Improving Conceptual Performance with Programmed Instruction

Daniel G. Reese
Western Michigan University
IMPROVING CONCEPTUAL PERFORMANCE WITH PROGRAMMED INSTRUCTION

by

Daniel G. Reese

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
August 1972
ACKNOWLEDGMENTS

To my committee members, Dr. Jack Michael and Dr. Robert Hawkins, I would like to express my sincere appreciation for their time, suggestions, and criticisms during the course of this study. I would like to especially thank my committee advisor, Dr. Richard Malott, for his extensive efforts, guidance, and encouragement. Finally, I am indebted to the Psychology Department for the assistantship, which made my graduate studies possible.

Daniel G. Reese
PLEASE NOTE:

Some pages may have
indistinct print.
Filmed as received.

University Microfilms, A Xerox Education Company
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INDEX OF TABLES.</td>
<td>iv</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Instructional Problems in Education</td>
<td>3</td>
</tr>
<tr>
<td>Research and Development</td>
<td>4</td>
</tr>
<tr>
<td>Solutions to the Instructional Problems of Education</td>
<td>6</td>
</tr>
<tr>
<td>Conceptual Learning</td>
<td>12</td>
</tr>
<tr>
<td>Conceptual Learning from Programmed Instruction</td>
<td>13</td>
</tr>
<tr>
<td>The Problem</td>
<td>14</td>
</tr>
<tr>
<td>The Concepts</td>
<td>15</td>
</tr>
<tr>
<td>DEVELOPMENT AND EVALUATION OF THE PROGRAMS</td>
<td>20</td>
</tr>
<tr>
<td>Program I: Discriminated Extinction</td>
<td>20</td>
</tr>
<tr>
<td>Program II: Conditioning of Behavior with Positive and Negative Reinforcement</td>
<td>29</td>
</tr>
<tr>
<td>Program III: Extinction of Behavior Maintained by Positive and Negative Reinforcement</td>
<td>34</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>38</td>
</tr>
<tr>
<td>CONCLUSIONS</td>
<td>41</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>42</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>43</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
INDEX OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The teaching sequence of elementary operant concepts</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>Final results, program I</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>Final results, program II</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>Final results, program III</td>
<td>39</td>
</tr>
</tbody>
</table>
Education is concerned with the establishment of specific stimulus-response repertoires, as well as the acquisition of more general conceptual skills. These goals can be achieved in many ways, however some methods are obviously more effective than others. The optimum learning system is one that maximizes learning, and minimizes student and administration cost, time, and labor (J. Michael, personal communication, July 26, 1971).

Until recently an empirically developed teaching technology based on learning theory has not been available to educators. As a result, many educators have proceeded haphazardly, or have perpetuated traditional teaching methods. Behavioral psychology has recently provided educators with laws of behavior relevant to education, with implications for improvements in existing methods.

Any successful learning situation has three necessary components: a program of tasks, differential consequence for participation and differential consequence for correctness (J. Michael, personal communication, June 23, 1971). The program of tasks must begin at the student's educational level (Malott, 1971). The student must have the necessary prerequisite skills in his repertoire before he can successfully initiate a given task. If this is overlooked, a student could fail to successfully complete a program of instructional tasks because he lacked prerequisite skills, rather than because the program was deficient. For example, a student may have difficulty reading a text if he starts in the middle, whereas he would have little or no difficulty if he started the book at
the beginning. The program of tasks must also progress toward the terminal behavior in small steps, and each step must receive some form of differential consequence if the program is to be maximally efficient (Malott, 1971). Finally, when the material to be learned is sequential, the program of tasks must proceed one step at a time, to insure that the students have the necessary prerequisite skills for more advanced tasks (Malott, 1971). If the material is sequential, and requires mastery of earlier lessons before the student can understand subsequent material, the teacher must require mastery of the current material before a student is allowed to proceed (Michael, 1969). The requirements of small steps, differential consequence for each step, and content mastery are logical assumptions which are commonly made. Effective instructional materials can exist without these properties, however the inclusion of these in the instructional process seems to aide the student.

The existence of instructional tasks alone, is not enough for learning to occur. Once the student's prerequisite skills have been determined, the instructional tasks arranged in small steps, and a content mastery criterion established, it is necessary to provide a source of differential consequence for the student's participation in the instructional tasks\(^1\) (Michael, 1969). In theory, educators seem to prefer a motivational system based on the positive control of students' behavior. In a system of positive

\(^1\)Differential consequence is the selective presentation, removal, or withholding of stimulus events (i.e., reinforcers, punishers) following the response to be affected.
control, a student's academic behavior is reinforced, and this results in a strengthening of a class of approach responses. In practice however, educators seem to use aversive control, which results in the strengthening of a class of avoidance responses. Possibly a better system of control would consist of a combination of positive and aversive control, reinforcing participation in the program of tasks, and punishing non-participation.

Finally, educators must provide for the differential consequence of correctness in the execution of the instructional tasks (Michael, 1969). It is not enough to consequeate participation in a series of well designed tasks; there has to be some payoff for the student to do well at those tasks. As with the consequence for participation, a combination of differential reinforcement and punishment is probably the most efficient system of control. Reinforcers should be available to the student who does well at the tasks, and punishers should be presented to the student who does not do well.

**Instructional Problems in Education**

Problems in achieving an efficient educational system are generally due to deficiencies in one or more of the above areas (i.e., program of tasks, differential consequence for participation, differential consequence for correctness).

The solutions to these three sources of difficulty are not simple, and generally require a great deal of restructuring in the educational environment. The exact form the environment should take
varies from one situation to the next. What works in one situation may or may not work in another. This problem of design can only be dealt with through systematic research and development of the instructional tasks, and the consequences and contingencies for correctly participating in these tasks.

Research and Development

Malott (1971) presents six phases to systems analysis that are adaptable to the problem of research and development. The first step in this process of research and development is an analysis of the existing system in terms of the necessary components for a successful learning situation. Perhaps the present system is adequate on several dimensions, or possibly on all dimensions. Does the program of tasks start at the student's present educational level, proceed in small steps, and require mastery of each step before the student is allowed to continue? Is the student's behavior affected by the contingencies in effect? Is the correctness of the student's response to the program differentially consequted?

Once the present system is analyzed, the next step is to state the terminal objectives of the system. These objectives should specify observable behavior, be attainable, and be functional. It is necessary to work with observable behavior for two reasons. First, it is necessary to apply the contingencies consistently. This is only possible if the behavior to be consequted can be

2A contingency is a specification of the conditions in which specific consequences occur.
observed in some way. Secondly, the contingencies and consequences used can only be evaluated in terms of their effect on the behavior. If the behavior is not observable, how can the contingencies and consequences be evaluated?

The objectives must be attainable. If too much is expected of the student, he will not succeed regardless of the remaining features of the system. The objectives must also be functional. This is important because the behavior generated in the instructional setting will extinguish outside of the classroom, unless some source of reinforcement is available. Functional behaviors established in the classroom will occur in the natural environment, and they will be reinforced due to their functionality.

The third step in research and development is to design the system. This entails the specification of the behaviors, consequences, and contingencies to be used in the system. This system should be designed to accomplish the behavioral objectives developed in step two. The behavior, consequences and contingencies specified should consider the present educational level of the student, require small steps, and require mastery of the materials before advancement. The contingencies and consequences should reinforce participation and correctness, and should punish non-participation and incorrectness.

The fourth step is to implement the system. This involves the observation of the behaviors, and the application of the contingencies and consequences specified in step three.

Extinction is the decrease in the rate of a conditioned response due to the absence of reinforcement.
Once the system is operating, it must be evaluated. It is necessary to determine if the procedures are being applied consistently. Are the contingencies and consequences being applied consistently, and as specified? If so, it is necessary to evaluate the behaviors previously specified. Are these behaviors affected by the contingencies and consequences in effect? Finally, determine if the system is accomplishing the objectives from step two.

If the evaluation in step five reveals any problem (i.e., the system is not doing what it was designed to do), then move on to step six. Step six is recycling through all of the previous steps. This recycling should be done as many times as is necessary, to achieve the objectives.

Solutions to the Instructional Problems of Education

Obviously there are different solutions to the problems of an efficient educational system. This first problem encountered is one of adequate source materials (Michael, 1969). Often the existing texts are poorly sequenced, contain irrelevant material, and are missing important items. Lectures and demonstrations are frequently less adequate than texts. One solution is to rewrite the texts. Although this is feasible, many existing texts can be more than adequate for teaching given repertoires if they are supplemented with reading objectives. Objectives aid the student in using difficult, poorly sequenced, and irrelevant material. They show him what is important, and often clarify difficult or obscure points. Lectures and demonstrations may also be supplemented...
by objectives. The objectives help the student identify the relevant points of a lecture or demonstration.\(^1\)

The next problem is one of motivation, or, how to get the student to participate (Michael, 1969). Possible solutions to this problem consist of arranging differential consequences for participation. Many different consequences may be used. Grades, class promotions, and degrees can be used effectively; or deliberately arranged extrinsic reinforcers, such as edibles, tangible items, and activities may also be satisfactory. The present educational system appears to rely on the possibility of parents, teachers, peers, or employers evaluating the student at some time in the future. This system generates little behavior and is considered sub-optimal. An optimal system would use immediate consequences such as those mentioned earlier, in a consistent specified manner, to maintain student participation.

The final problem considered here is the source of consequation for correctness, while responding to the instructional material. Again, objectives aid in this area. The student's reading response is reinforced if he can respond to the objectives after reading the relevant section in the text. The inability to respond to the objectives punishes reading errors, because the student should then reread the particular section of the text associated with the difficult objective. Another source of reinforcement for correctness

\(^1\)The general notions about the use of objectives to overcome the problem of inadequate source materials, were from an educational technology course at Western Michigan University, offered by Jack Michael in the spring of 1971.
is simply the knowledge of the correct answer. Through socialization, success is generally established as a reinforcer for most individuals. Unfortunately, success alone may not be a sufficiently powerful consequence to control an adequate amount of behavior. Due to this inadequacy, other sources of differential consequence for correctness should be utilized.

There is also some overlap concerning differential consequence for participation and correctness. Participation can be conseuated indirectly by conseuating correctness. To make a correct response, it is usually necessary to participate. Therefore, reinforcement for correct responses may also reinforce participation, due to the temporal and functional association involved.

Another possible solution to the above problems (i.e., adequate source material, consequence for participation, consequence for correctness) is programmed instruction. Empirically validated programmed instruction appears to directly overcome at least two of the three major problems in education—adequate source material and consequence for correctness.

Programs, as information sources, are not necessarily better than texts, however they usually have some advantages. Programs proceed in small steps; each step may require active responding and each response may be differentially conseuated (Pipe, 1966).

There are many different programming techniques, each with its advantages and disadvantages (Espich & Williams, 1967). A discrimination frame sequence consists of three different types of frames. The first type of frame in the sequence requires the
student to make a discrimination between various $S^D$'s and $S^A$'s based on the prompts supplied. After the student receives corrective feedback (consequation for correctness) he proceeds to the next frame. The second type also requires a discrimination, however no prompts are supplied but consequation occurs via corrective feedback. Finally in the last type of frame, the student is asked to define and give examples of the concept which he has been discriminating. Correctness is then consequated in one of many ways.

Espich and Williams (1967) report that this type of programming technique is especially useful when teaching concepts, although they present no data.

Another programming technique is retrogressive chaining. This involves a backward chaining procedure to establish a stimulus-response chain. Initially, all but the last response of the chain are supplied in the frame. As the student progresses through the program, he is gradually required to supply more of the responses in the chain. Finally by the end of the program, the student is making all of the responses in the chain. Corrective feedback is supplied after each frame (Espich & Williams, 1967).

A final programming technique considered here is adjunct programming. The purpose of this technique is to allow students to learn as efficiently as possible from an existing text. This technique can take two different forms. First, the program and text can be used independently. The program is read either before or
after the text; it clarifies and expands on the text. Another way to use adjunct programming is to present sections of the book verbatim in the program. This allows the book and the program to be read simultaneously. Adjunct programming is very similar to the use of objectives with a standard text, however corrective feedback is often more complete with the program (Espich & Williams, 1967).

All of the programming techniques mentioned above provide differential consequence for correctness. This feedback follows each frame in which a response (other than reading) was required. Reading is probably conseqeated by the ability to respond to the questions, and responding to the questions is consequence by the subsequent corrective feedback.

The programming techniques above do not directly solve the problems of inadequate source material and consequence for participation. The consequence for participation still presents the same problems as mentioned earlier, and the solution to this problem is no more apparent. However the problem of adequate source material can be solved by the empirical validation of the program. Programs that are developed following the research and development guidelines presented earlier, can overcome many of the deficiencies of non-validated materials. By testing, revision, more testing, etc., many of the inadequacies of the program can be overcome.

How do programmers know which technique to use? Sometimes the subject matter or the specific behavior being taught is more
amenable to one technique contrasted with another. One such
criterion for programming techniques might be what Espich and
Williams (1967) label the levels of learning.

The first level is "exposure." This is information that is
often presented for enrichment. It is not very functional, but
teachers often express the notion that students should be exposed
to it. Most programs would not teach this subject matter or at this
level.

The second level is "recognition." Learning at this level
requires only that students make very broad discriminations
(e.g., between correct and incorrect statements). This type of
learning still lacks a great deal of functionality, and few programs
are directed toward this level.

The third level is "recall." This enables students to repeat
what they have learned in their own words. At this level the learn-
ing is becoming functional, as it may be of much greater use to
the student.

The fourth level is "memory" and is similar to the recall level,
except that the student is now able to repeat information verbatim.
It is questionable whether memory or recall is a higher level of
learning; unfortunately the authors failed to explain their
rationale behind this hierarchy.

The final level is conceptual learning. It provides the
student with the behavioral repertoire necessary to apply his know-
ledge. He can generalize or make fine discriminations; he can solve
problems; and he can generate novel examples for the concept he has learned. The primary advantage of conceptual learning over lower levels of learning, is that it provides the student with the skills necessary to respond correctly in novel situations (Whaley & Malott, 1971; Psych 150 and Friends, 1971). If a student has learned a concept, he can respond correctly to an instance of that concept, even if he has had no prior contact with that specific instance. This is possible because conceptual behavior is correctly discriminating the presence of all of the defining characteristics of a given concept. Thus, conceptual learning reduces the probability that a student will be handicapped by stimulus variations found outside the initial training situation (Whaley & Malott, 1971).

Conceptual Learning

Englemann defines a concept as "the set of characteristics shared by a set of instances in a given universe of concepts, and not shared by other instances in that universe." A concept has been taught when all instances of that concept are responded to in the same way,"...even though some were not in the teaching set, and any or all (not-instances) are responded to in a different way" (Englemann, Becker, Thomas, 1971, 238-241).

Englemann, et. al. (1971) claim that concepts are taught by the presentation of instances (S<sup>D</sup>) and not-instances (S<sup>A</sup>), such that the range of instances, not-instances, and irrelevant stimulus characteristics are taught. He presents these guidelines:
"1. It is not possible to teach a concept through one instance and one not-instance. A set of instances and not-instances is required.

2. The set should be constructed so that all instances have all essential concept characteristics, and not-instances possess none or only some of these characteristics.

3. Within the set of instances and not-instances, it is necessary to vary stimulus characteristics that are not essential to instances or not-instances" (Englemann, et. al., 1971, 240-246).

A common procedure for conditioning conceptual behavior is described by Whaley and Malott (1971): one response should be reinforced in the presence of one stimulus class (concept), and that same response should be extinguished in the absence of that stimulus class.

Espich and Williams (1967), Whaley and Malott (1971), and Englemann, et. al. (1971) all describe similar procedures for conditioning conceptual behavior. All three procedures require the student to discriminate between concept instances ($S^D$) and not-instances ($S^A$). Also, Espich and Williams (1967) and Whaley and Malott (1971) specify a need for differential consequation for correctness.

**Conceptual Learning from Programmed Instruction**

Assuming these procedures are adequate, some programming techniques are more amenable to the teaching of concepts than others. The discrimination frame sequence most closely resembles the procedures above. It presents $S^D$'s and $S^A$'s for the concept being taught. Correct responses are reinforced in the presence
of the $S^p$, and extinguished or even punished in the presence of the concept $S^A$, by corrective feedback following each frame.

Although discrimination frames are the most amenable to concept teaching, the other techniques mentioned may also be helpful. The retrogressive chaining technique could be used to teach the student to analyze an example to determine if it is an instance or not-instance of a given concept. And adjunct programming could help utilize any existing instructional materials that deal with the concepts to be taught.

The Problem

Hubbard (1971) analyzed the final exam performance of students, in an introductory psychology course at Western Michigan University. The results indicated that the texts (i.e., Whaley & Malott, 1971; Psych 150 and Friends, 1971), and the accompanying class procedures (Hubbard, 1971), were adequate for students to learn at the level described by Espich and Williams (1967) as "recognition." The results also indicated that the text and procedures were adequate for learning to take place at a conceptual level, however the amount of conceptual learning was less than desired. The class utilized a standard text and daily multiple-choice quizzes. The large enrollment required the course administrators to use a proportionally large number of undergraduate teaching apprentices. The difficulty in obtaining a satisfactory level of quality control in the grading of daily quizzes led to the use of multiple-choice questions.

The problem attacked by this study, was the development of supplementary instructional materials, to be utilized in conjunction
with the regular text, that would improve conceptual performance, and still retain the use of a multiple-choice format.

The Concepts

Initial testing revealed that conceptual learning did not occur with a satisfactory level of accuracy. This difficulty in obtaining a high level of conceptual performance was not isolated to select concepts, rather there was a general inadequacy of the text and course procedures concerning conceptual emphasis. Based on these results, "discriminated extinction" was arbitrarily chosen as the first programming topic. Difficulty in obtaining satisfactory conceptual performance with this program, was partially attributed to a lack of necessary prerequisite skills.

"Discriminated extinction" is very similar to three other concepts presented in the text. First, the text presents "discriminated extinction" as a special type of stimulus control, in which the absence of reinforcement for a response functions as an $S^A$ for that response. For a student to learn discriminated extinction in this way, he must first learn "stimulus control" at a conceptual level, because he must be able to identify the defining characteristics of stimulus control, and be able to identify these characteristics in instances of discriminated extinction. Also, since discriminated extinction is a special type of stimulus control, a student must also be able to discriminate between stimulus control that is discriminated extinction, and stimulus control that is not discriminated extinction.
Discriminated extinction is also easily confused with the simple extinction process. The discriminated extinction procedure is very similar to the simple extinction procedure. In both cases, the response is allowed to occur and reinforcement is withheld. The procedural difference is that in discriminated extinction, the response is reconditioned and extinguished many times (Whaley & Malott, 1971).

The final concept that is usually confused with discriminated extinction is punishment. The difficulty arises because both procedures can result in an immediate cessation of responding. This is due to the response contingent presentation or removal of stimuli in the case of punishment and noncontingent withholding of reinforcement in the case of discriminated extinction.

All of the above concepts create considerable difficulty for the programmer. Not only is it necessary to teach what the concept is, but it is also necessary to teach what the concept is not. To facilitate the formation of discriminations among these closely related concepts, a more careful consideration of prerequisite skills was necessary.

One of the advantages of the concepts being taught was the inherent procedural and behavioral symmetry that results from an operant analysis of behavior. In general, behavior is strengthened by reinforcement and weakened by punishment. The procedural symmetry is centered around positive and negative procedures. In a positive procedure a stimulus or event is presented following the response to be affected; in a negative procedure a stimulus or event is removed.
By combining the behavioral and procedural components, four procedures that affect behavior exist.

Positive and negative reinforcement both strengthen the response they follow. Positive reinforcement strengthens a response due to the contingent presentation of stimuli or events, whereas negative reinforcement strengthens a response due to the contingent removal of stimuli or events. Positive and negative punishment procedures both weaken responding. This is due to the response contingent presentation of stimuli or events in positive punishment, and removal of stimuli or events in negative punishment.

The distinction between positive and negative procedures is a common one, however some theoretical problems seem to be created by this distinction. Problems arise when the compliment of the observable environmental consequence is considered. For example, is the response of washing an automobile reinforced by the presentation of a shiny vehicle, or is it reinforced by the removal of the dirt? The interpretative problems seem to result from a study of the relative stimulus change following a behavioral act. Washing the car removes the dirt, however relative to the dirty car, a clean car is presented.

It was decided that although this theoretical issue existed, there was a procedural difference. When one considers the observable environment, the only stimulus change that resulted from washing the car was the removal of the dirt. The car was always present.
Furthermore, students would be handicapped if they could not create examples of these two procedures, regardless of the theoretical source of reinforcement. Ideally, a student should be able to devise a procedure to facilitate responding, due to either the presentation or removal of observable environmental events.

There are several procedural dimensions that may be altered independently of reinforcement and punishment, yet determine the behavioral effect of these procedures. These procedural dimensions comprise the teaching sequence devised. It was decided to teach the symmetrical concepts of reinforcement and punishment, while changing the various conditions in which these concepts were taught.

The first procedural dimension was labeled conditioning. Under this, the first program would be Conditioning by Positive and Negative Reinforcement. The next program would be Conditioning by Positive and Negative Punishment. The second condition considered was extinction. Therefore the third program written would consider the Extinction of Behavior Maintained by Positive and Negative Reinforcement and the fourth program would be Extinction of the Behavioral Effect Produced by Positive and Negative Punishment. Then the teaching sequence would progress to the consideration of antecedent stimuli, see table 1.

The prerequisites for each topic appear in table 1, horizontally to the left of the topic. For example, the prerequisite for learning the concept of extinction of conditioned behavior, is learning
Table 1
Teaching Sequence of Elementary Operant Concepts

<table>
<thead>
<tr>
<th></th>
<th>non-cued</th>
<th></th>
<th>cued</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>conditioning</td>
<td>extinction</td>
<td>conditioning/extinction</td>
</tr>
<tr>
<td>positive</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>reinforcement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>punishment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>punishment</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the concept of conditioning. Reinforcement and punishment are only dealt with simultaneously to aide the student with the symmetrical aspects of the procedures, and is independent of the prerequisite skills necessary for learning the concept.

Development and Evaluation of the Programs

Program I: Discriminated Extinction

Phase one. The first step in the research and development of educational materials, had already been performed by Hubbard (1971). She analyzed the existing system and concluded that it was not satisfactorily teaching behavioral principles at a conceptual level.

The second step was to specify the behavioral objectives. Stated another way, what will the student be able to do once he has learned the principle of discriminated extinction at a conceptual level? Conceptual learning has occurred when the student can respond the same way to all instances of a given concept, and respond in another way to all not-instances of that concept, even though some instances were not in the teaching set. Another level of conceptual learning should prepare the student to generate his own original instances of a given concept. With this in mind, the following objectives were specified for the program on discriminated extinction. After reading the program, a student should be able to: (1) correctly discriminate between instances and not-instances of discriminated extinction, and he should be able to (2) write original examples of discriminated extinction.
The third step was to design the system. A program was written that primarily resembled a discrimination frame sequence. The program started with two short paragraphs describing discriminated extinction; a series of examples followed these paragraphs. Once the student responded to each five examples as either instances or not-instances of discriminated extinction, corrective feedback was provided. The feedback included an analysis of each example in terms of the characteristics of discriminated extinction, and if the example was a not-instance, the feedback told the student what the example was and why. The entire program consisted of twenty examples. Eight of the twenty examples were instances of discriminated extinction, and the remaining examples illustrated stimulus control, punishment, extinction, and response differentiation.

Two test forms were also written; each test consisted of a paragraph describing the concept, then ten questions. The first nine questions consisted of two instances of discriminated extinction, two instances of discrimination behavior, two instances of extinction, one of punishment, one of ratio strain, and one that illustrated a post-reinforcement pause. The tenth question required the student to write an original example of discriminated extinction.

The paragraph describing the concept was included on the tests to check the possibility that students could learn the concept from only the definition and a short prose explanation. Since this paragraph was on both the pretest and posttest forms, the definition alone could not be responsible for any increase in conceptual accuracy.
Only two of the test questions were instances of discriminated extinction; because it was considered equally important that students know what discriminated extinction is not. This was later changed to three questions. No direct attempt was made in the program to teach all of the above concepts; however an attempt was made to present examples that students were likely to confuse with discriminated extinction. The tests also included concepts that had not been dealt with in the program. It was reasoned that if the program could indirectly teach the student that these concepts (i.e., ratio strain, post-reinforcement pause) were not-instances, it would be an even more effective program.

The program was to be tested utilizing a pretest-posttest design, with two similar, but different test forms. Half of the students would take one form for the pretest, and the other half would take the second form. Then, for the posttest, each student would take the remaining form. The students were to receive bonus points based on their overall (pretest and posttest) performance. These bonus points could be applied to their grade in the introductory psychology course, and they could earn up to five points. (A daily quiz in this course was worth ten points; there were four quizzes each week.) Therefore, students could earn the equivalent of one-half of a daily quiz.

The program was written with the assumption that students would read it after they had read and been quizzed on the primary text. The prose explanation included a summary of the main points
in the text, thus attempting to maximize the probability that students' initial repertoire was adequate for the program. It was assumed that the corrective feedback and points earned, would be adequate sources of differential consequence for participation and correctness.

The fourth step in the development of the program was to test it to determine if it accomplished its objectives. Fourteen volunteers from the introductory psychology course at Western Michigan University were tested. The pretest mean was 50% and the posttest mean was 73.6% yielding a t value of 3.69 (p < .01, df=13). An error analysis of the pre and posttest indicated that students were having the most difficulty discriminating between examples of discriminated extinction and extinction. The test results also indicated that correct example writing could be taught using the multiple-choice format of the program. On the pretest, the mean percentage for example writing was 14%, and this increased to better than 64% on the posttest.

Phase two. Evaluation of the program indicated that additional practice discriminating between discriminated extinction and extinction was necessary, as the average error rate for this distinction did not decrease from the pretest to the posttest. Ten examples, and their respective corrective feedback, were added to the program. In addition, a half page clarification on punishment, extinction and stimulus control was added to the middle of the program.

Eighteen student volunteers were tested under the same conditions described above. Pre and posttest means of 43.3% and 71.1%
respectively were obtained which also yielded a significant $t$ of 5.56 ($p < .01$, $df = 17$). An error analysis of the test items revealed a more even distribution of errors, as compared to the first test. The problem at this point was to improve the posttest performance in general, in an attempt to obtain posttest means of at least 90%.

**Phase three.** The next step was to determine what sections of the program, if any, hindered the students' learning of the concept. Three versions of the program were prepared, two of which eliminated one aspect of the program. The first altered program eliminated definitions and prose explanations; the second eliminated a series of five examples. Three groups of students were tested ($n_1 = 9$, $n_2 = 9$, $n_3 = 8$). The complete version of the program yielded pre and posttest means of $48.9\%$ and $71.1\%$. The second version yielded means of $50\%$ and $74.5\%$, and the final draft yielded means of $51.3\%$ and $82.5\%$. A detailed error analysis of posttest scores for each of the three groups revealed that students were still confusing punishment and extinction with discriminated extinction. An analysis of program errors yielded similar results.

**Phase four.** At this time, all of the program frames having an average error rate over 10% were either rewritten or replaced. Similarly, two test items were rewritten in an attempt to make the examples clearer. And finally, the section of the program that was eliminated in the third version of the program above, was replaced with a sequence of five new examples and their respective corrective feedback.
These efforts were justified by the fourth testing of the program. Six student volunteers had a pretest mean of 43.3% and a posttest mean of 88.3%, which yielded a \( t \) value of 8.5 \((p< .01, \text{df}=5)\). Posttest performance was very near the arbitrary 90% goal.

**Phase five.** The fifth draft of the program was identical to the fourth draft, however the blocks of five examples and five answers were broken down so that each example was followed by its answer. Previously the student had to read five examples before he received feedback, now feedback was more immediate. Thirteen student volunteers were tested, and the data were analyzed in three groups: the complete program, the first nine questions (example identification), and the last question (example writing). The complete program yielded pre and posttest means of 45.4% and 86.9%. The nine objectively scored questions had pre and posttest means of 49.5% and 87.1%. The last question (original example) had pre and posttest means of 7.7% and 84.5%. All three analyses produced significant \( t \) scores \((7.54, 7.68, 7.70\) respectively, \(p< .01, \text{df}=12)\).

**Phase six.** At this point, testing was conducted on a one to one, programmer to student ratio. This approach was an attempt to uncover problem areas in the program that were not revealed by detailed error analyses. As a result of this testing, minor revisions were made on the program; however, the general format remained unchanged.
Phase seven. The final testing of the program included three groups. The program group took the pretest, read a relevant section copied from the course textbook, and took the posttest. The test only group took the pretest, remained in their seats for 30 minutes without any reference material, and took the posttest. Each group was divided into two sub-groups, each taking a different pretest quiz form.

The purpose of the text group described above was to test for the possibility that the pretest "sensitized" the reader to the text material. If this were the case, the student could read the standard text on the concept being tested, after completing a pretest, and improve significantly on the posttest. The third group described above tested for the possibility that posttest scores improve simply as a function of exposure to the pretest.

For the final testing of the program the testing procedures changed somewhat. Rather than asking for volunteers, students were required to participate in this activity for several different sections of the introductory psychology course. This decision was based on the need for a greater number students necessitated by the text and test only groups. Due to the fact that this was a required class activity, it was necessary to increase the point value of the tests to be consistent with the existing class procedures. Each question on both tests was equal to 1/2 of a quiz point. Therefore, a perfect score on both tests gave the student 10 points for the activity, or the equivalent of a normal daily quiz over an hour-long reading assignment.
Six different types of packets were assembled, two types for each of the three groups. The two types of packets for each group varied only in the form of the pre and posttests. The six types of packets were arranged in a random order, and distributed to the students in this way as they entered the classroom. The pretest form was indicated on the outside of the envelope, and students were not allowed to sit next to another student with the same pretest form. All of the test procedures were conducted and monitored by the students' daily instructor, to minimize the fact that it was a thesis project. The experimenter was available in a nearby room to handle any special problems that occurred during testing.

One additional qualification is necessary. The introductory psychology course was structured to allow students with the letter grade of "A" at the end of the twelfth week (of a sixteen week semester), the option of not attending class for the remainder of the semester. Consequently, the remaining students at the time of this test, were those students who had achieved less than 90% on all of their daily activities and/or had failed to score 90% on the final exam. Only 206 of the original 1127 students remained.

The results of the final test are presented in table 2. The program group achieved significant differences in all three areas: overall program, example identification, and example writing. The remaining two groups did not achieve significant differences in any of the three areas.
### Table 2

**Final Results: Program I**

<table>
<thead>
<tr>
<th>Questions Analyzed</th>
<th>Group</th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Mean Diff.</th>
<th>t Score</th>
<th>s df</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 multiple choice questions</td>
<td>program</td>
<td>43.4%</td>
<td>69.7%</td>
<td>26.3%</td>
<td>6.38</td>
<td>.37</td>
<td>21</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>53.7%</td>
<td>56.3%</td>
<td>2.6%</td>
<td>.65</td>
<td>.37</td>
<td>22</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>57.7%</td>
<td>57.7%</td>
<td>0.0%</td>
<td>0.00</td>
<td>.46</td>
<td>20</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td>1 essay question</td>
<td>program</td>
<td>4.0%</td>
<td>59.0%</td>
<td>55.0%</td>
<td>5.50</td>
<td>.10</td>
<td>21</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>22.0%</td>
<td>30.0%</td>
<td>8.0%</td>
<td>.57</td>
<td>.14</td>
<td>22</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>24.0%</td>
<td>28.0%</td>
<td>4.0%</td>
<td>.28</td>
<td>.14</td>
<td>20</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td>complete test</td>
<td>program</td>
<td>39.5%</td>
<td>68.6%</td>
<td>29.1%</td>
<td>7.10</td>
<td>.41</td>
<td>21</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>50.4%</td>
<td>53.9%</td>
<td>3.5%</td>
<td>.79</td>
<td>.44</td>
<td>22</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>51.8%</td>
<td>54.3%</td>
<td>-.5%</td>
<td>.10</td>
<td>.48</td>
<td>20</td>
<td>p &gt; .01</td>
</tr>
</tbody>
</table>
Reliability checks were taken on the grading of question ten (example writing). All essay questions were graded independently by two graduate students, without consultation or knowledge of the test group the student was in. Items that were left unanswered by the students were not included in the computation of the reliability figure. Both graders agreed on the grading of 89.1% of the essay test items.

Program II: Conditioning of Behavior with Positive and Negative Reinforcement

Phase one. Generally the rationale and procedures were the same as those for Program I.

After reading the program, a student should be able to: (1) correctly discriminate between instances of positive reinforcement, negative reinforcement, and not-instances of reinforcement, and he should be able to (2) write original examples of positive and negative reinforcement.

The first eight questions of the tests were examples. The students had to identify each example as one of three possibilities: an instance of positive reinforcement, negative reinforcement, or non-reinforcement. The eight examples included two instances of positive reinforcement, two of negative reinforcement, two of punishment, one of extinction, and one of discriminated extinction. Students generally confuse negative reinforcement with the concepts of punishment and extinction, due to the connotations of the word "negative." For this reason, most of the distractors
in the tests (non-reinforcement) dealt with those behavioral processes that decrease the rate of an ongoing response. The last two questions on each quiz required the student to write original examples of positive and negative reinforcement.

After the two test forms were completed, the program was written. In the present program, the concepts were broken down to the individual components of the operational definitions, then the program attempted to teach these components. For example, the program began by having the student discriminate between instances and not-instances of "the presentation of stimuli."

The program was written in the format of a discrimination frame sequence, with feedback provided after each frame of five examples. The feedback consisted only of the correct answers--no explanation was given.

The program first dealt with the components of the concept of positive reinforcement, then negative reinforcement, and finally, the concepts of positive and negative reinforcement simultaneously.

Eleven student volunteers generated pre and posttest means of 53.6% and 61.8% respectively, yielding a t value of 1.67 (p > .01, df=10).

Phase two. At this time the parts of the program dealing directly with the components of the concepts, were abandoned, and the program was completely rewritten in a form similar to that of Program I. The program initially presented definitions and summaries from the text, followed by a series of twenty examples.
Five of these examples were instances of positive reinforcement, five were negative reinforcement, and the remaining were not-instances that included the concepts of punishment, extinction, discrimination behavior, and examples that resembled positive and negative reinforcement, but lacked one or more of the defining components. Each example required a response from the student. Following the response, the student was provided with the correct answer and an analysis of the example in terms of the components of the definitions of positive and/or negative reinforcement.

In the middle of the program there was a half page prose explanation on analyzing examples in terms of the necessary components of the two concepts. This appeared in the middle of the program, rather than the beginning, so the student would have some experience at attempting the required discrimination before reading this section. In this manner, it was hypothesized that the explanation would be more meaningful, and thus control more behavior.

This draft of the program was tested with seven student volunteers. A t value of 1.33 (p > .01, df=6) was obtained from pre and posttest means of 70.0% and 82.8% respectively. This particular test was conducted near the end of the Fall semester 1971. The poor performance could be due to the students' grades being relatively unaffected by their scores. Also, it was very difficult to obtain student volunteers at this point in the semester. Another possible source of difficulty is the students at the time of this test had completed three weeks of remedial work. This apparently
affected the pretest scores, as they were abnormally high. Consequently, testing was discontinued until the following semester.

The same draft of the program was then tested with fourteen student volunteers. A $t$ score of 4.23 ($p < .01, df=13$) was obtained. The pre and posttest means were 56.4% and 81.4% respectively.

Phase three. Many of the program and test examples were re-written, and the directions on analyzing examples were moved to the beginning of the program. Many examples were rewritten to include: baseline information, a description of the contingencies and consequences, and their effect. In addition, several program and test items were changed on the basis of a detailed error analysis.

Sixteen student volunteers were tested with the revised program and tests. The pre and posttest means were 53.7% and 85.6% overall. The eight objectively scored questions (identification of examples) were analyzed independently. The pre and posttest means were 43.1% and 86.8%. The subjectively scored questions (example writing) had pre and posttest means of 56% and 81%. All three analyses yielded significant $t$ scores ($p < .01, df=15$). The $t$ scores obtained were 7.97, 8.15, 3.85 respectively.

Students were still having difficulty with instances of punishment and negative reinforcement. Also, most of the incorrect, original examples of negative reinforcement illustrated either extinction or punishment.
Phase four. The corrective feedback for examples of punishment and negative reinforcement were expanded. Eighteen students were tested, and significant $t$ scores were obtained. The pre and posttest means were 65.5% and 81.5%. Since the overall performance was not improved, it was decided that any further development of the program should be conducted on a one to one, student to programmer ratio.

Phase five. During this phase (one to one), the program was expanded from twenty examples to twenty-three. Two of the new examples were instances of negative reinforcement; and one was punishment. In addition to the expansion of the program in this way, several frames were added to the beginning of the program. These frames dealt with the analysis of examples. Then, following each example in the program, there was a short series of questions related to the analysis of the concepts. Students then transferred their answers to these questions to a flowchart. If they answered the questions correctly, the flowchart provided the correct answer. Also after each example, the behavior to be analyzed was specified. In examples where several different responses occurred, this helped to eliminate unnecessary ambiguity.

Each frame included an example; specification of the behavior to be analyzed; a series of questions pertinent to a behavioral analysis of the example; a reference to the flowchart, so students could determine their answer; a multiple-choice question, where the student marked his response; and the answer with a detailed explanation.
Phase six. One important difference existed between the test conditions for the two programs. Program II was tested one day prior to the first final exam; therefore the test population included all of the students but the grades of approximately 80% of the students tested could not be affected by their performance. Many students did not have to take the quiz at all, and would have still received an "A" for their final grade.

The program group (see table 3) achieved significant differences in all three areas: overall program, example identification, and example writing. The remaining two groups did not achieve significant differences in any of the three areas.

Reliability checks on the scoring of example writing, revealed grader agreement on 90.1% of the items.

Program III: Extinction of Behavior Maintained by Positive and Negative Reinforcement

Phase one. Generally the rationale and procedures were the same as those for the first two programs.

After reading the program, a student should be able to: (1) correctly discriminate between instances of extinction of behavior maintained by positive or negative reinforcement, (2) correctly discriminate between instances and not-instances of extinction of an operant response, and (3) write original examples of extinction of behavior maintained by positive and negative reinforcement.

Next, both forms of the criterion tests were written. At the top of each form was an operational definition of the extinction procedure and extinction results for behavior maintained by
Table 3

Final Results: Program II

<table>
<thead>
<tr>
<th>Questions Analyzed</th>
<th>Group</th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Diff.</th>
<th>t score</th>
<th>S</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 multiple choice</td>
<td>program</td>
<td>55.2%</td>
<td>80.5%</td>
<td>25.3%</td>
<td>3.49</td>
<td>.58</td>
<td>24</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td>questions</td>
<td>text</td>
<td>62.0%</td>
<td>62.0%</td>
<td>0.0%</td>
<td>0.00</td>
<td>.38</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>50.0%</td>
<td>57.0%</td>
<td>7.0%</td>
<td>2.06</td>
<td>.29</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td>2 essay questions</td>
<td>program</td>
<td>38.6%</td>
<td>62.0%</td>
<td>23.4%</td>
<td>2.82</td>
<td>.17</td>
<td>24</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>41.5%</td>
<td>42.0%</td>
<td>.5%</td>
<td>.07</td>
<td>.14</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>30.0%</td>
<td>38.0%</td>
<td>8.0%</td>
<td>1.60</td>
<td>.10</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td>complete</td>
<td>program</td>
<td>51.6%</td>
<td>76.8%</td>
<td>25.2%</td>
<td>3.50</td>
<td>.72</td>
<td>24</td>
<td>p &lt; .01</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>58.7%</td>
<td>59.2%</td>
<td>.5%</td>
<td>.14</td>
<td>.34</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>46.0%</td>
<td>53.2%</td>
<td>7.2%</td>
<td>2.20</td>
<td>.32</td>
<td>24</td>
<td>p &gt; .01</td>
</tr>
</tbody>
</table>
positive and negative reinforcement.

Next, both forms of the criterion tests were written. At the top of each form was an operational definition of the extinction procedure and extinction results for behavior maintained by both positive and negative reinforcement. As with the previous programs, each quiz had ten questions. The first eight questions were examples that the students had to identify as: extinction of behavior maintained by positive reinforcement; extinction of behavior maintained by negative reinforcement; a not-instance of extinction. The first eight questions consisted of two examples of extinction of behavior maintained by positive reinforcement; two examples of extinction of behavior maintained by negative reinforcement; two examples of punishment; one example of avoidance behavior; and one example of positive reinforcement. The remaining two questions required the students to write original examples of extinction.

After both quiz forms were completed, the program was written. The format very closely resembled the later reinforcement programs. The program began with operational definitions of the extinction procedures and results. Then the student was presented with several questions he should answer to determine if a given example illustrates extinction. Once the student has answered these questions he was provided with a flowchart. If he followed the flowchart correctly, he arrived at the correct answer. Following this section was a series of examples, and each example required a response. Following the example was a behavioral analysis in terms of extinction. There were fifteen examples in the program, seven of which were extinction.
The remaining examples were punishment, reinforcement, and nonsense examples.

Sixteen student volunteers were tested under conditions identical to those in which the first two programs were developed. Pre and posttest means of 65.6% and 73.7% were obtained. This performance did not produce a significant t score (p > .01, df=15). A further analysis of the data after the scores were broken down into objective (first eight questions) and subjective questions, also failed to reveal a statistically significant difference for either part.

Phase two. An analysis of the pre and posttest results indicated that students were having the most difficulty with examples that illustrated extinction of behavior maintained by negative reinforcement, and punishment by the contingent removal of stimulus events. Also, many of the incorrect examples written illustrated reinforcement of incompatible behaviors, rather than extinction.

The program was revised accordingly. Instances of reinforcement of incompatible behavior were added to the program as well as more examples of the other two problem areas. Several of the analyses following the examples were also expanded.

Eighteen student volunteers were tested with the revised program. Pre and posttest means of 72.1% and 80.0% were obtained. A statistical analysis did not reveal any significant differences in the two distributions.
Phase three. At this time, development of the program progressed on a one to one, programmer to student ratio. Drastic revisions were made in the flowchart, and corresponding changes were made in the questions prior to the flowchart. Many examples were rewritten to eliminate ambiguities, and several more examples of extinction were added to the program.

Final testing was conducted under conditions similar to the two previous programs. As with discriminated extinction, final testing occurred after the final exams, therefore only the students who did not exhibit "A" level performance participated.

The results of the final test are presented in table 4. None of the three groups achieved significant differences.

Reliability checks on the grading of the last two questions (example writing) revealed agreement on 92.8% of the test items.

Discussion

Lower absolute scores were obtained during the final tests, than during the developmental phases. This could be due to the change in the test populations. During the development of Program II, the same program was tested at the end of the Fall semester 1971, and several weeks into the Winter semester 1972. The students tested during the Fall semester were students who had less than an "A" as of the twelfth week. The students tested during the Winter did much better. The final tests for Programs I and III were conducted under conditions similar to the Fall semester. Since the test population was comprised of the bottom 20% of the students, the absolute scores should be lower.
<table>
<thead>
<tr>
<th>Questions Analyzed</th>
<th>Group</th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Mean Diff.</th>
<th>t score</th>
<th>df</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 multiple choice</td>
<td>program</td>
<td>58.1%</td>
<td>68.4%</td>
<td>10.3%</td>
<td>1.67</td>
<td>.49</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>55.0%</td>
<td>54.1%</td>
<td>-.9%</td>
<td>.14</td>
<td>.47</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>51.9%</td>
<td>57.8%</td>
<td>5.9%</td>
<td>1.27</td>
<td>.37</td>
<td>12</td>
</tr>
<tr>
<td>2 essay questions</td>
<td>program</td>
<td>29.5%</td>
<td>41.0%</td>
<td>11.5%</td>
<td>1.05</td>
<td>.22</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>13.5%</td>
<td>23.5%</td>
<td>10.0%</td>
<td>1.17</td>
<td>.17</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>42.5%</td>
<td>33.0%</td>
<td>-9.5%</td>
<td>1.11</td>
<td>.17</td>
<td>12</td>
</tr>
<tr>
<td>complete test</td>
<td>program</td>
<td>52.3%</td>
<td>62.9%</td>
<td>10.6%</td>
<td>1.86</td>
<td>.57</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>text</td>
<td>46.7%</td>
<td>48.0%</td>
<td>1.3%</td>
<td>.25</td>
<td>.51</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>test only</td>
<td>50.0%</td>
<td>53.1%</td>
<td>3.1%</td>
<td>.86</td>
<td>.36</td>
<td>12</td>
</tr>
</tbody>
</table>
Final testing for Program II was conducted prior to the twelfth week; however, just before the final exam. Due to this, approximately 80% of the students' final grades were not affected by their scores on the pre and posttests. This motivational variable could account for the lower absolute scores during the final test.

The data indicate that the last several phases of the development of each program accomplished very little. The posttest scores for Programs I and II were approaching the arbitrary 90% goal, then decreased drastically for the final test. The revisions made during the one to one development phases, were not of sufficient magnitude to account for this decrease. Primarily, the revisions consisted of cleaning up ambiguous examples. Therefore, this decrease in performance is probably not due to changes in the programs.

Failure to obtain significant t scores for Program III was probably a function of time. The program had less developmental phases than the other programs, and the results were improving with each phase. Probably an effective program could be developed from the existing material, given enough time.

There seems to be a general misconception about the use of a multiple-choice format in education. The argument implies that learning above the recognition level, probably will not occur, due to the requirements placed on the students. The programs developed in this study utilized a multiple-choice format, yet conceptual learning was demonstrated in both the identification and writing of examples. Clearly this contradicts the notion that a multiple-choice
format cannot teach above recognition. This does not mean that anything in this format will be successful; rather, the content is probably as important as the format in an instructional sequence.

Conclusions

This study has traced the development of three programs. Perhaps the next step should be an analysis of the existing programs, to reveal which particular aspects contribute to their success. It would also be interesting to retest the programs, to determine if motivational variables and a change in the test population did contribute to the lower final test results obtained.

The problem addressed by this study was the development of adjunct programs that would improve the conceptual performance of students in an introductory psychology course. Two of the three programs achieved their objectives, and the third suffered from a lack of time. Finally, the multiple-choice format was successful in obtaining an increased level of conceptual performance.
References


Appendix I
Sample Test

FORM B
POSITIVE & NEGATIVE REINFORCEMENT

Directions: If this is your first quiz (pretest), then answer questions 1-8 below in the spaces on the answer sheet marked 1-8. If this is your second quiz (posttest), then answer questions 1-8 below in the spaces on the answer sheet marked 41-48.

Write your name, student no., and lab section on the answer sheet. Also indicate the pretest and posttest forms on the answer sheet (Form A or B).

Positive reinforcement: the presentation of a stimulus following a response, with a resultant increase in the rate of that response. The stimulus is a positive reinforcer.

Negative reinforcement: the removal of a stimulus following a response, with a resultant increase in the rate of that response. The stimulus is a negative reinforcer.

Answer questions 1-8 either a, b, or c as indicated.

1. You didn't buy new tires for your car, even though all four of your tires were bald. All four tires blew out on a long trip, so you vowed never to ride on bald tires again. "Riding on bald tires" receives____________________.
   a) positive reinforcement  b) negative reinforcement  c) neither

2. You were finally caught cheating on an exam after 3 years. You got an "E" in the course, and now you don't cheat anymore. "Cheating" receives____________________.
   a) positive reinforcement  b) negative reinforcement  c) neither

3. You told your dog to roll over, and he rolled over. You said "good boy", and scratched him behind the ears. He now rolls over on command, consistently.
   a) positive reinforcement  b) negative reinforcement  c) neither

4. You tried to explain behavioral psychology to Spirot A. Gnew, and he didn't understand. You tried telling him about behaviorism another way, and another way...finally you gave up, figuring it was all a bad dream.
   "Explaining behavioral psychology to Gnew" receives____________________.
   a) positive reinforcement  b) negative reinforcement  c) neither
1. Joe Spat used to shine his shoes daily, because he always received compliments on them. Then Joe's tour of duty in the Marines ended and Joe returned to civilian life. Joe's frequency of shoe shining decreased because no one in the real world ever complimented him on his shoes.

REFER TO THE FLOWCHART FOR YOUR ANSWER.

In this example, "shoe shining"
   a) extinguishes due to the absence of positive reinforcement.
   b) extinguishes due to the absence of negative reinforcement.
   c) does not extinguish.

--------

ANSWER: This is extinction of a behavior maintained by positive reinforcement. Shoe shining was originally reinforced by the presentation of compliments. When the compliments stopped, so did shoe shining.

Extinction Procedure: no more compliments presented

Extinction Results: no more shoe shining

Mort was a brick layer, and he was good at it. Mort could build a wall quicker than four other bricklayers put together, and still use only one hand. Then one day mort's scaffold collapsed under him as he was putting a brick in place, and he fell 50 feet to the ground. Since that day Mort has slowed down, and is one of the slowest brick layers in the state. Even though Mort has slowed down, he still receives his weekly paycheck.

REFER TO THE FLOWCHART FOR YOUR ANSWER.

In this example, "brick laying"
   a) extinguishes due to the absence of positive reinforcement.
   b) extinguishes due to the absence of negative reinforcement.
   c) does not extinguish.

--------

Example two is not an example of extinction. The behavior (brick laying) decreases in rate, however, the decrease is not due to the withholding of reinforcement. Rather, the decrease is due to the presentation of a stimulus event (50 foot plunge). This is an example of punishment.