2 and 3, having color and shape arrangements respectively, should provide estimates of the relative grouping value of the two dimensions. Condition 4 gives the subject two conflicting modes of organization. If one mode is much stronger than the other, the condition should be about as fast as the condition employing only the strong variable. Nearly equal variables might lead to instability of grouping, perhaps forcing the subject to start over in his counting. Condition 5, having no principle of grouping available, should lead to counting rates close to those observed in Experiment II, scrambled arrangements for college students.

Results: Time to count was a linear function of total number of objects. The results here are very close to those obtained with college students for the scrambled condition in Experiment II. This indicates that the general findings are probably quite stable with successive samples from this pool of subjects.

Insert Table 3 about here

As predicted, the Color & Shape condition is counted fastest. Condition 2, Color only, is almost as fast, and Condition 3, Shape, is slower, so it appears that color is a much more effective dimension than shape in this experiment. This agrees with many experimental studies comparing the two variables, varied approximately as in this experiment.

In these data, shape segregation has little advantage over random, Condition 5. Therefore, it is predictable that the ambiguous condition, having one strong (color) and one very weak (shape) organizing variable in competition, would be almost as fast as color alone. Notice that it is faster than shape (Condition 3) or random (Condition 5), though a small effect of shape is found in the fact that Condition 4 is slower than Conditions 1 or 2.

An analysis of variance showed that both the number of objects and the five conditions produce significant, ($P<.01$), effects on speed of counting. These results establish that rapid counting depends upon grouping the material into subgroups. It is probable that the subjects segregated the whole display into color groups, "subitized" the number in each gorup, and then added these numbers to obtain the result. When no such grouping was possible, the subject was forced to rely on a relatively difficult and ambiguous grouping by position, or to count one by one.

Discussion: The experiments reported above show that in counting, the set of objects is grouped perceptually, in accord with Wertheimer's rules of propinquity, good continuation, and similarity. This grouping then serves as a support or basis for counting, in that the subject performs his counting operation within groups, and then in some fashion connects the various partial counts. One hypothesis is that college students subitize within each group, that is, determine the number by a somewhat mysterious but very rapid and accurate "perceptual" method. The subject might collect such numbers and add them at the end, or might subitize each group and add the result to a running total.

The fact that fairly young children show great sensitivity to the organization of the visual field, as shown in Experiment 1, suggests that grouping may play a more general role in counting. That is, even when a child is enumerating one by one, he may work rapidly and routinely within one group, then pause and consolidate or "store" his result in some way, and then attack the next group. The pausing, and the ability to divide the task into suitable parts, is possibly a generally important part of a long serial task.

References

Bourdon, Bertran. Sur le temps nécessaire pour nominer less nombres. Revue Philosophique, 1908, 65, 426-431.

Bourne, L. E., Jr., & Restle, F. Mathematical theory of concept identification. Psychological Review, 1959, 66, 278-296.

Bower, G. H., & Trabasso, T. R. Concept Identification. In Atkinson, R.C. (Ed.), Studies in Mathematical Psychology, Stanford, Cliifornia; Stanford University Press, 1963.

Casperson, R. C., & Schlosberg, H. Monocular and binocular intensity thresholds for fields containing 1--7 dots. Journal of Experimental Psychology, 1950, 40, 81-92.

Fernberger, S. W. A preliminary study of the range of visual apprehension. American Journal of Psychology, 1921, 32, 121-233.

Freeman, F. N. Grouped objects as a concrete basis for the number idea. The Elementary School Teacher, 1912, 12, 306-314.

Hunter, W. S., & Sigler, M. The span of visual discrimination as a function of the time and intensity of stimulation. Journal of Experimental Psychology, 1950, 26, 160-179.

Jensen, E. M., Reese, E. P., & Reese, T. W. The subitizing and counting of visually presented fields of dots. Journal of Psychology, 1950, 30, 363-392.

Kaufman, E. L., Lord, M. W., Reese, T. W., and Volkmann, J. The discrimination of visual number. Americal Journal of Psychology, 1949, 62, 498-525.

Restle, F. A theory of discrimination learning. Psychological Review, 1955, 62, 11-19.

Restle, F. A metric and an ordering on sets. Psychometrika, 1959, 24, 207-220. (a)

Restle, F. Additivity of cues and transfer in discrimination of consonant clusters. Journal of Experimental Psychology, 1959, 57, 9-14. (b)

Restle, F. Psychology of judgment and choice. New York: Wiley, 1961.

Restle, F. The selection of strategies in cue learning. Psychological Review, 1962, 69, 329-343.

Restle, F., Rae, J., & Kiesler, C. The probability of detecting small numbers of dots. Journal of Experimental Psychology, 1961, 61, 218-221.

Saltzman, I. J., & Garner, W. R. Reaction time as a measure of span of attention. Journal of Psychology, 1948, 25, 227-241.

Schlosberg, H. A probability formulation of the Hunter-Sigler effect. Journal of Experimental Psychology, 1948, 38, 155-167.

Taves, E. H. Two mechanisms for the perception of visual numerousness, Archives of Psychology, 1941, No. 265.

Trabasso, T. R. Additivity of cues in discrimination learning of letter patterns. Journal of Experimental Psychology, 1960, 60, 83-88.

Trabasso, T. R. Stimulus emphasis and all-or-none learning in concept identification. Journal of Experimental Psychology, 1963, 65, 398-406.

Von Szeliski, V. Relation between quantity perceived and the time of perception. Journal of Experimental Psychology, 1924, 7, 135-147.

Warren, H. S. The reaction time of counting. Psychological Review, 1897, 4, 569-591.

Woodworth, R. S., & Schlosberg, H. Experimental Psychology (2nd Ed.) New York: Holt, 1954.

Footnotes

1. This research was supported in part by NSF Grant GB 2848 from the National Science Foundation to the second author and USOE Grant 5-10-398 from the U. S. Office of Education to the second author. The authors wish to express their appreciation to the University School, Bloomington, Indiana and to Dr. Richard C. Pell, Principal, for his cooperation.

Table 1

Mean Time to Count, (Sec.) by Young Children

Number of Objects	Arrangement of Objects															
	Line				Circle				Rectangle				Scrambled			
	Same	Error	Diff.	Error	Same	Error	Diff.	Error	Same	Error	Diff.	Error	Same	Error	Diff.	Error
12	6.58	1	6.36	1	8.05	5	6.84	1	6.10	1	5.62	1	8.77	4	7.36	3
15	9.28	2	8.45	0	9.45	3	9.13	3	7.75	2	8.32	0	9.98	7	11.02	6
16	9.11	3	8.63	3	9.04	5	10.51	4	8.13	2	8.68	3	10.49	6	11.42	8
18	11.14	3	8.94	2	10.94	4	12.38	6	10.47	4	9.52	1	12.36	7	13.79	6
Proportion of Errors	.104				.215				.097				.326			

Table 2

Mean Time to Count, (Sec.) by College Students

Number of Objects	Arrangement of Objects															
	Line				Circle				Rectangle				Scrambled			
	Same	Error	Diff.	Error	Same	Error	Diff.	Error	Same	Error	Diff.	Error	Same	Error	Diff.	Error
12	3.92	0	3.48	0	4.24	2	4.23	2	1.92	1	2.15	2	3.57	1	3.55	2
15	5.28	4	4.83	3	5.51	4	4.70	3	3.20	0	2.67	0	4.59	2	4.67	2
16	5.59	4	4.97	2	5.99	1	5.83	1	2.19	1	2.70	5	5.45	5	5.45	4
18	6.07	2	5.63	0	6.53	3	6.31	7	3.34	2	3.81	0	5.54	0	6.13	5

Table 3

Mean Time to Count (Sec.), Experiment III

Number of Objects	Conditions									
	Color and Shape		Color		Shape		Ambiguous		Random	
	Time	Errors	Time	Errors	Time	Errors	Time	Errors	Time	Errors
12	3.18	1	3.17	1	3.56	3	4.39	2	3.80	3
15	4.39	1	4.52	5	5.37	2	4.76	5	5.16	4
16	4.55	0	4.32	1	5.56	4	4.47	2	5.21	0
17	4.82	9	5.44	1	6.93	5	5.83	2	6.72	5
18	5.39	7	5.96	4	6.40	6	5.93	1	7.12	6

Experiment II

REDUCING COMPLEX LOGIC EXPRESSIONS

At the foundation of all logic and mathematics is the simple sentential calculus or logic of propositions. It begins with whole sentences, unanalyzed, which may be true or false, and builds and analyzes compound sentences. Compound sentences are built from simple or atomic ones by use of the connectives AND , OR , NOT , and some optional ones.

A characteristic of sentential logic is that one can construct very complex combinations of true and false sentences, and then step-by step reduce them to atomic form, either T or F . The process is to find the innermost parentheses, and reduce the innermost sentence using the truth tables.

This process was studied in detail. In order to improve the perceptual situation, we used not conventional parentheses, but loops drawn entirely around the expression, forming concentric eggs. Thus, finding the innermost parentheses did not require counting parentheses, but merely finding the bull's eye.

Further to make the problem easy, experimenters presented it step-by-step, so that substitutions and always correct substitutions were actually made for the subjects. Mean time per step as a function of the size of the current display, the substitution to be made, and stage of practice were measured [illegible] teen-age students in a special summer reading class.

METHOD

Each of two women Graduate students in the Department of Psychology at Indiana University served as experimenters. Subjects were children, 8—15 years old, enrolled in a remedial reading program at the University School, Bloomington, Indiana.

Design: Each child was tested on 32 problems, four each of lengths 1—8 steps. A permutation of the eight lengths was given, then another and another until four permutations had been administered, so it is natural to divide training into four stages.

Each problem consisted of one or more stages. A stage consisted of finding the central parentheses in the given display, and noting its contents. The contents could be C•C , C•W , W•C , or W•W : or C*C , C*W , W*C , or W*W . The child was then to announce the correct substitution, shown to him and available on a separate card. The correct substitutions were C for C•C ; C•W , W•C , and C*C . The correct substitution was W for W•W , C*W , W*C , and W*W .

An interpretation of the stimuli, to give structure to the experiment, is this. Let C be a true sentence, and W a false one. Let • be the connective "OR" , and * be "AND" . Then the above "substitutions" are merely the truth table for these connectives.

The particular substitutions to be made were counterbalanced over the lengths of the stimulus expressions and over stages of practice. To accomplish this, four different sets of stimuli (displays from size 1 to 8) were prepared. These were permuted independently for each subject, and then the four different sets were given in different orders for different subjects ensuring that each set was equally often in each section of the experiment, as nearly as was possible with two sets of [illegible] subjects.

Procedure

The apparatus from Experiment I, the counting experiment, was again used. Each child was taken from class by agreement with the teacher and tested in a 6x6 ft. experimental cubicle used for testing and experimental purposes in the school building.

The instructions, which reveal most of the procedure, are quoted below.

Logic Experiment: Instructions

Today I am going to show you some puzzles. I will also teach you how to solve them. I am going to ask you to solve each one as we go along.

Here is an example of the kind of problem we will use. This symbol and these letters can be replaced by one single letter, either a C or a W. To find out which one is correct, we simply look up the problem on this chart. Do you see, it tells us that C*C can be replaced by C? Now, let's look at another problem. This says W·W. Can you find W·W on the chart and tell me what letter can take it's place? Let's see what the answer is. (Yes, you are right! That was very good./ No, you see, the answer says W. Will you look again at the chart and see if you can find the answer?) Now let's take the next problem and see if you can solve it all by yourself.

--2 operations:

That was very good. Let's look at another kind of problem. Part of this problem is in a box, isn't it? Will you point to the box for me? Now, this box tells you where to start work on the problem. We must find the answer to what is inside of the box first. Will you do that for me? Fine. Now, that makes the puzzle easier. We have shortened it to this. Will you find the answer on the chart? Fine. Now let's look at the puzzle we started with. You see, you really solved two puzzles which had been hooked together. Here's another problem. Show me where you start--and tell me what will take its place.

--3 operations:

When we have more than one box, we begin with the smallest one. For example, where would you start on this problem?

--Final Test:

Now I am going to show you some more problems. Some will be longer and some will be shorter than those you have just been doing. All of them can be solved in the same way, however. Simply remember to look for the smallest box and begin there. When you see a card I want you to point to the starting box and then to tell me what letter will take its place. Try to work as quickly as you can without making a mistake.

If a problem consisted of seven steps, then the experimenter had a deck of eight cards, starting with the original problem, next the problem reduced by making the correct substitution in the correct place, etc., down to the final "answer". A typical problem is shown in Fig. 2-1.

Insert Fig. 2-1 about here

The first card was shown by pulling away the aluminum shield, and the timer started at the same time. The child was required to point to the proper (inner) box on the card, and then announce the substitution. E would then cover the display, write down the reaction time in tenths of a second, place the next card of the problem (or the first card of the next problem) under the shield, and begin the next trial.

Results

It took children much longer to look up the substitutions for C·W, W·C, C*W, and W*C than for the elements having the same symbol twice, C·C, W·W, C*C, and W*W. The reason should be obvious; whenever two Cs are together, C·C or C*C, the correct substitution is always C; similarly, W·W and W*W yield W. Therefore, these substitutions are strongly supported by the natural hypothesis that C's lead to C, W's lead to W. These steps took an average of 2.26 sec.

The pairs C·W and W·C lead to C, but the pairs C*W and W*C lead to W. This discrimination is obviously somewhat more difficult to handle, since the strategy requires (a) notice that the two terms are unlike, and then (b) respond C if the middle connective is ·, and W if it is *. Such steps took an average of 4.22 sec., almost twice the time of the C·C and W·W steps.

Improvement occurred during the four sections of the experiment; a learning curve is shown in Fig. 2-2.

Insert Fig. 2-2 about here

The main point of the experiment was to determine if the difficulty of a step depended upon the complexity of the display or the size of the problem. Unfortunately, the importance of the C-W and C-C displays was not sufficiently appreciated at the time of the design of the experiment. The result is that different sizes of display differed, over-all, with respect to the proportion of C-W versus C-C displays. The raw data of mean time, as a function of display size, shows a saw-tooth shape correlating closely with the proportion of C-C and W-W displays.

The only indication of the results that can be obtained is to make an empirical correction. Using the data given above, that C-W displays average 4.32 and C-C displays average 2.26 sec., and using the known proportion of C-C elements at each size of display, a "predicted" score was calculated for each size of display. If there is no effect of size of display, the obtained mean latency minus the predicted should average near zero. The results of this calculation are shown in Fig. 2-3. Notice that mean latency is above the predicted score when the display is sufficiently large, but the effect appears only with displays larger than 5. Unfortunately, since the result depends upon a simple proportional correction factor, which may be somewhat in error, the results cannot be interpreted with confidence.

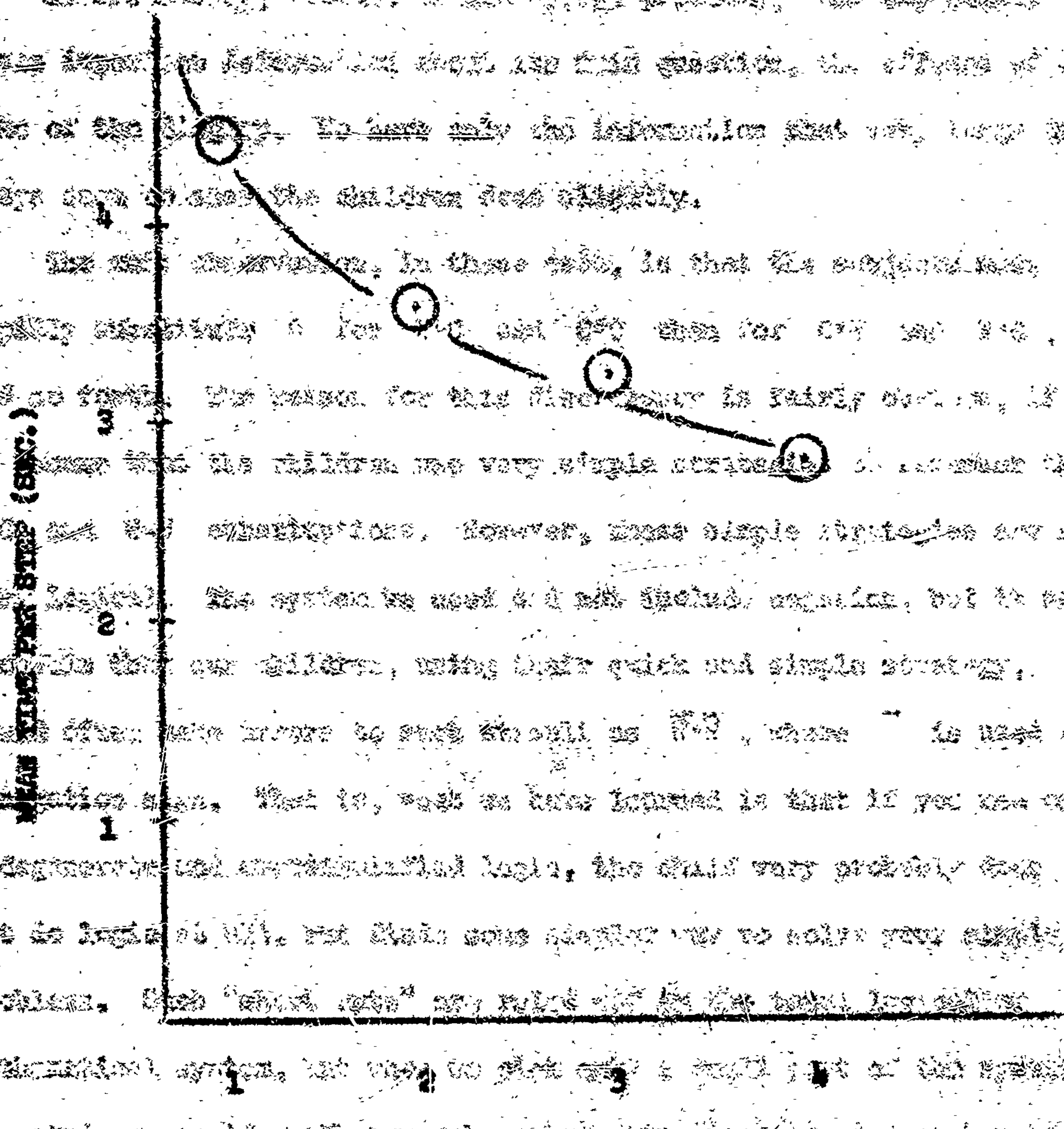

MEAN TIME PER STEP (SEC.)
1
2
3
4
1
2
3
4
QUARTERS OF EXPERIMENT (TRIALS)

Fig. 2-2

Insert Fig. 2-3 about here

Unfortunately, because of the design problems, this experiment gives imperfect information about its main question, the effects of the size of the display. We have only the information that very large displays seem to slow the children down slightly.

The main observation, in these data, is that the subjects more rapidly substitute C for C·C and C*C than for C·W and W·C, and so forth. The reason for this discrepancy is fairly obvious, if we assume that the children use very simple strategies to remember the C-C and W-W substitutions. However, these simple strategies are not very logical. The system we used did not include negation, but it seems probable that our children, using their quick and simple strategy, would often make errors to such stimuli as $\bar{W} \cdot \bar{W}$, where $\bar{\ }$ is used as a negation sign. That is, what we have learned is that if you use only a degenerate and oversimplified logic, the child very probably does not do logic at all, but finds some simpler way to solve your simple problems. Such "short cuts" are ruled out in the total logical or mathematical system, but when we pick only a small part of the system for study we accidentally may make other cues relevant, and so happen to make improper strategies correct. In this experiment, any time the child saw two Ws together, he could make the substitution W. That is not a correct strategy in a more complete logic using negation, but is a correct and very efficient strategy for our sub-problem.

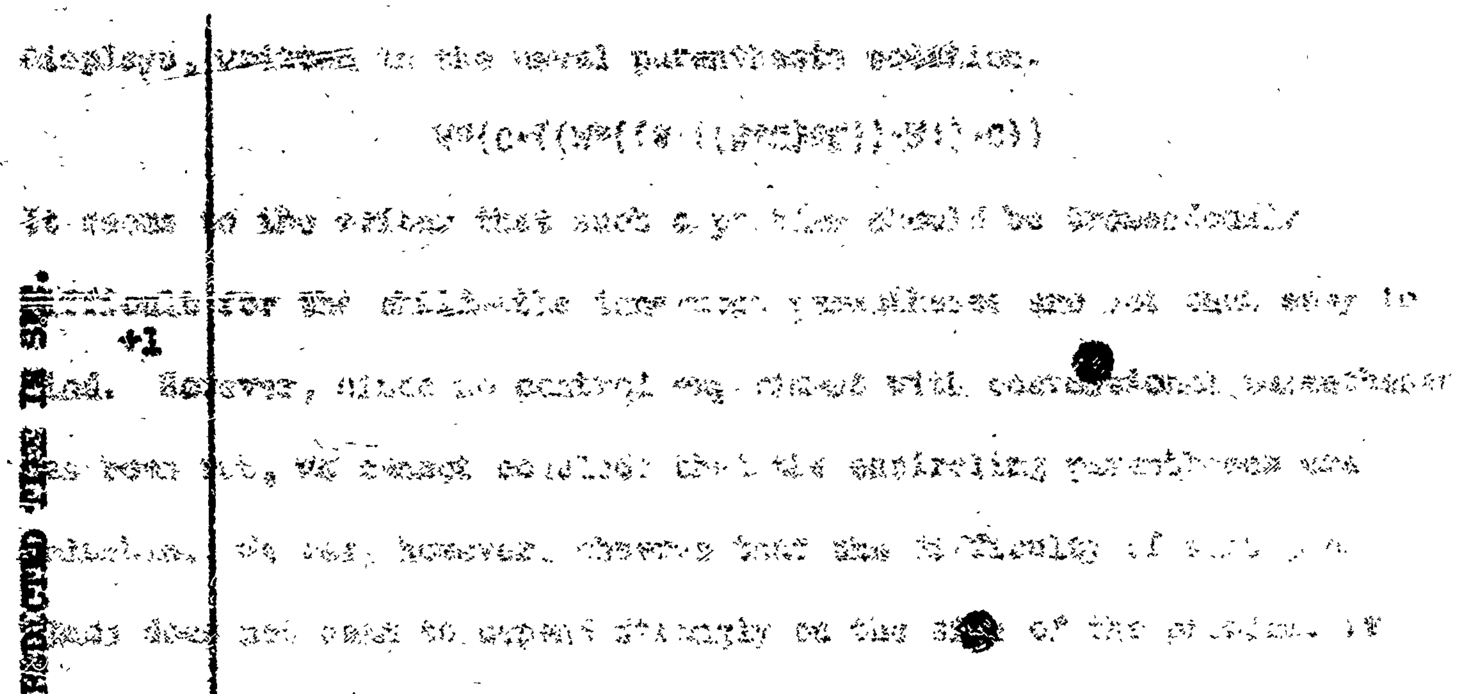

MEAN DISCREPANCY OBS.-PREDICTED TIME IN SEC.
+1
0
-1
1
6
8
SIZE OF DISPLAY IN C-C OR C-M ELEMENTS

Fig. 2-3

Perhaps more striking is the relatively small effect of the size of the display. We expected a large and powerful effect of size of display, but apparently using "boxes" or closed loop parentheses is tremendously easier for the children than the usual arrangement. For demonstration, notice the appearance of the second of the Eight-Component displays, written in the usual parenthesis notation.

W*(C·((W*((W·((W*C)*C))·W))·C))

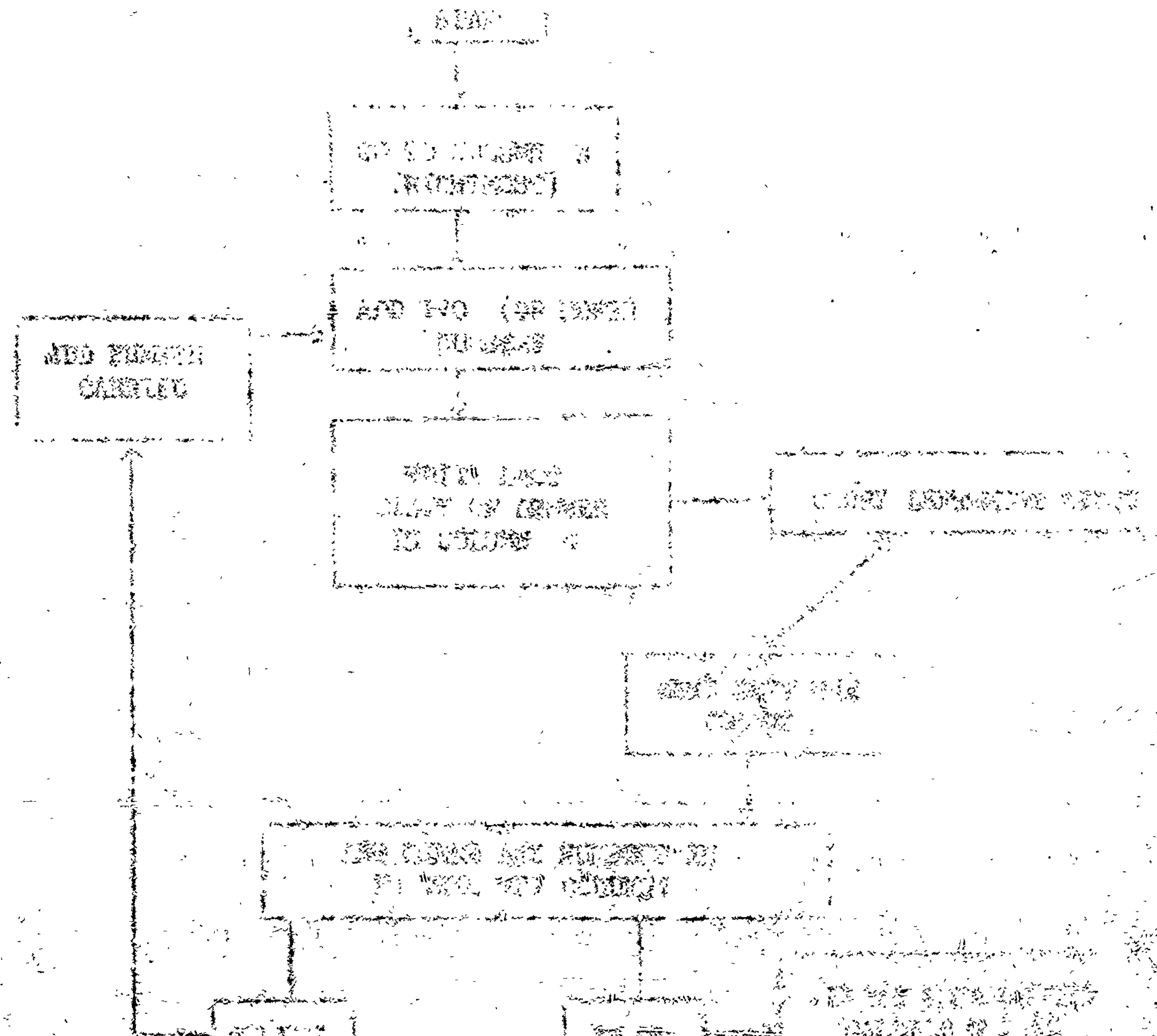
It seems to the writer that such a problem should be tremendously difficult for the child—the innermost parentheses are not that easy to find. However, since no control experiment with conventional parentheses has been run, we cannot conclude that the encircling parentheses are valuable. We can, however, observe that the difficulty of such problems does not seem to depend strongly on the size of the problem, if the encircling parentheses are used.

This last point may be of some convenience and use to the instructor of elementary logic, especially when it is taught in the lower grades. The indication is that children can rapidly learn routine simplifications of very large expressions. This might turn out to be a useful way to overcome the fear of a large symbolic expression that so often appears among college students.

Apparently, a large amount of the time in this experiment was used, not finding the center of the display, but looking up the substitutions. The learning curve shows the progress of learning of the truth table—notice that progress is relatively slow and the students are still improving considerably after the first whole sequence, in which they looked up a total of 36 substitutions, more than four trials per element. It must be concluded that mastery of the truth table, including only conjunction and disjunction, may take quite a while.

To add large numbers, the subject must take the [illegible] right-hand digits, add them, write the second digit of the answer, "carry" the first digit, and move one step to the left. If he has reached the end of the number he writes down the number he has [illegible]. Otherwise he adds it to the sum of the two numbers in that column.

The process can be [illegible] as follows:

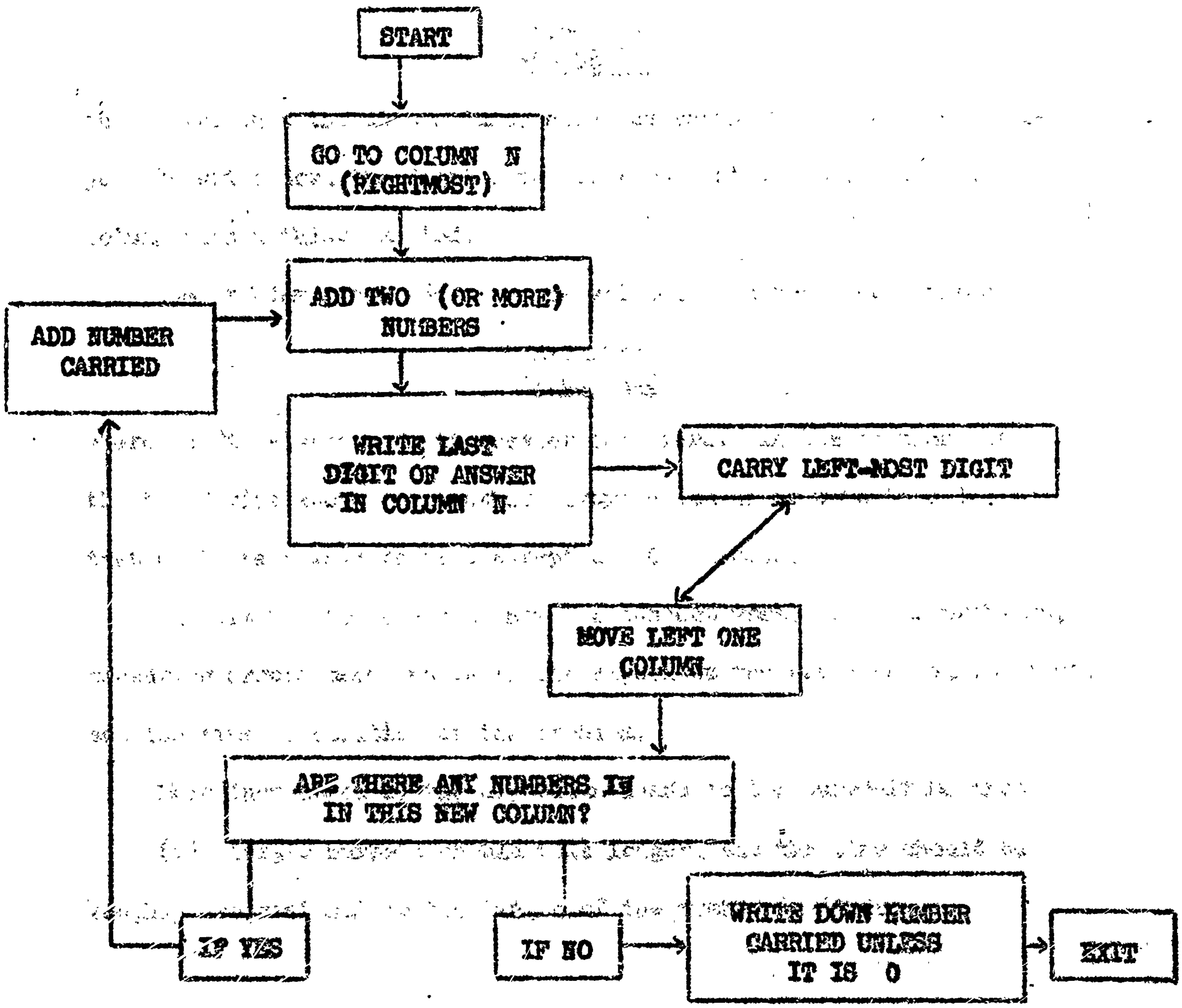


Experiment III

ADDITION OF TWO LARGE NUMBERS

To add large numbers, the subject must move the the right-most digits, add them, write the second digit of the answer, "carry" the first digit, and move one step to the left. If he has reached the end of the number he writes down the number he has carried. Otherwise he adds it to the sum of the two numbers in that column.

The process can be diagrammed as follows:



The psychological process may diverge from that of the algorithm or program listed above, in that the child adding a large number may need to organize the task into parts larger than particular columns. This would be reminiscent of the findings of Exp. I on counting, that a set, to be counted accurately and fast, must be capable of being organized into parts.

In the present studies, designed and carried out by Miss Penelope Peterson in the summer of 1965, the task was divided up by inserting a column of zeros in the problem. Consider the problem

$$\begin{array}{r} 870940740 \\ +\underline{640680690} \end{array}$$

and notice that the zero columns serve as rests; they permit the subject to write down the digit being carried, and to begin the next column with nothing carried.

The problem given above can be written in more general terms as

$$\begin{array}{r} xx0xx0xx0 \\ +\underline{xx0xx0xx0} \end{array}$$

where x is always a digit greater than zero. In the problems used, the two digits above one another always added up to more than 10, so that a 1 is always carried except at 0 columns.

The aim was to determine how the subject worked on such problems; measuring errors made, where in the algorithm the error seemed to occur, and the time to solution of the problem.

Mere inspection of the algorithm leads to the supposition that

(a) Larger numbers should take longer, and the time should be roughly proportional to the length of the number in digits.

(b) Zero-columns should speed the process since there is no addition step on that step, and no digit to carry the next column. However, since something must be written down and the move to the next column is required, at least, zero-columns should require some time. In addition, our ideas about organization of the task would suggest that introducing zero columns might not only reduce the average performance by introducing a faster operation, but might also speed the other columns—at least, a long unbroken sequence of additions without a zero column may lead to very long delays and disruptions.

Method

A total of 12 problem-structures were generated, as follows:

Table 1. Problem Structures

Problem Number	Structure
1	xxxx
2	0xx0
3	xxxxxxxxx
4	xxxx0xxxx
5	xx0xx0xx0
6	xxxxxxx0x
7	xxxxxxxxxxxxxxxxxx
8	xxxxxxxx0xxxxxxxxx
9	xxx0xxxx0xxxx0xxxx
10	0xx0xx0xx0xx0xx0xx0
11	xxxxxxxxxxxxxxxx0x
12	x00xxxxxxxxxxxxx0x

Notice that those problems are of length 4, 9, or 19. Problems 1, 4, and 9 give blocks of four numbers separated by zeros. Problems 2, 5, and 10 give blocks of two numbers so divided. These six problems constitute the problems that are divided, by zeros, into sizes of sub-problems that are likely to be most efficient. Problems 3 and 7 are undivided, and problem 8 is divided into two blocks of length 9. Finally, Problem 6 has eight numbers and a 0, but the 0 is so near the end as not to divide the problem into convenient pieces. Therefore, it is a control for Group 4, that also has eight numbers and a 0, but is divided in half. Similarly, Problem 11 is a control against 8, and Problem 12 is a control against 9; in each case the control has the same number of 0s and numbers, but the 0s are not placed so as to divide the problem into equal-sized and convenient parts.

A total of 12 school children, reading problems attending University School in Bloomington, Indiana, were tested. Six Ss were ages 10-12, and the other six were 13-15. Each child was tested separately in a testing cubicle, face to face with E. The experimenter showed the problem printed large on a ditto paper, and the subject worked directly on that problem. Timing was with a Standard Timer in seconds.

Each step of the child's work was followed by E, who noted whether the child started correctly, added the numbers, wrote down the unit's digit, carried any value in the ten's place, and then shifted to the next column. All errors were recorded and located in the algorithmic scheme if possible, at the time they were committed.

The twelve problem-types were given in a latin-square order, one order to each of the 12 subjects.

Results

Table 2 shows the problems, the mean time to solution of each type of problem, and the mean errors, both for younger (10-12) and older (13-15) children. In order to make the relationships easier to see, the mean time to complete each problem has been divided by the number of columns in the problem, since each column of the addition problem is a single cycle through the algorithm. Thus, if all problems were equally difficult except for their lengths, and time and error were therefore proportional to length of the problem, all values in the table would be equal.

Table 2

Mean Time Per Column, and Mean Errors Per Column

Problem	Younger Children		Older Children	
	Time	Errors	Time	Errors
1. xxxx	9.3	0.00	3.8	0.08
2. 0xx0	6.2	0.13	2.9	0.00
3. xxxxxxxx	9.0	0.00	4.3	0.02
4. xxxx0xxxx	9.6	0.11	4.1	0.04
5. xx0xx0xx0	9.1	0.07	3.5	0.06
6. xxxxxxx0x	10.5	0.06	4.1	0.09
7. xxxxxxxxxxxxxxxxxxx	10.7	0.11	5.2	0.02
8. xxxxxxxxx0xxxxxxxxx	10.1	0.06	4.7	0.03
9. xxx0xxx0xxx0xxx0xxx	9.3	0.03	4.3	0.04
10. 0xx0xx0xx0xx0xx0xx0	8.4	0.05	3.5	0.03
11. xxxxxxxxxxxxxxxxx0x	11.2	0.08	4.6	0.05
12. x0xxxxxxxxxxxxxxx0x	9.8	0.05	4.3	0.05

From Table 2 we see that, in general, the younger children appear to have shown the effects of grouping and the older children did not.

Consider the following control pairs. Problems 4 and 6 have the same distribution, 8 numbers and a single 0. Problem 4 is divided in half, Problem 6 is cut near the end. For the younger children, the mean times are 9.6 and 10.5, in favor of the better division. The older children gave mean times of 4.1 versus 4.1, no difference.

Problem 8 gives 19 numbers divided by a single zero in the middle, Problem 11 has its single zero near the end. Mean times for younger children were 10.1 versus 11.2. The older children show times 4.7 versus 4.6, a slight advantage for the less-well-organized problem.

Problems 9 and 12 have three 0s in 19 columns, dividing into blocks of four numbers or clustered near the ends. The younger children show times of 9.3 versus 9.8, whereas the older children show means of 4.3 for both problems.

Thus, in all three cases, the younger children do the more evenly divided problem in from 5 to 10 percent less time, whereas the older children show no difference as great as 2 percent between comparable experimental and control problems.

However, the mean time per column does not vary over a very wide range as one goes down the columns, and this shows that the mean time to solve such an addition problem is roughly proportional to the number of steps in the problem.

In Experiment I on counting, we found that larger assemblies took more time <u>per item</u> in the process of enumeration. The same thing happened in Experiment III on addition, for the mean time per column became higher with more columns. The only uncontaminated comparison is between problems with no zero columns. The results are shown in Fig. 1.

Insert Fig. 1 about here

As more zeros are introduced into a problem, the time shortens as a linear function of the number of zeros. This relationship is shown in Fig. 2.

Insert Fig. 2 about here

Generally speaking, this means that the time taken to add two large numbers is roughly proportional to the number of digits involved, except that the subject is slightly slower with larger numbers. Zero columns speed the process, taking only about 1.2 sec. per column, as compared with other number pairs that consume about 4.5 sec./column, for older children. For the older children, it does not matter where the zeros are put.

The younger children show a somewhat different pattern. As mentioned above, dividing the problem into subsequences of two or four columns separated by zeros, seems to help the younger children. As is seen in Figure 3, younger children also show an increase in time almost proportional to the length of the problem, with a slight slowing of the

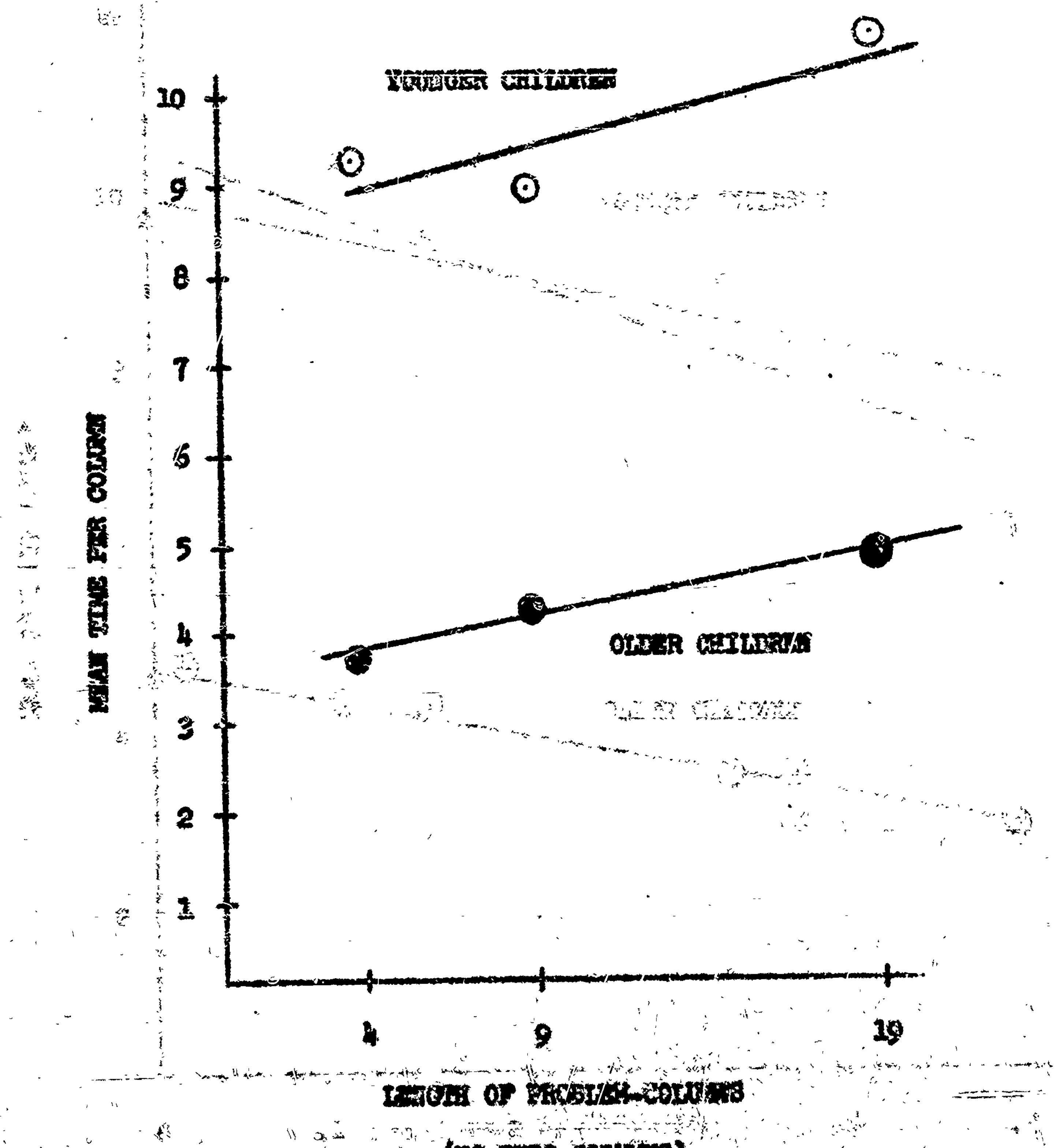

YOUNGER CHILDREN
OLDER CHILDREN
MEAN TIME PER COLUMN
10
9
8
7
6
5
4
3
2
1
4
9
19
LENGTH OF PROBLEM-COLUMNS
(NO ZERO COLUMNS)

Fig. 1

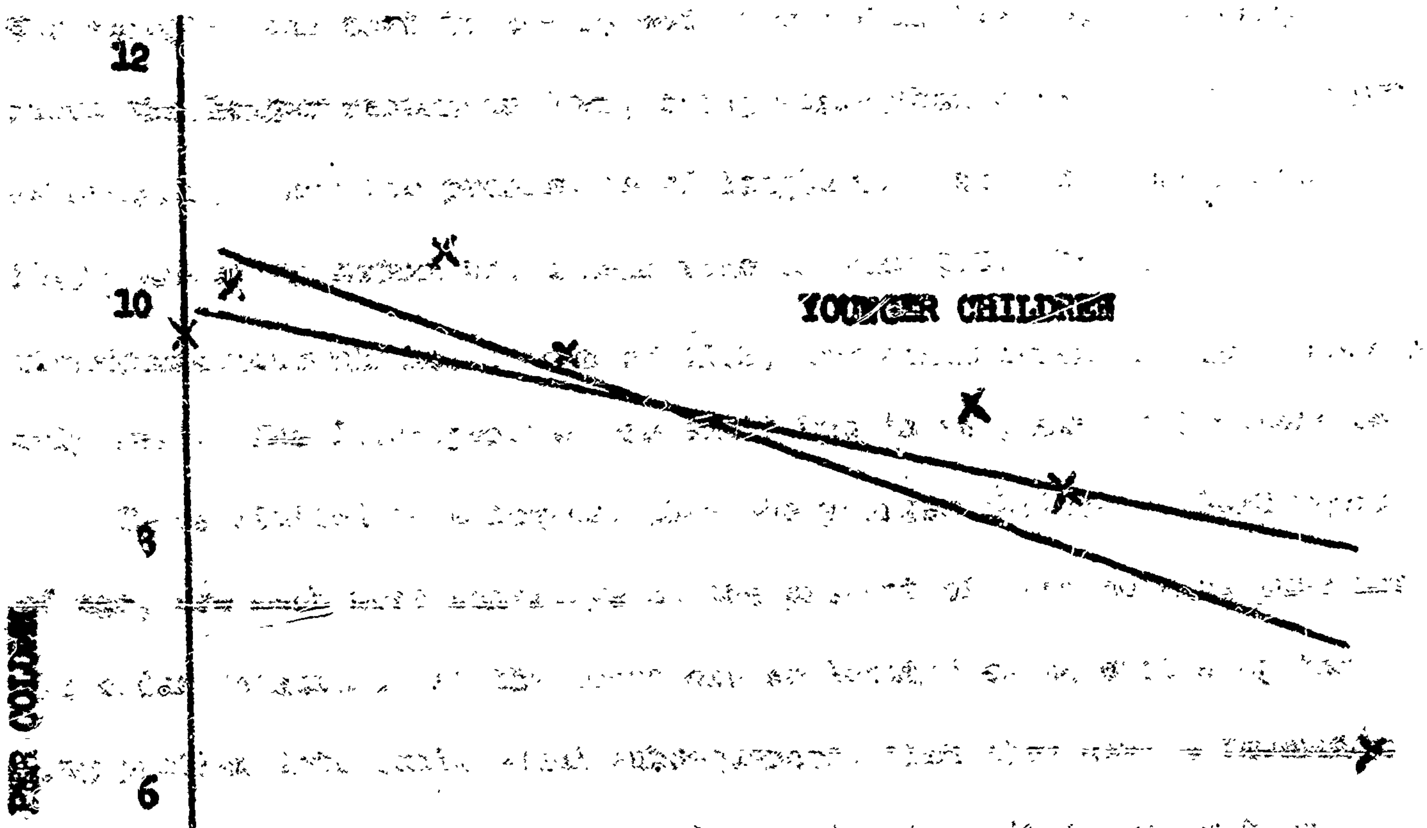

12
10
8
6
4
2
YOUNGER CHILDREN
OLDER CHILDREN
MEAN TIME PER COLUMN
0
.1
.2
.3
.4
.5
PROPORTION OF O COLUMNS

Fig. 2.

process when the problems are long. The effects of increasing the number of zeros is ambiguous, mainly because so much depends upon whether the zeros divide the problem into small equal segments or not. For example, one zero in the second column from the right actually slows the longer sequences down, being worse than a whole solid problem of numbers. When the problem is of length 9, xxxxxxx0x has a mean of 10.5 whereas xxxxxxxxx has a mean time of only 9.0. Similarly, xxxxxxxxxxxxxxx0x has a mean of 11.2, xxxxxxxxxxxxxxxxx has a mean of only 10.7. The discrepancies are small but in an unexpected direction.

These limited data suggest that the younger children, 10-12 years of age, are much more sensitive to the present of zero columns than are the older children. If the zeros are so located as to divide up the long problem into small equal subsequences, then they have a favorable effect. If the zeros are asymmetrically located, it is as if they break up the "set" or rhythm of the subject, and actually slow him down. The younger children, being somewhat less experienced in calculation on such long problems, may set up a temporary set of carrying ones, and may be disrupted by occasional and otherwise useless interruptions of this repetitive task.

In Table 2 it will be noticed that the total error data are not very different in the two groups. One rather striking finding is that the longest problems, numbers 7, 8, and 11, that have the fewest 0s, draw relatively large numbers of errors from the young subjects. Errors are almost uniformly distributed over problems by the older subjects.

It is possible that the subjects make errors by failing to carry a one from the previous column. This particular error can easily be identified, for it occurs only in columns that are not at the right edge of the problem and do not have a zero column to the right, (for they have no digit to carry), and should always be errors in which the answer given is one too small, since the number to be carried, in adding two numbers as constructed here, is always one.

Totaling the data for all 12 subjects, there are a total of 1512 instances in which a one is carried. In 40, or .0254 of these cases, the answer was too low by 1. There were 384 cases in which there was no 1 to carry, and of these 9, or .0234, were too low by 1. The difference in proportion is only .003, and appears negligible. Furthermore, when there is a one to carry, there were a total of 80 errors, of which 40 or .50 were errors too low by one. In the cases in which there was no number to carry, only 15 errors were made, of which 9 or .60 were answers too low by one. Thus, the particular error that would be symptomatic of failing to carry a digit, giving an answer too low by one, is actually slightly more common when there is no digit to carry.

Discussion

From the above analyses, we conclude that adding two large numbers by hand is a step-wise process, taking approximately 4 seconds per column for older children 12-14 and about 10 seconds for younger children, 10-12 years of age. Zero columns are relatively fast for the older children, taking only 1.2 sec. Zero columns serve to break up the long problem for the younger children. This helps if the long problem is chopped into equal-sized parts, and hinders if the zero column is near the right (beginning of the problem) in which case it apparently disrupts a set. There is no tendency for the subjects to err consistently by forgetting to carry the ones.

Apparently, this task, which is familiar for the younger children and very easy for the older ones, shows a late stage in the evolution of a serial or algorithmic skill. The younger children cannot carry the long sequential task out very efficiently unless it is divided into segments. However, they also make some use of a temporary "mental set" and apparently "carry one" efficiently if that is required all across the problem. Thus, we may envisage the younger children treating such a problem as a whole, with the various steps strongly interrelated and either organized or homogenized.

The older children, for whom this is an easy and routine task, apparently handle it much less as an integrated task. That is, they add even very large numbers and are hardly slowed down by having a 19-digit number to work with. Dividing such a task up with zeros has no particular organizing effect, for the distribution of zeros makes no difference to such older children. Since the zeros enable the subject to leave out part of a step, zeros speed up the problem by taking less time, and have no other detectable effect.

In the usual psychological language, it would appear "mechanical" or "non-intuitive" to handle the problem as the older children do, neither organizing nor taking advantage of special, repetitive patterns. The very considerable speed advantage of the older children belies such a conclusion. Furthermore, it must be realized that there is not, in arithmetic in general, any pattern to the digits that must be added. The "pattern" of equal spacing of zeros, etc., is purely accidental and meaningless in the context of addition. The skill of adding large numbers should not be influenced by accidental patterns of the numbers, for such patterns are in general irrelevant to the problem.

Thus we must not be too hasty in assigning a positive value to organizing processes. In the present experiment, children 10-12 have "organized" information that is not generally useful in addition, and as a result give more interesting data but are far less efficient than older children, 13-15 , who showed almost no signs of the organizational tendencies made available by the experimenter.

Experiment IV

PREPARATORY SET IN MULTIPLYING

There is a story told about an ancient university in Germany, back when it was young and raw as an American State University. It is said that a successful burgher visited the professor of mathematics at this German University asking advice in the education of his oldest son, whose talents the father wished to advance. He inquired whether it would be satisfactory for his boy to study at the German university, or whether he must be sent all the way to Italy. The professor is supposed to have replied, "If it will be enough that he be able to add and subtract, then your son will be educated well enough in Germany—but if he must also multiply and divide, then I fear he must be sent to study in Italy."

Whether this story is true or not, it illustrates the inherent difficulty of multiplying and dividing, as compared to addition. In public schools for some time, the most powerful weapon employed in teaching children to multiply has been the "multiplication table", memorized and performed with reasonable speed and skill by many, if not all, children.

A few years after this training the same child is in college, and if asked a simple problem like "What is 6 x 4" will give the answer fairly quickly, and will probably be correct. If asked how he knows the answer, the student replies that he just remembers it. Most of our college students perform the multiplication table as a well-practiced skill, and one might as well ask them how they walk as how they multiply—the two skills are almost equally unconscious.

However, any skill must have some limitations that may tell us something about its structure. If the college student has memorized the multiplication table and performs such tasks from memory, then we should determine the function of memory as well as possible. What, exactly, is remembered, how is it looked up, what errors are possible, and what is the procedure of search of memory (if search is necessary)?

It is natural to contrast remembering with calculating. In the psychological literature there is a dichotomy between thinking and memory, between memorizing and organizing, between productive and reproductive thinking. The subject who says he merely remembers the answer is also saying that he does not perform any conscious calculation.

However, remember that in the multiplication tables 1—9 , there are a total of 81 answers, not all different, ranging from 4 to 81. As the name of the multiplication table suggests, the factors and products can be arranged in a square table. When a subject "remembers" one of these items, it is natural to think of him as looking it up in a table inscribed in his memory. If that is true, then he actually may engage in some sort of search, and as a result may be faster in arriving at some products than others. Under time pressure some subjects will make errors—their errors should form a pattern, if we believe as seems likely that errors are often "near misses" in searching for the correct answer.

Of course, this search of memory could be thought of as a kind of calculation. In the language of computers, the subject must either calculate the answer or, if the answers are located in a table, he must calculate the address of the answer. In more psychological terms, we may imagine a subject using a variety of cues and discriminations to lead him to the correct memory and away from certain possible errors. This sequence of discriminations, especially if they direct a stream of behavior tha mainly comes from the subject's own ratiocinations rather than as responses to externally-changing stimuli, may be called a "calculation".

From this it may be conjectured that the process by which a college student multiplies two numbers is either a calculation or a process somewhat like a calculation. Since the task is skilled we do not know what the components may be and must study them by experiment.

The present experiment studied 50 college students doing the multiplication-table as rapidly as possible. The experimenter presented the two factors and as soon as possible after the second was presented (verbally), the subject gave the product verbally. Differences between different problems in time to answer, and in accuracy, may help us understand the structure of the skill.

However, if we think of multiplying as a search process, then it is apparent that the speed of completing a given search depends not only on where the answer is, but also on where the searcher starts. If the problem is 8 x 7 and the subject has started a search at 2 x 2 it may take him a long while to arrive at the solution. If, on the other hand, the subject began his search at 9 x 9 , then he may soon arrive at 8 x 7 .

Therefore, to understand the search process we must attempt to start the subject at different points, and also give a variety of problems, and see if the subject's times and errors reveal anything about the path taken between the two points. Of course, this theory is mainly metaphorical, but it does suggest that we wish to vary not only the problem, but also the subject's preparatory set or starting point.

To control preparatory set each subject was given a whole sequence of multiplication problems such as 7 x 3 , 7 x 6 , 7 x 4 , 7 x 9 , etc., intended to set up an "expectation" or "set" for problems of the form 7 x n , where n would vary. When this set was pretty well established, variations were rung such as 4 x 7 (reversing the order to n x 7) and also entirely extraneous problems such as 9 x 5 . If multiplying is a search process of any sort, then the "preparatory set" should be established as the subject learns where to begin his search, and exceptional items in the list should be slowed down because they are not to be found in the immediate search area.

Method

These experiments were performed in the Psychological Laboratories at Indiana University, in a small laboratory room. Apparatus consisted mainly of a small tape recorder on which both the experimenter's "stimulus" and the subject's "responses" were recorded.

__Subjects:__ Subjects were 50 students from elementary psychology lecture and laboratory courses at Indiana University, working for class credit. There were 22 male and 28 females, 17-25 years old.

__Procedure:__ Subjects were instructed that this was a study of problem-solving and they were to multiply the two given numbers as fast as possible. E gave the two numbers verbally about 1/2 sec. apart, and S responded verbally with the answer. The next problem was presented about 3--5 sec. after the response, so as to prevent E and S falling into a rhythm.

Responses were recorded on a tape recorder, and later the tape was played before a voice-key system and the times recorded manually.

__Design:__ Each series consisted of 25 problems, consisting of 21 items of the set-inducing type, two "reversals", and two "exceptions". In the list 7 x n, for example, the 21 set-inducing items would be 7 x n, that is, 7 x 3, 7 x 9, etc. The reversals would be n x 7, for example, 8 x 7. The "exception" item has no 7 in it, being for example 2 x 3. Each subject was given 5 such lists, the common elements being 2, 3, 5, 7, and 9, given in permuted orders for different subjects. The set-inducing items had their constant digit first, (7 x n) in half the lists, and second (n x 7) in the other half of the lists.

With this design it is possible to compare set-inducing, reversal, and exception performance because the same item (e.g., 7 x 5) would appear at almost-corresponding points in various series, being set-inducing in one series, a reversal in another series, and an exception in still a third. Furthermore, it is possible to study the development of the set, and the time and errors made to problems of various sizes, by singling out only set-inducing items for consideration.

Because a "reversal" is not a reversal until the set is well established, the first exception or reversal did not appear until trial 10-12 of a series of 25. The two reversal and the two exception items were then spread out quite evenly over the last 13-15 items so that each was preceded by at least three consecutive set-inducing items.

Results

The over-all mean time to multiply two numbers in the set-inducing group was 0.86 sec.; reversed items averaged 0.88 sec., only slightly longer; but exception items averaged 1.00 sec., considerably slower. From this we may conclude that the effort to establish a determining tendency or set was successful. The results are shown in Fig. 4-1.

If we now consider only the set-inducing items, it is apparent that larger numbers take longer to multiply than smaller ones. Several possibilities can be entertained. For example, it may be that the child learns the multiplication table up to 3, then up to 4, etc., and is more skillful and rapid at those smaller numbers. If so, it is natural to suppose that the time-to-calculate should depend only on the larger of the two numbers given, for the larger number determines the largest table that must be known. Fig. 4-2 suggests strongly that this conjecture is incorrect. The reason is that if anything, time to

Insert Figs. 4-1 and 4-2 about here

calculate seems to depend more on the smaller than on the larger digit; that is, the lines are displaced up-and-down as much if not more than the lines tilt, increasing with the larger digit.

Another possible means of multiplying is for the subject to look the answer up in a two-way table. If so, then when given the first number, the subject might select a row of the table, then when given the second number might scan down that table looking for the answer. There is no reason to suppose that time-to-multiply, here measured as time from the second factor to the answer, would depend upon the first factor. No matter what the first factor is, the subject can select his row before the second factor, and the measured reaction time, begins. Figure 4-3 shows that the reaction time does depend strongly on the magnitude of the second digit, as might be expected from this hypothesis, but also depends heavily on the first digit, for notice that the curves in Fig. 4-3 are separated widely, as well as being tilted.

Insert Fig. 4-3 about here

Finally, it may be supposed that the subject actually calculates the solution to the problem by a very rapid process of addition or counting. If so, the time required should depend upon the magnitude of the answer itself. In Fig. 4-4 is plotted that mean reaction time as

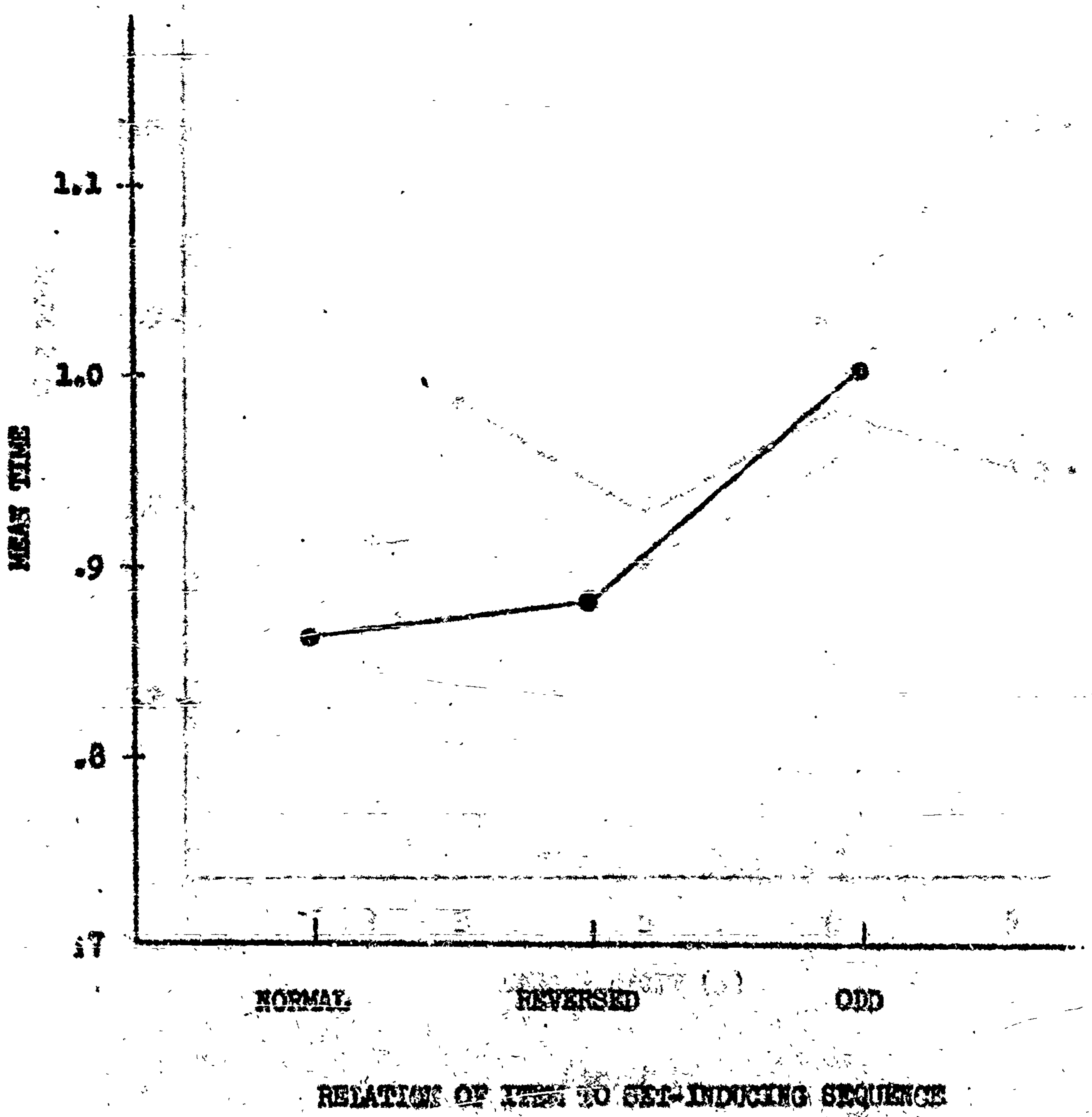

MEAN TIME
1.1
1.0
.9
.8
.7
NORMAL
REVERSED
ODD
RELATION OF ITEM TO SET-INDUCING SEQUENCE

Fig. 4-1

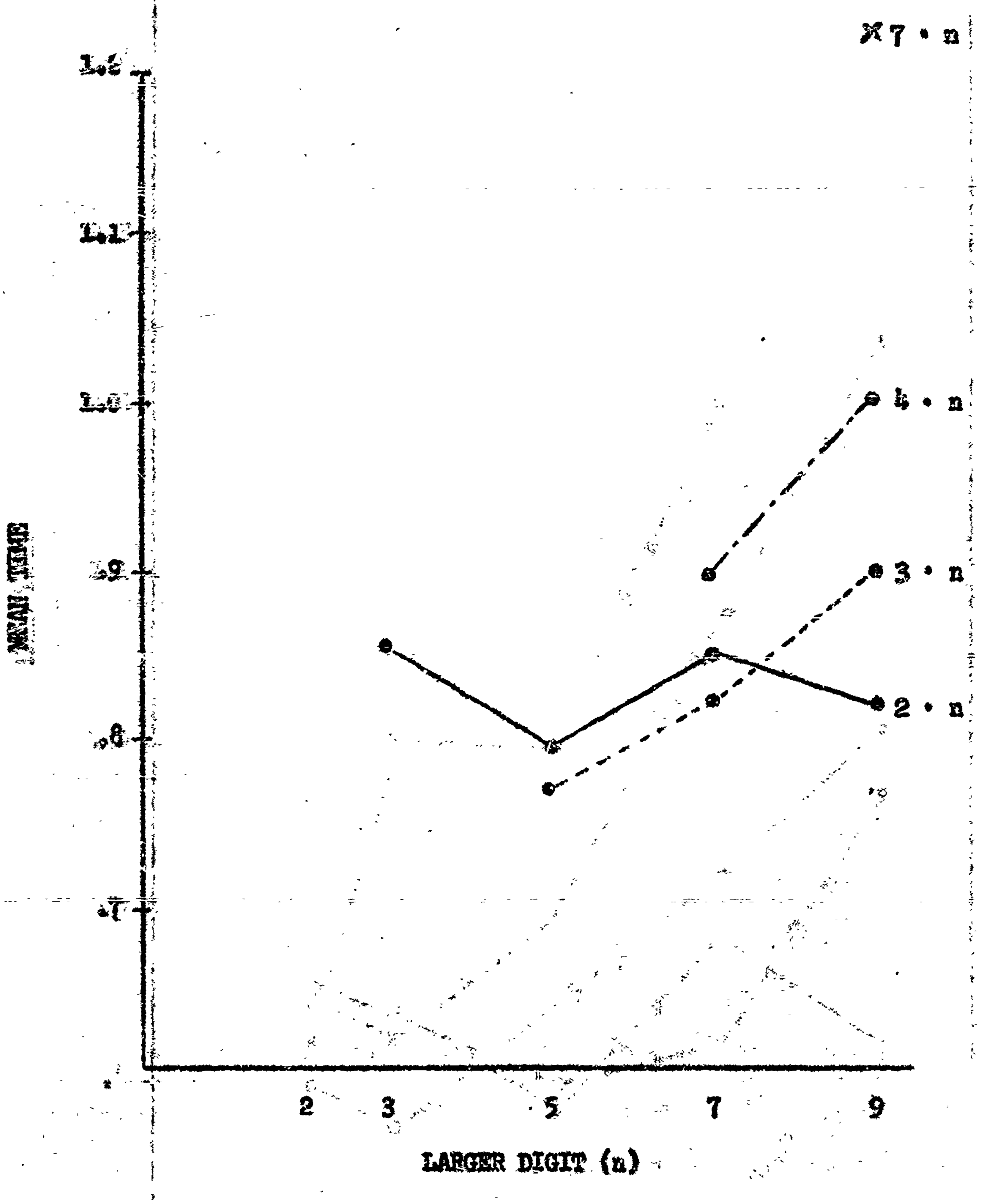


Fig. 4-2

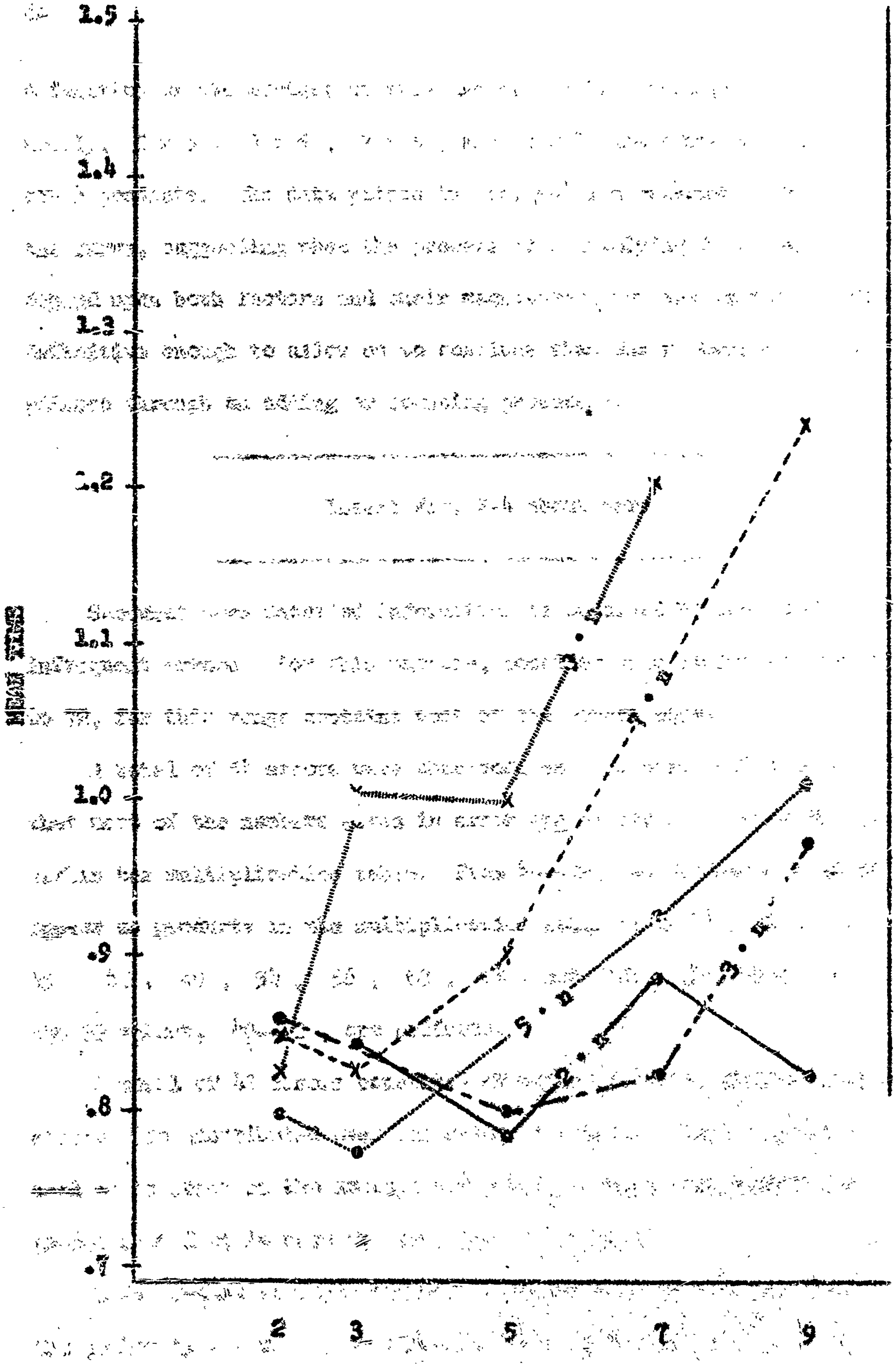


Fig. 4-3

a function of the product or true answer. Notice that products of 5, namely, 2 x 5, 3 x 5, 7 x 5., and 9 x 5 are somewhat faster than other products. The data points in Fig. 4-4 lie reasonably close to the curve, suggesting that the process of multiplying does seem to depend upon both factors and their magnitudes, but the result is hardly definitive enough to allow us to conclude that the process is accomplished through an adding or counting procedure.

Insert Fig. 4-4 about here

Somewhat more detailed information is obtained by considering the infrequent errors. For this purpose, consider only responses from 40 to 72, fo. this range contains most of the errors made.

A total of 54 errors were observed, and the most striking fact is that most of the numbers given in error are products of other values within the multiplication table. From 40—72, the following numbers appear as products in the multiplication table up to 9x; 40, 42, 45, 48, 49, 54, 56, 63, 64, and 72.. Thus only ten of the 33 values, 40—72, are products.

A total of 42 errors consisted of wrong products, whereas only 12 errors were distributed over the other 23 numbers. Each product was used as an error on the average 4.2 times, whereas each number that is not a product is used an average of 0.5 times.

This result strongly suggests that subjects do not just "calculate" the number by a counting or adding process that might run short or

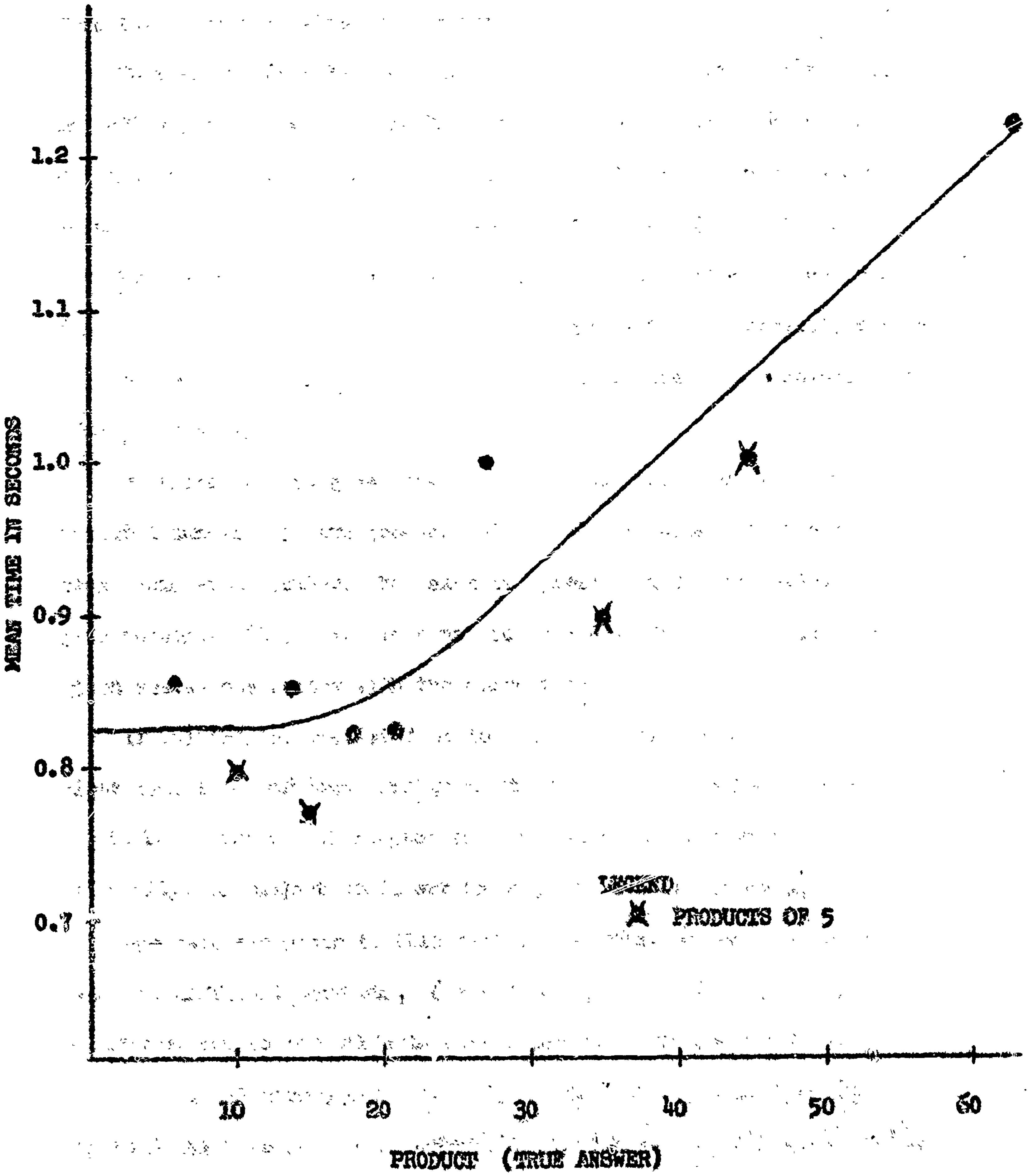

MEAN TIME IN SECONDS
1.2
1.1
1.0
0.9
0.8
0.7
10
20
30
40
50
60
PRODUCT (TRUE ANSWER)
LEGEND
PRODUCTS OF 5

Fig. 4-4

over-run. On the contrary, subjects seem to have available the list of answers or products, and attempt to select the one correct answer from this list when given the factors.

Further evidence that we deal with something other than a counting or adding process is given by the fact that 10 of the 54 errors were of magnitude greater than 10 , that is, the difference between the answer given and the true product was greater than 10 . Considering the high proportion of accurate responses, (a total of 6,250 responses were elicited, of which 6,196 or about 99 percent were correct), errors as large as 10 should be very rare if produced merely by a counting or adding mechanism.

Of those answers given that are wrong but are products, 42 responses in total number , 31 are products of one of the numbers in the set times some wrong number. For example, given 7 x 6 the subject might give response 63 , which is equal to 7 x 9 . Therefore, the answer given shares one factor with the correct answer.

If multiplying consisted of looking up values in a table, one might expect the subject usually to obtain a correct reading of the row but to be in error with respect to the column, or vice-versa. If so, ordinarily the subject would err in only one of the two factors.

The main exception to this pattern was found in the confusion of two very difficult products, 6 x 9 = 54 and 7 x 8 = 56 . It was relatively common for subjects to confuse these two; a total of nine of the 54 total errors were confusions. However, notice that these two products have no common factors in the table, though both are even.

The fact that they are numerically very close seems to have been enough to produce a high frequency of confusions.

We now have, emerging, a picture of the typical subject having memorized and learned to generate a list of numbers that are the products or answers to the multiplication table. Given the factors he somehow uses them to search this list, trying to pick out the correct answer. The process takes an average of about one second.

One final idea that would relate the multiplication table to calculation would be the use of valid cues, about the factors, to choose (or eliminate) possible products. One simple rule is this; any two odd factors give an odd product, and if there is an even factor the product must be even. If the subject used this fairly-obvious rule, then he would make mainly odd errors when given two odd factors, and even errors otherwise. Table 4-1 shows the results of a tabulation.

Table 4-1

Frequency of Odd and Even Errors as a Function of Whether the Product Itself was Odd or Even

Subject's Error	Correct Product	
	ODD	EVEN
ODD	3	11
EVEN	16	24

As can be seen from Table 4-1, the subjects showed a strong preference for even over odd products as erroneous answers. This may reflect the fact that of the 10 possible wrong products, 7 are even and only 3 are odd. If the subject has been given a problem with an odd product, there are only two other odd products that can be errors, and 7 even products, so the probability (other things equal) of an odd product being given as an error should be 2 / (2 + 7) = .222 . In the Table, 3 of 19 or .158 of the errors were odd. If the subject has been given a problem with an even answer there are 6 even products that can be wrong, and 3 odd. The probability of an odd product being given as an error, other things equal, would be 3 / 9 = .333 . In the table it is shown that 11 of 35 , or .314 of the errors are odd.

These two calculations give an accurate account of the values in Table 4-1. To complete the argument, notice that a total of 19 out of 54 errors are made when the correct product is odd, that is, .352 . Since one-third of the true answers are odd, one might have predicted .33 of all errors would be made to odd products, if there were no bias, and the difference is slight.

In summary, errors to an "odd" problem have a tendency to be even errors, if there is no bias, simply because one of the three possible odd products is already used up as the correct answer. Errors on "even" problems have a corresponding though smaller tendency to be odd. The calculations given above agree very closely with such predicted values, considering the small number of observations. This, in turn, agrees with the hypothesis that the subjects, when trying to find an answer in the multiplication table, do not use the cue of "EVEN-ODD" in selecting the answer they will choose. There is no correlation between even-ness of the answer given and even-ness of the true product.

Conclusions

This study has shown, first, that clear patterns may emerge when we study performance on a highly-skilled mathematical task, even so simple a task as multiplying two small numbers, using the multiplication table.

Second, a set or determining tendency can be established over a sequence of 10 such problems, strong enough to cause some delay when a problem outside the set is introduced. The delay, from approximately .82 to 1.00 sec., was not very large but seemed detectable and probably significant.

The exact nature of this set remains somewhat unclear. After a series of trials on 7 x n, the subject is slowed only slightly by begin given n x 7, but is more slowed down by n x n. It appears, then, that a set to find 7 x n is very similar to, and compatible with, a set to find n x 7. The "multiplication table" is, of course, symmetrical because the operation of multiplication is commutative; a x b = b x a.

It is obvious that larger numbers require more time to calculate. So far as we have been able to discern, it is not the first, nor the second, nor the larger of the two factors that determines time, but both of the factors, i.e., the magnitude of the product.

One possible explanation of this fact would be that the subject unconsciously calculates the answer, by a process of successive adding or counting. Clearly, counting up to 72, even at a very high speed and unconsciously, should take longer than counting to 27.

Analysis of errors indicates that errors are not merely the results of stopping short or overrunning of a counting mechanism. With remarkable consistency subjects choose, as erroneous answers, numbers that appear in the multiplication table as products.

The errors did not consistently lie just above or just next to the correct answer in the table—that is, subjects did not show a very strong and consistent tendency to substitute 8 x 7 for 8 x 6, etc., for a common confusion was between 8 x 7 and 9 x 6.

Subjects showed no discernible tendency to choose EVEN errors when the true product was EVEN, so the ODD-EVEN property was apparently not used in selecting the answer from the table.

A process of estimation, if used, was only fairly successful, since 10 of the 54 errors were more than 10 in magnitude.

The results of this experiment are, therefore, somewhat negative, in that we have eliminated many possible mechanisms for the process of multiplying small numbers, without arriving at a very clear picture of the process actually used. We know that it is not a matter of a simple, random-access memory, since set and magnitude of the numbers, at least, have a large effect on time to respond. A set for 7 x n also worked almost as well for n x 7, indicating that commutativity of multiplication is used by the subjects in organizing the task. However, analysis of errors shows that subjects do not use the ODD-EVEN rule to select answers, but do pick most of their errors from values in the table.

One may picture the process approximately as follows. The subject in his childhood training on the multiplication table, has mastered a general skill in which he can generate any of the products, 40 , 42 , 45 , 48 , 49 , 54 , 56 , 63 , 64 , 72 , and 81 very rapidly when given a multiplication-table problem. His problem, therefore, is merely to select among these values when given the specific factors.

A difficulty facing the subject may be that the possible answers are spaced widely with many non-answers between. It may be relatively easy to find the answer when answers are more closely packed and there are fewer impossible values to be ignored. Inspection of Figure 4-5 shows that the answers to problems in the multiplication tables, 1--9, are concentrated at the lower values and become quite sparse above 50. This may be the source of the difficulty students have calculating the larger products. This hypothesis would take account of the fact that the speed seems to depend upon the product rather than either factor, yet errors are patterned as they are.

Nonetheless, it must be admitted that the picture of multiplying given by this experiment is very incomplete. The fact, emerging from the data and not forseen, is the great importance of what may be called the "response production" aspect of the process. The importance of producing responses, in ordinary paired-associates learning, has been emphasized especially by Underwood (for example, in the book by

Insert Fig. 4-5 about here

Underwood and Schulz). In the task of multiplying, at least when reciting the 1—9 table rapidly, it appears that students become especially ready to emit those numbers that are products in the table. By what mechanism these numbers are integrated and given strong readiness seems, to the writer, most obscure. If all of these numbers are available, then some mechanism is needed to suppress all but the correct response. This process worked successfully 99 percent of the time, in the present experiment, but we found that when subjects did err they often got neither factor right, and paid no detectable attention to the odd-even principle. This may merely mean that these mechanisms happened to slip, producing an error, and that they are usually sufficient to yield a correct answer.

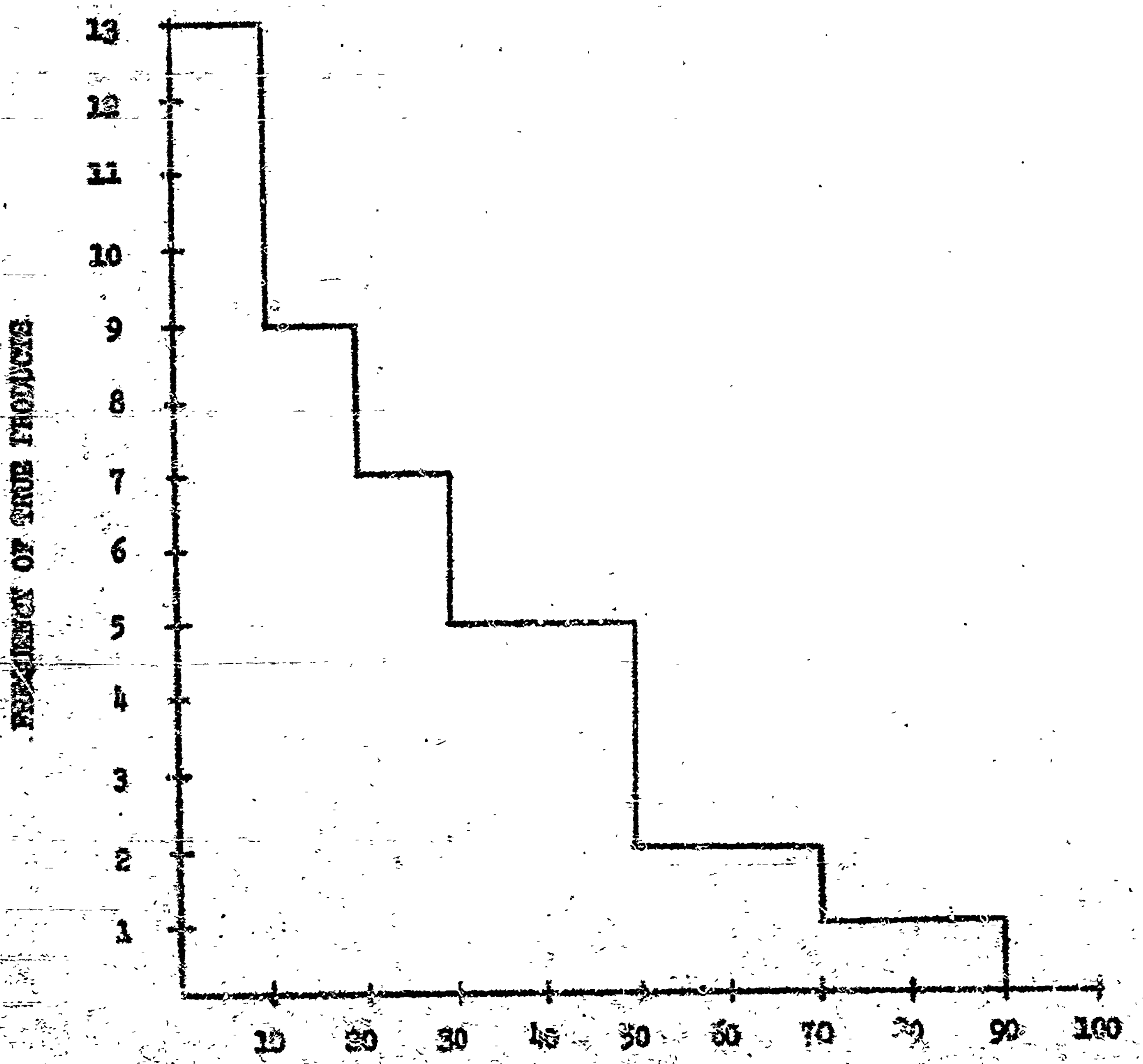

FREQUENCY OF TRUE PRODUCTS
1
2
3
4
5
6
7
8
9
10
11
12
13
10
20
30
40
50
60
70
80
90
100
NUMBERS (POSSIBLE PRODUCTS)

Fig. 4-7