DOCUMENT RESUME

ED 063 756 EM 009 931

AUTHOR	Gay, Lorraine R.
TITLE	Delay of Feedback and Its Effect on the Retention of Three Types of Computer-Assisted Mathematics Instruction.
INSTITUTION	Florida State Univ., Tallahassee. Computer-Assisted Instruction Center.
SPONS AGENCY	National Science Foundation, Washington, D.C.
PUB DATE	72
NOTE	23p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*College Mathematics; *Computer Assisted Instruction; *Feedback; Intermode Differences; Man Machine Systems; Reinforcement; Set Theory

ABSTRACT

Delay of feedback studies with animals or with tasks which are simple for human subjects have led to the common assertion that computer-assisted-instruction (CAI) systems should incorporate immediate feedback in order to maximize learning. However, studies using more complex learning tasks have suggested that delays of up to at least a day actually benefit learning. A study was designed to investigate the optimal feedback delay for different types of mathematical material in an undergraduate mathematics course taught by CAI. The results indicated that immediate feedback was not superior to delayed feedback, although the optimal delay is very much dependent on the type of material involved. The more complex the material to be learned is, the shorter is the feedback delay which can be tolerated without a decrement in retention. Thus, the implication of this study was that time-shared CAI systems need no longer be concerned about the resulting delays in input-output processes. (RH)

DELAY OF FEEDBACK AND ITS EFFECT ON THE RETENTION OF TWO TYPES OF COMPUTER-ASSISTED MATHEMATICS INSTRUCTION[1]

Lorraine R. Gay[2]

Florida State University

INTRODUCTION

Feedback interval has been defined as the length of time intervening between the subject's response to a learning or test time and the occurrence of an informative event that tells him whether his response is right, wrong, or somewhere in between (Brackbill, Wagner, and Wilson, 1964). One of the major issues in feedback research is now, and has been for quite some time, the question of the ideal interval between response and feedback. This problem has been thoroughly investigated in the field of animal learning; considerable research has shown conclusively that for sub-human organisms duration of feedback delay is directly related to decrease in learning efficiency (Rennner, 1964). With delays of but a few seconds, learning may not occur at all. Studies reporting the superiority of immediate knowledge or results over delayed feedback for human subjects have also appeared with some regularity during the past three decades. This research, however, has been concerned primarily with non-meaningful verbal materials. In the last few years, a number of studies have cast some doubt upon the principle of immediate knowledge of results for certain highly important types of learning in relatively mature human organisms (English and Kinzer, 1966).

Rats and men have yielded surprisingly similar results for a variety of simple learning and conditioning situations (Brackbill, Wagner, and Wilson, 1964); it is ludicrous, however, to expect parallel results for studies involving complex aspects of behavior. The human's response to feedback delay involves just such complexity. Sassenrath and Yonge (1968) have questioned the logic of inferring from a rat learning to press a bar or a pigeon learning to peck at a disk, to a child learning to speak his native language, or an adult learning matrix algebra. The implication is that the principle of reinforcement may not have the generality it has commonly been assumed to have. In addition, they point out that in animal studies of delayed reinforcement the animal has to remember his response over the delay interval and is then given only the reinforcement without being again presented with the original task with alternatives. In studies with humans, the subject may also have to remember his response over the delay period but when receiving feedback he is usually presented with the question or task as well as the alternatives originally available.

Evidence from studies on human motor skills first called into question the superiority of immediate feedback on learning; approximately eighty percent showed no significant differences in learning efficiency (Sassenrath and Yonge, 1968). Further, classroom applicability was limited by at least two factors. First, the majority of these studies involved either acquisition criteria or immediate retention criteria (More, 1969). The value of such criteria is limited because long-term rather than short-term retention is generally a primary instructional objective. Secondly, the results of many studies which did utilize long-term retention criteria were confounded by the amount of feedback given; for example, groups which received immediate feedback also received delayed feedback in the form of discussion along with the delayed feedback groups.

In recent studies, the principle of immediacy of feedback has been further opened to question (Brackbill, Wagner, and Wilson, 1964; English and Kinzer, 1966; Sassenrath and Yonge, 1968). These studies have found a greater amount of retention by students for whom feedback had been delayed than by students who received immediate feedback. The realization has grown that the complete acceptance, without qualification, of delayed reinforcement or information feedback generalizations, which suggest that a few seconds delay between response and reinforcement may mean the difference between maximal learning and no learning whatever, is unwarranted in human learning.

It is also clear, that the ability of a teaching machine to give immediate knowledge of results may not be a valid reason for preferring automated instruction to traditional instruction. The above discussion has major implications for computer-assisted instruction (CAI). Under CAI instruction, the student finds out immediately upon responding whether his response was correct (in contrast to traditional instruction which typically involves delays of one or more days). The provision of immediate feedback has long been presented as a rationale for the use of CAI.

The most obvious question that arises is why such an erroneous conclusion has been perpetuated through many hundreds of teaching machine studies. One reason is that virtually all of the research has involved immediate retention of non-meaningful materials (Brackbill and Kappy, 1962). As Brackbill and Kappy have pointed out, although most psychological variables that affect both learning and retention have directionally correlated effects on the two processes, there are exceptions, e.g., partial reinforcement. Also, at least part of the answer lies in the fact that the teaching machine makes simultaneous use of several learning conditions (e.g., active rather than passive responding). Whenever two or more variables are allowed to operate simultaneously in any situation, the outcome of that situation cannot be pinpointed to the influence of any one of those variables because their separate contributions to the outcome are inextricably confounded (Brackbill, Wagner, and Wilson, 1964). Obviously, the way to increase the efficiency of CAI - or any other complex learning situation - is to systematically study the separate contribution of each of the identifiable variables. Only by weeding out the ineffective variables, such as immediate feedback, can the complex be improved.

Review of the Literature

The literature on delay of feedback contains investigation of two dimensions of the phenomenon: feedback received during a learning situation (acquisition feedback) and feedback received following a testing situation (test feedback). Acquisition feedback studies have generally utilized discrimination and paired-associate learning tasks. Test feedback studies have typically involved meaningful material and multiple choice evaluation techniques.

Acquisition Feedback. Most of the research on this dimension has been conducted by Brackbill and her associates with feedback delays up to 30 seconds. As their studies have involved discrimination and paired-associate learning tasks, the index of retention has typically been relearning scores, or trials to reach criterion. The results of two studies with third grade males on a drawing discrimination task favored delaying feedback. These studies used delays of 0.5 and 10.0 seconds (Brackbill and Kappy, 1962) and 0.0, 2.5, 5.0, 10.0, 20.0, and 30.0 seconds (Brackbill, Adams, and Reaney, 1967). In every case, the longer feedback delay improved retention when measured one day after original learning; no consistent results, however, were found when retention was measured eight days following original learning.

Another pair of studies also involved the discrimination behavior of third grade males but involved more difficult tasks (Brackbill, Bravos, and Starr, 1962; Brackbill, Wagner and Wilson, 1964). Both studies used feedback delays of 0.0 and 10.0 seconds. When the task involved drawings, a 10.0 second delay again improved retention as measured 1 day later; for the learning of English-French equivalents, there was no difference. In both cases, however, the 10.0 second delay resulted in greater delayed retention as measured seven or eight days following original learning.

In 1966, in a series of studies using undergraduate males, Lintz and Brackbill found the delay of feedback effect to be dependent on the nature of the task. Using a 0.0 or 10.0 second delay in a discrimination task, there were no retention differences either 1 or 7 days after original learning; for a paired-associate task, however, a 10.0 second delay resulted in greater retention both 1 day and 7 days after original learning.

Based on the above findings, the following conclusions seem warranted for acquisition feedback:

1. For a task of medium difficulty, the short-term retention of young males is improved by feedback delays of up to 30 seconds;

2. For a task of greater difficulty, the long-term retention of young males is improved by feedback delays of at least 10.0 seconds;

3. For a discrimination task, the retention of male undergraduates is unaffected by a 10.0 second feedback delay; and

4. For a paired-associate task, the retention of undergraduate males is improved by a 10.0 second delay.

Test Feedback. The results are more consistent for multiple choice testing situations. Typically in such studies, the student responds to multiple choice questions receiving feedback either immediately after responding or after a pre-determined time interval has elapsed. Reported feedback delays range from 1 hour to 1 week (1 hour, 2.5 hours, 1 day, 2 days, 4 days, 1 week). While some studies have investigated completeness of feedback, this report will be concerned only with feedback situations in which the item stem and all alternatives are re-presented along with the indication of the correct alternative. Also, as the current study was concerned primarily with delayed retention, the results of previous research will be discussed with respect to their delayed retention findings.

With the exception of one study (More, 1969), the subjects for the studies of delayed testing-feedback have been college undergraduates. The subject matter involved has spanned a range of curriculum areas from introductory psychology to statistics, and retention has been measured either three days, five days, or one week following original learning. In every case, immediate feedback has been found to be inferior to feedback delayed for 1 hour, 2.5 hours, 1 day or 2 days. Prolonged feedback delays of 4 days or 1 week, however, have been found to be inferior to delays of 1 hour, 2.5 hours, 1 day, or 2 days. These results suggest the existence of a non-linear relationship between feedback delay and retention.

Given such a relationship, a delay of 2.5 hours, for example, would be expected to result in higher retention than a delay of 0.0 seconds, whereas a delay of 4 days would be expected to result in lower retention than a delay of 1 day. In other words, we would expect that increasing the feedback delay would correspondingly increase retention (up to some ideal delay period), and that further delay would begin to decrease retention until a retention level equal to that resulting from immediate feedback were reached. Examination of the results of previous studies supports this hypothesis:

Feedback Delay Results	Author(s)
0.0 < 1 hour, 2 days > 1 week	English & Kinzer (1966)
0.0 < 2.5 hours, 1 day > 4 days	More (1969)
0.0 < 1 day	Sassenrath & Yonge (1968)
0.0 < 1 day	Sturges (1969)

It appears that delayed testing-feedback actually improves retention of the learned material while both immediate feedback and overly-prolonged feedback delays impair retention.

The purpose of the present study was 1) to systematically investigate the feedback delay variable, and 2) to extend the research on feedback delay to mathematical learning. In a CAI learning task, three types of mathematical material were investigated: definitions, algorithms and proofs. The four feedback conditions were as follows:

1. Immediate feedback (IF) - students received feedback immediately following a within-program response (WPR);

2. 30 second feedback (30S) - students received feedback 30 seconds after making a WPR;

3. End of session feedback (ES) - students received feedback on all WPR's at the end of each day's session; and

4. 24 hour feedback (24H) - students received feedback on all WPR's made during one session 24 hours later at the beginning of the next session.

Based on previous findings with college students, the following hypotheses were formulated:

1. For definitions and algorithms, 30S, ES, and 24H feedback will result in greater retention than IF. Also, the means for students in the 24H condition will be systematically higher than the means for students in the ES and 30S feedback conditions; similarly, the means for the ES condition will be systematically higher than the means for the 30S condition.

2. Due to the sequential dependencies involved in a mathematical proof, the IF and 30S conditions will result in greater retention than the ES and 24H conditions.

METHOD

Subjects

Subjects for this study were lower-division undergraduate volunteers from two institutions: Florida State University (n = 40), and Tallahassee Community College (n = 40). The subjects came from a variety of curriculum areas and all received an honorarium for their participation in the study.

Learning Materials

This study utilized portions of CAI materials previously described in detail (Love, 1969), which were adapted for the present study. The instructional program was composed of approximately four hours of instruction (20 lessons) on Boolean algebra, primarily set theory. The material included definitions, algorithms, and proofs. The content of the lessons was as follows:

Lesson 1 - Introduction to Set Theory
Lesson 2 - Venn Diagrams With Two Sets
Lesson 3 - Venn Diagrams With Three Sets
Lesson 4 - Special Cases
Lesson 5 - The NOT Operation
Lesson 6 - Determining A Set Given Any Expression
Lesson 7 - Boolean Algebra Introduction
Lesson 8 - The Commutative Laws
Lesson 9 - The Associative Laws
Lesson 10 - The <u>AND</u> Distributive Law
Lesson 11 - The <u>OR</u> Distributive Law
Lesson 12 - The Universal Element
Lesson 13 - The Null Element
Lesson 14 - Complements
Lesson 15 - Proof of Theorem - Universal <u>OR</u>
Lesson 16 - Proof of Theorem - Null <u>AND</u>
Lesson 17 - Proof of Theorem - <u>OR</u> Indempotent
Lesson 18 - Proof of Theorem - <u>AND</u> Indempotent
Lesson 19 - Proof of Theorem - <u>OR</u> Absorption
Lesson 20 - Proof of Theorem - <u>OR</u> Identity
(See Appendix A for a Content Outline)

Within-program problems were presented in a multiple choice format.

Apparatus

The learning materials were presented by an IBM 1500 CAI system. Terminals for the system consist of a cathode-ray tube, a light pen, and a keyboard. The terminals were located in an air-conditioned, sound-deadened room. The CAI system administered the instruction and recorded the students' responses.

Experimental Design

All students in all groups were pretested, received their instruction on days 1-4, and were tested for delayed retention fourteen days from the first day of original learning (See Figure 1). All students received the same instruction and the same within-program testing frames; the groups differed only as to when they received their feedback for within-program testing frames and with respect to occurrence of quizzes.

Experimental Procedures

This study was conducted over a six-month period. Each week between 12 and 16 students were processed from one of the 2 participating institutions. These students were randomly assigned to one of the four feedback conditions. Of the forty students from each institution, 10 were assigned to the IF condition, 10 were assigned to the 30S condition, 10 to the ES condition, and 10 to the 24H condition. For all students, the procedure was the same:

1. At a time convenient to the student, he was administered a 100-item multiple choice pretest convering all aspects of the instructional program; the pretest was a re-ordered version of the delayed retention (DR) test;

DAY

Group	0	1	2	3	4	5 - - - - - 15	
IF	X_0	$(X\text{-}F)_1Q_1$	$(X\text{-}F)_2Q_2$	$(X\text{-}F)_3Q_3$	$(X\text{-}F)_4Q_4$		X_5
30S	X_0	$(X\text{-}30\text{-}F)_1Q_1$	$(X\text{-}30\text{-}F)_2Q_2$	$(X\text{-}30\text{-}F)_3Q_3$	$(X\text{-}30\text{-}F)_4Q_4$		X_5
ES	X_0	$X_1(FX_1)Q_1$	$X_2(FX_2)Q_2$	$X_3(FX_3)Q_3$	$X_4(FX_4)Q_4$		X_5
24H	X_0	X_1	$(FX_1)Q_1X_2$	$(FX_2)Q_2X_3$	$(FX_3)Q_3X_4$	$(FX_4)Q_4$	X_5

X_0 = Pretest

X_1 = Lessons 1-5

X_2 = Lessons 6-10

X_3 = Lessons 11-14

X_4 = Lessons 15-20

X_5 = Delayed Retention Testing

Q_1 = Quiz 1 (Lessons 1-5)

Q_2 = Quiz 2 (Lessons 6-10)

Q_3 = Quiz 3 (Lessons 11-14)

Q_4 = Quiz 4 (Lessons 15-20)

(X-F)= **Feedback immediately following WPR**

(X-30-F)=Feedback 30 seconds following WPR

X(FX)= **Feedback for all WPR's at conclusion of lessons**

(FX)QX= **Feedback for previous day's WPR's, Quiz for previous day, new WPR's**

Figure 1.--Experimental Design for the Four Feedback Conditions

2. On the first day of instruction, the student worked through lessons 1-5; each lesson consisted of instruction, illustrative diagrams, and multiple choice test questions;

3. On the second day of instruction, the student worked through lessons 6-10;

4. On the third day of instruction, the student worked through lessons 11-14;

5. On the fourth day of instruction, the student worked through lessons 15-20;

6. On the fifteenth day, the student was administered a 100-item DR test.

During the instruction, students in the IF group received their feedback immediately following each WPR; the feedback was of the form "correct, the answer is ___", or, "incorrect, the answer is ___". They were administered a quiz at the conclusion of each daily session. Each quiz was an alternate form of a portion of the DR test; the four quizzes had a total of 100 items. Students in the 30S group received feedback 30 seconds after making a WPR and were administered a quiz at the conclusion of each session. Students in the ES group did not receive feedback on their WPR's until the end of the session; at that time each question and its alternatives were presented. For each question, the WPR that the student had chosen was indicated. At the conclusion of the feedback phase, the student was administered the quiz for that session. Students in the 24H group also did not receive feedback on their WPR's during the instruction. The following day, at the beginning of the next session, they received feedback for the previous day's lessons as described for the ES group. Following feedback, they were administered the quiz for the previous day's lessons.

The pretest, all quizzes, and the DR test were administered via paper and pencil; they were not part of the CAI instruction.

RESULTS

Reliability

Reliability was computed for each institution to insure that the criterion test was equally reliable for the two student populations. The criterion test consisted of one item for each concept included in the learning materials and therefore contained three sub-strata corresponding to the three types of mathematical learning investigated, definitions, algorithms and proofs. The KR-20 reliability was computed for each sub-stratum as well as for the total test (See Table I).

TABLE I

Reliability of Criterion Test, by Type of Material and for Total Test

Institution	Reliability			
	Definitions[a]	Algorithms[b]	Proofs[c]	Total
Florida State University	.87	.90	.96	.96
Tallahassee Junior College	.81	.88	.93	.95

[a]Number of items = 23
[b]Number of items = 50
[c]Number of items = 27

As Table I indicates, reliabilities were equally high for both student populations. Consequently, data for the two groups were combined for analysis.

Retention

The two factors of interest were feedback delay and type of material. The ideal analysis design would be a 4 x 3 analysis of covariance with repeated measures on the second factor (covarying for pretest differences). However, the scores for the three types of material were not based on the same number of items, there were 23 definitional items, 50 algorithm items and 27 proof items. In Order to equate the scores, it would be necessary to either covert them to proportions or to apply a transformation. Neither of these approaches is desirable, as they affect both the reliability and the generalizability of outcomes.

As no acceptable analysis procedure was available, the adjusted means[3] for each group, by type of material, were computed and converted to percents (See Figure 2). While no probability levels can be attached to such statements, the following observations are offered:

1. There appears to be an interaction between feedback delay interval and type of material.
2. Up to 12% more definitions were retained by students in the 30S group, than by students in the other three conditions.
3. Students in the 24H group retained up to 23% fewer proof items than students in the other three conditions.
4. Students in each of the four feedback groups retained algorithms to approximately the same degree.

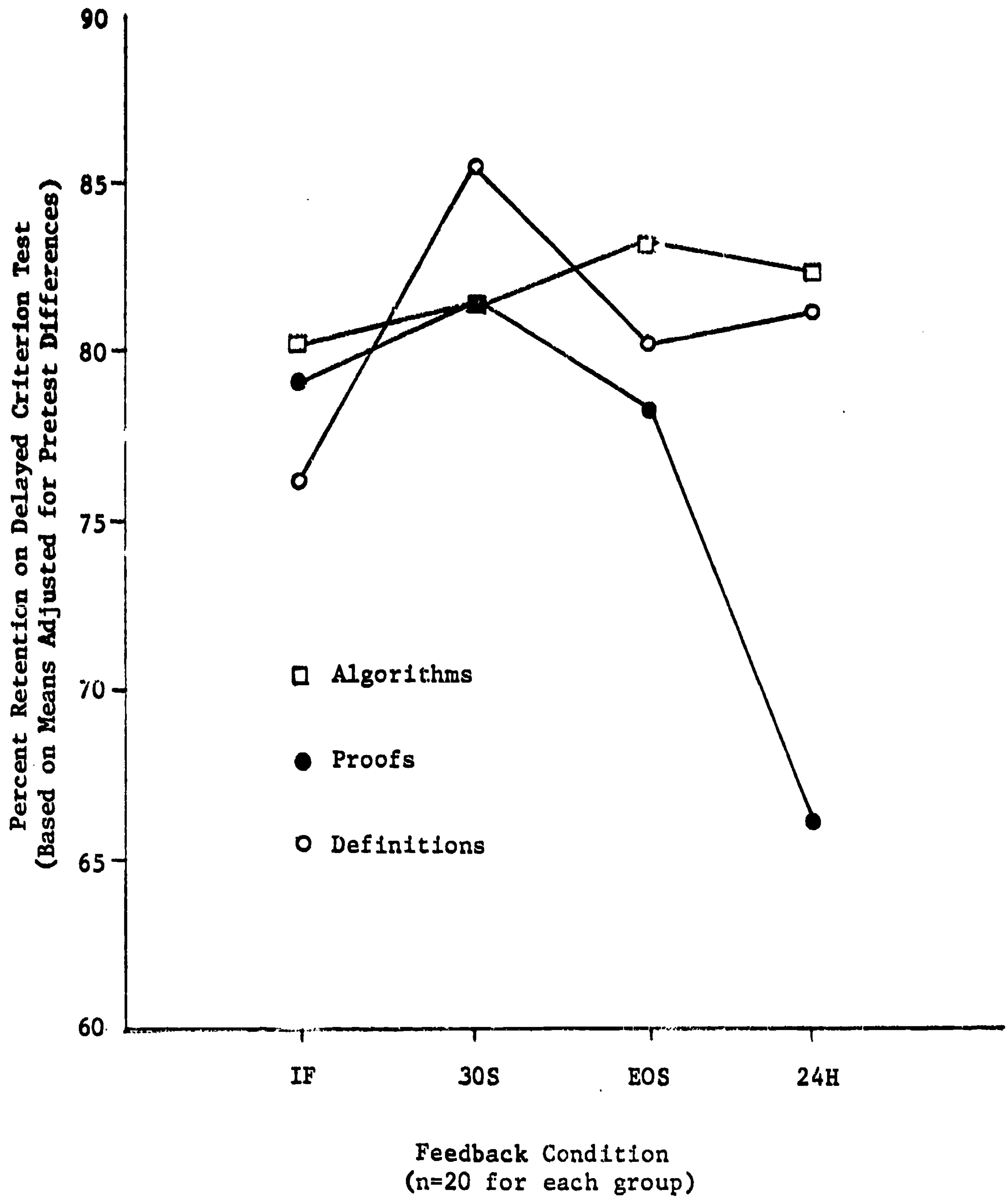

Fig. 2.--Percent retention for each feedback condition by type of material.

DISCUSSION

Past research on delay of feedback has been exclusively involved with a variety of verbal learning tasks. This study investigated the phenomenon for three types of mathematical learning, definitions, algorithms and proofs. Results indicate that the optimal feedback delay is very much dependent on the nature of the learning material, although they do support previous findings with regard to the non-superiority of immediate feedback.

Verbal learning research has generally found a one-day feedback delay to be optimal and delays of four or more days to be harmful; this study on mathematical learning has found that while a delay of thirty seconds may be beneficial, a one-day delay is equally effective at best, and may have a seriously debilitating effect on retention. In the present study, the definitional material was most like the learning material in past research and it was the only type of material for which a delay of feedback appeared to facilitate retention. The instruction on proofs contained the material most unlike that utilized in past research; correspondingly a one day feedback delay adversely affected retention of this material by as much as 23%. In conclusion, it appears that the more complex the material is, the shorter is the feedback delay which is defined as being "overly-prolonged."

The implication of this study for computer-assisted mathematical instruction is that immediate feedback is neither necessary nor desirable; a slight delay may even facilitate retention. One of the problems generally associated with simultaneous operation of terminals by many users, is the accompanying delay in input-output processes. This study supports previous research in the contention that this need not be a consideration.

REFERENCES

Boersma, F. J. Effects of delay of information feedback and length of post-feedback interval on linear programmed learning. Journal of Educational Psychology, 1966, 57(3), 140-145.

Brackbill, Y., Adams, G., & Reaney, T. P. A parametric study of the delay-retention effect. Psychological Reports, 1967, 20, 433-434.

Brackbill, Y., Bravos, A., & Starr, R. H. Delay-improved retention of a difficult task. Journal of Comparative and Physiological Psychology, 1962, 55(6), 947-952.

Brackbill, Y., & Kappy, M. S. Delay of reinforcement and retention. Journal of Comparative and Physiological Psychology, 1962, 55(1), 14-18.

Brackbill, Y., Wagner, J. E., & Wilson, D. Feedback delay and the teaching machine. Psychology in the Schools, 1964, 1, 148-156.

English, R. A., & Kinzer, J. R. The effect of immediate and delayed feedback on retention of subject matter. Psychology in the Schools, 1966, 3, 143-147.

Lintz, L. M., & Brackbill, Y. Effects of reinforcement delay during learning on the retenti n of verbal material in adults. Journal of Experimental Psychology, 1966, 71(2), 194-199.

Love, W. Individual versus paired learning of an abstract algebra presented by computer-assisted instruction. Technical Report No. 5, Computer-Assisted Instruction Center, Florida State University, July, 1969.

More, A. J. Delay of feedback and the acquisition and retention of verbal materials in the classroom. Journal of Educational Psychology, 1969, 60(!', 339-342.

Renner, K. E. Delay of reinforcement: A historical review. Psychological Bulletin, 1964, 61(5), 341-361.

Sassenrath, J. M., & Yonge, G. D. Delayed information feedback, feedback cues, retention set, and delayed retention. Journal of Educational Psychology, 1968, 59, 69-73.

Sassenrath, J. M., & Yonge, G. D. Effects of delayed information feedback and feedback cues in learning on delayed retention. Journal of Educational Psychology, 1969, 60(3), 174-177.

Sturges, P. T. Verbal retention as a function of the informativeness and delay of informative feedback. Journal of Educational Pscyhology, 1969, 60(1), 11-14.

APPENDIX A

CONTENT OUTLINE OF CAI INSTRUCTION

Lesson 1 - Introduction to Set Theory

Definition of a Set

Complements

Universal Element

Null Element

Lesson 2 - Venn Diagrams With Two Sets

The AND Operation on Sets

The OR Operation on Sets

Lesson 3 - Venn Diagrams With Three Sets

The AND Operation

The OR Operation

Lesson 4 - Special Cases

Combining AND and Ø

Combining OR and Ø

Combining AND and I

Combining OR and I

Combining AND and Self, e.g. $A \cdot A = A$

Combining OR and Self, e.g. $A + A = A$

The AND Complement

The OR Complement

Lesson 5 -- The Not Operation

Lesson 6 - Determining a Set Given Any Expression

Equivalent Expressions - DeMorgan's Laws

Lesson 7 - Boolean Algebra Introdution

Variables

The Universal, I

The Null, Ø

AND, OR, and NOT

Introduction to the Laws of Boolean Algebra

The OR Commutative Law
The AND Commutative Law
The OR Associative Law
The AND Associative Law
The OR Distributive Law
The AND Distributive Law
The Universal Property
The Null Property
The AND Complement
The OR Complement

Lesson 8 - The Commutative Laws

The OR Commutative Law

The AND Commutative Law

Lesson 9 - The Associative Laws

The OR Associative Law

The AND Associative Law

Application of the Commutative and Associative Laws

Lesson 10 - The AND Distributive Law

Lesson 11 - The OR Distritutive Law

Lesson 12 - The Universal Element

The Universal Element

The Universal Property

Lesson 13 - The Null Element

The Null Element

The Null Property

FOOTNOTES

[1]This research was supported in part by a grant to Duncan N. Hansen, Computer-Assisted Instruction Center, from the National Science Foundation (GJ-623).

[2]Now at the Board of Regents, Division of Academic Affairs, State University System of Florida, Tallahassee, Florida.

[3]Means were adjusted using an analysis of covariance.