418

TO THE PUBLIC

AIR-649-1/70-FR

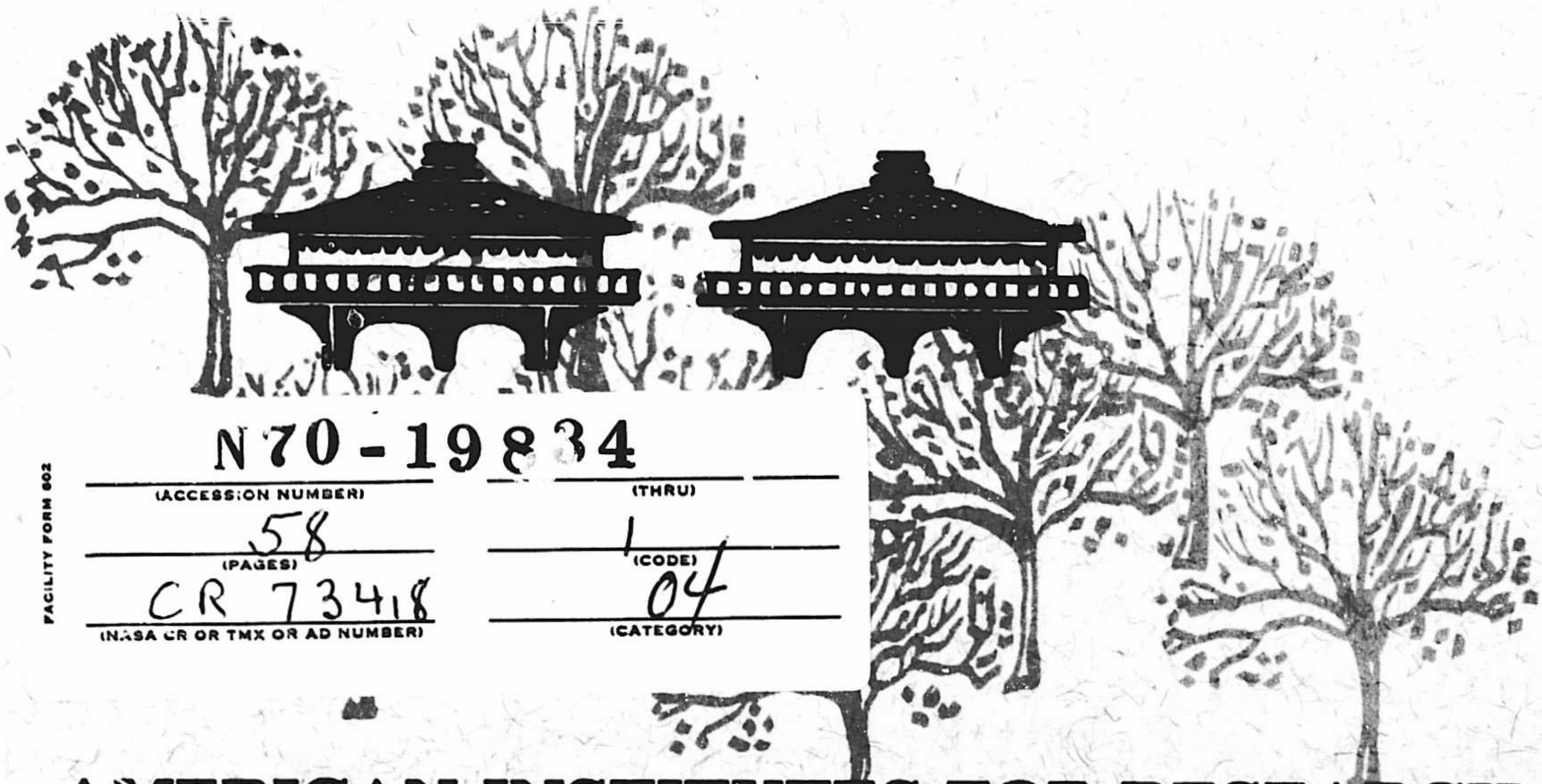
A STUDY IN IMPROVING PERFORMANCE OF COMPLEX TASKS

Lloyd O. Brooks

Final Report to

National Aeronautics and Space Administration

Ames Research Center



AMERICAN INSTITUTES FOR RESEARCH

Post Office Box 1113 / Palo Alto, California 94302

AMERICAN INSTITUTES FOR RESEARCH

SOCIAL AND EDUCATIONAL RESEARCH PROGRAM

January 1970

A STUDY IN IMPROVING PERFORMANCE OF COMPLEX TASKS

Lloyd O. Brooks
Principal Investigator and
Project Director

Contract No. NAS2-3916

National Aeronautics and Space Administration

Ames Research Center

Edward K. Crossman
Technical Monitor

Robert Dolan
Contract Administrator

mailing: P.O. BOX 1113, PALO ALTO, CALIFORNIA 94302 (LOCATION: 1791 ARASTRADERO ROAD) • TEL. (415) 328-3550

TABLE OF CONTENTS

LIST OF FIGURES

Page

THE CUEING OF SPEED REQUIREMENTS

A STUDY IN IMPROVING PERFORMANCE OF COMPLEX TASKS

Lloyd O. Brooks

ABSTRACT

Laboratory experiments were conducted, with college students as subjects, in which contingencies of reinforcement were the experimental treatments. Speed and correctness of responses to instructional items were the dependent variables. There were eleven independent groups; and the numbers of subjects ranged from five to twenty-one (who were in a Base-line Group). Subjects participated individually, completed the first 142 items of a psychology program under a base-line condition, the next 489 items under one of the experimental treatments, and then responded to a posttest.

A goal of the experiments was to identify procedures for improving performance of complex tasks. Subgoals were the development of techniques for accurately predicting later performance impairment from an early, pretreatment, sample of response times and the development of procedures for increasing performance speed without reducing correctness. In achieving the first subgoal, later impairment was defined in terms of the number of posttest errors on the posttest which followed the experimental treatment. Statistically significant predictions (multiple R's in excess of 0.5, cross validated with two samples of 35 subjects) were obtained on the basis of response times from three instructional items (selected from among the first four dozen at the beginning of baseline). That is, subjects who responded quickly to these instructional items made fewer posttest errors. Those who responded slowly made more.

Achievement of the second subgoal was through refining the experimental treatment which best provided inevitable and immediate conditioned reinforcement, since this was found to benefit correctness. With refinements, which included revised estimates of typical response times and were the basis for time criteria, the prolonging of stimulus consequences of correct and fast correct responses, etc., correctness on the posttest as well as during learning seemed to improve. Each subject of the two groups which incorporated the refinements made fewer than predicted posttest errors. They also increased their speed, reducing their response times by one-fourth to one-third of their baseline values.

The cueing of speed requirements did not further increase speed of responding nor did cueing increase correctness. Two cueing procedures were used with independent groups. One cueing procedure indicated to the subject how long he would be allowed for fast responding at the beginning of each item, as well as time remaining. The other cueing procedure merely indicated the last time unit at the end of the interval allowed. Differences in performance attributable to cueing procedure were not of practical or statistical significance.

A STUDY IN IMPROVING PERFORMANCE OF COMPLEX TASKS

Lloyd O. Brooks

American Institutes for Research, Palo Alto, California

INTRODUCTION

In this final report, the first major section contains findings based upon the first two years of study. Some of this material, and additional details which relate to it, was reported previously in interim and progress reports. The effects of experimental treatments on the number of posttest errors was not included previously. Nor were multiple regression analyses used to demonstrate statistically significant predictions of the number of posttest errors from response times on a few pretreatment components of the subject's learning task, although validities of individual items were noted. The recent cross validation also adds to the development of this topic. These findings have implications for personnel selection and for techniques of monitoring the quality of ongoing learning and performance.

As a commonsense observation, particularly slow performance on relatively simple elements of a task would not be expected to be to a candidate's credit or advantage. But one might expect early slowness to show later only in exaggerated slowness on a more difficult portion of the task, rather than in number of errors. That early slowness related to the number of later posttest errors raises a number of questions concerning where slowness needs to be measured in order to make predictions of later performance. As reported here, the investigator needed to go through the task item by item, first finding items for which response times had high validities for the criterion (number of posttest errors). A few with high validity were then combined in a multiple-regression equation, for improvement in accuracy of prediction over the accuracy of individual items. Cross validation was also part of the process, with two halves of the data. As will be explained below, slowness on a difficult early item would not have the same implications for the number of later posttest errors that slowness on an easy item would have.

Research conducted outside this project pointed up some of the complexities to be expected in using response times for predicting the number of errors. Subjects made fewer errors who spent longer on difficult elements than they did on easier ones. This finding appeared in several experiments in which steps were taken to ensure the selection of unambiguous easy or unambiguous difficult items. Ambiguity was regarded as indicated by disagreement between measures of task difficulty. The unambiguous easy item was responded to quickly and correctly. The unambiguous difficult item was responded to slowly and often incorrectly.

Of the three items from which response times were obtained to maximize successful prediction of the number of posttest errors, one item had been repeatedly identified among unambiguous items in several experiments in which there were statistically significant predictions of the number of posttest errors. Item 48 of Set 1 maintained such standing in a series of studies, even in a study in which the number of observing responses was used as a measure of task difficulty.

However, the main approach to improving performance was not through selecting the best candidate from a sample of early performances. Instead, it was through the identification of some of the critical details of procedures which contributed to the evolution of improved performance. There is a long history of experiments showing benefits from knowledge-of-effects feedback and other kinds of reinforcement. Our findings show that with reinforcement contingencies, which deal simultaneously with speed and correctness, incompatibility between the two is not inevitable. It seems particularly easy, of the two, to increase speed, even by simple instructions to subjects. However, the increase is superior and better maintained if reinforcing consequences are added to strengthen performances which are consistent with instructions. When speed is increased and the increase well maintained, even when correct responding is required in order for fast responding to be reinforced, correctness of responding varies during learning and on the posttest with other variables. One of these variables seems to be an immediate and unavoidable stimulus change which is correlated with reinforcement: placement of a conditioned reinforcing stimulus. It also seems to benefit performance (number of posttest errors) to prolong this stimulus. Finally, the adjustment of speed requirements, to demand more speed of the subject when he is capable of faster correct responding and less when he is not, seemed to prolong increasing of speed.

SPEED AND CORRECTNESS OF RESPONSES TO INSTRUCTIONAL ITEMS

Lloyd O. Brooks and John W. Clark

American Institutes for Research, Palo Alto, California

ABSTRACT

College students responded to a psychology program. Automated equipment awarded points worth money contingent upon correctness and speed of keyboard responses. The third letter of the last word of the subject's covertly composed answer was usually the required response. After 142 baseline items, the subject responded to 489 items under one of eight contingencies or under Baseline. A posttest followed. During treatment, subjects whose correct responses were always reinforced, showed the greatest increase in percentage correct. The greatest decrease was by subjects whose fast-correct responses were intermittently reinforced. Subjects whose fast-correct responses were always reinforced and whose conditioned reinforcement was contiguous with their correct responding approached Baseline subjects in percentage correct. Of these subjects, those responded fastest and continued to reduce their response times throughout treatment whose speed and correctness of responses adjusted their time criteria. These subjects also made fewer posttest errors than Baseline subjects. The subjects' response procedure can be adapted for the classroom by instructors who want the advantages of machine scoring without the disadvantages of multiple-choice items.

INTRODUCTION

College students completed an instructional program at a teaching machine as their laboratory task in experiments concerned primarily with speed and correctness of ongoing performance, rather than instructional benefits. Programmed learning was used as an example of a complex task. As regarded here, complex tasks characteristically entail mixtures of new and previous learning.

A programmed learning task seemed especially suitable for sustaining performance for laboratory study. Per hour of experimental session, this task provided for numerous responses by subjects. Therefore, there were many opportunities for the application of reinforcement contingencies. While designed to occasion correct responses by most students to most items, there were the inherent task complexities of variety of stimuli and heterogeneity of responses. In demonstrating control of reinforcement contingencies over many kinds of responses, studies have dealt mainly with homogeneous responses, recurrences of a particular response or members of a limited response class. Blough and Millward (1965) commented, ". . . there is relatively little research on complex response topographies, still less on the rapid shift observed in shaping" (p. 74).

The present report is related to an earlier experiment (Brooks, 1965; 1967) which dealt with shaping faster question answering by means of automated procedures which provided points on a counter. These points were worth money and served as reinforcement. Among these procedures were two arrangements -- related to the adjusting schedule (Ferster & Skinner, 1957) -- for adjusting the speed requirement of a speed and correctness contingency on the basis of the subject's performance. According to both arrangements, greater speed was required for the next reinforcement following each fast, correct, reinforced response. The first correct response was reinforced if it occurred in less than average read-and-answer time for a given item. Averages were based on data from a previous experiment and typically differed from one program item to the next. After the first response was reinforced there was a change in value of the time criterion for the next item. The change was from its initial average value to a lesser value. The amount of change, in seconds, was not always the same. Rather, the lesser value represented a fraction of any item's average value, in order to maintain proportionately among items of different average response times.

The difference between the two arrangements for adjusting time criteria was in the effect or lack of effect of non-reinforced responses. According to one arrangement, each non-reinforced response caused a reset of time criteria to initial, average values. According to the other arrangement, non-reinforced responses had no effect on time criteria. With progressive reductions in time criteria brought about by reinforced responses, only exceptionally fast correct responses came to be reinforced. This scheduling of a progressive decrease in time requirements had its parallel in an investigation of a progressive increase in ratio requirements by Hodos and Kalman, 1963.

In the previous study (Brooks, 1965) subjects responded faster when a fast correct contingency was in effect than when the reinforcement contingency was discontinued on counterbalanced halves of sets of items. (There was no exteroceptive stimulus to indicate presence or absence of reinforcement contingency on halves of lessons.) However, subjects responded faster on reinforced than non-reinforced halves of lessons whether the adjustment of time criteria was with or without reset by non-reinforced responses. The former (with reset) provided a greater number of reinforcements, but this excess in number over the non-reset procedure were reinforcements of slower responses. The latter (without reset) provided fewer reinforcements, but more of these were of exceptionally fast correct responses. Perhaps a better procedure than either of these would relax time requirements when requirements became adjusted out of reach of the subject but relax them only slightly. An advantage would be increased precision in the tracking of performance by response time criteria. Preliminary results obtained after the 1965 study (Brooks, 1967) suggested the superiority of this third procedure.

Accordingly, in the present experiments time criteria were relaxed by the small increments by which they had been rendered more challenging. That is, when time criteria had been reduced by a series of reinforced responses, it took a series of non-reinforced responses to increase them back to previous values.

In both the 1965 and 1967 experiments, and present experiments, initial time criteria were approximately average response times, as established for individual items. Allowance for differences among items in response times made it possible to accommodate differences in response classes in order to apply procedures derived from experiments concerned with repeated instances of a response of a single class. However, not all applications depend on such accommodation.

Berman (1966) applied an adjusting procedure in an experiment in which Mexican and Yaqui Indian students used a program in order to learn English. His application did not require automating of procedures, nor did it require that average response times be established for individual instructional items. Berman worked with blocks of instructional items. Adjustment was a matter of requiring greater speed or fewer errors of the student on a next block of items after he met the time or error requirement of a given block. The student's failure to satisfy time or error requirements resulted in the relaxation of requirements. Reinforcement became contingent upon less speed or a greater number of errors. Berman's results indicated that reinforcement contingencies which were adjusted by time or errors could control speed or accuracy. He later investigated the subject's preferences for speed versus accuracy contingencies (Berman, 1967). Adjustment through smaller steps seemed superior to that through larger steps in maximizing speed and accuracy (Berman, 1968).

Logan (1963), in emphasizing learning of response speed as an important part of the content of learning, noted similarity between his account and the earlier account of Ferster and Skinner (1957). Respectively, the two accounts seem to interpret speed as well as form of behavior as automatic consequences of learning variables and of reinforcement contingency variables. In expressing concern that common practices often fail to take all content of learning into account, Logan gave an example of the arithmetic teacher who, in teaching the pupil to avoid mistakes by adding slowly, might, unintentionally, teach slowness in adding. Ferster and Skinner's account might raise the question of whether slowness would be maintained after the contingency provided by the teacher came to be replaced by other contingencies of the pupil's subsequent environment.

Instructions to students were investigated in the experiment (Brooks, 1965) which provided findings on which the present experiments built. The joint effects of (1) instructions to the student to perform quickly and correctly, when accompanied by (2) reinforcement (points worth money), produced faster performance that either instructions or reinforcement alone. When fast and correct responses were reinforced without the subject being informed of the contingency, subjects differed significantly among themselves in their rates of responding. From their volunteered comments, subjects seemed to infer erroneously a variety of relationships between reinforcement and characteristics of their performances, including reinforcement when particular keys were pressed or when any key was pressed particularly firmly. Their comments are suggestive of superstitious behaviors developed in experiments with lower organisms (Skinner, 1953). For some subjects the reinforcement contingency was effective in increasing their fast, correct responding without their being informed of it. For others this contingency may, in time, have become similarly effective. However, it seemed evident that a description of the contingency would result in its having a more uniform and immediate effect. Therefore, the usefulness of instructions was tentatively accepted for the present experiments, which address new issues.

Among the new issues were the effect on speed of responding of an adjusting procedure which either increased or decreased response-time criteria by the same size of incremental unit, a comparison of performance adjusted time criteria with fixed criteria, how time criteria were best adjusted by the subject's performance (e.g., whether criteria were increased by a slow or slow incorrect response), how response correctness could be increased (as by adding conditioned reinforcing stimuli), and the effect on speed and correctness of intermittent reinforcement. Of particular interest were the long-term effects on speed and correctness of experimental treatments. These presently were to be continued while the subject completed a dozen sets of items. (In the 1965 study, at most, the experimental treatment was in effect from the last half of one set of items through the first half of the next.)

Since one objective was to compare contrasting procedures for improving performance, only fast responses which were also correct were reinforced. The 1965 experiment demonstrated that increases in speed did not necessarily come at the expense of a decrease in correctness. Overall, within a fraction of one per cent, the same number of correct responses occurred during treatment as control condition. Since correctness was of fundamental importance, the manner in which the subject responded was changed. In the 1965 study the subject responded by giving the first letter of his answer word. In the present study, he was usually required to give the third letter. (There seemed to be less redundancy in the third letters of correct and incorrect responses than in the first letters.)

GENERAL PROCEDURES

Subjects and Design

An approximately equal number of male and female Foothill College students participated individually as paid volunteers in laboratory studies conducted at the Palo Alto office of the American Institutes for Research. All subjects began under a baseline condition which was later followed by one of several experimental treatment conditions (with the continuation of baseline as one of the possible treatments). After the experimental treatment each subject completed a posttest. Throughout the study subjects were self-paced.

Laboratory Task

Subjects completed the first 17 sets of The Analysis of Behavior (Holland & Skinner, 1961). Filmstrip copies of the published version of this program were prepared, with permission of the publisher (McGraw-Hill). Each subject completed the first four sets under the baseline condition, the next 12 under one of the experimental treatment conditions, and Set 17 as the posttest. All subjects responded to the first 16 sets of items by means of a keyboard. All wrote answers to items of the posttest.

Each subject in three of the nine groups first wrote his answer then responded to the same item of the posttest using the keyboard. The purpose of obtaining both kinds of response was to consider the validity of keyboard responses using written responses as the criterion. The number of write-answers errors of each subject were to be paired with his number of keyboard errors in an analysis of correlation between the two kinds of errors. Sixty-nine written errors were possible (since some items had more than one blank). Only 39 keyboard response errors were possible.

Apparatus

The subject used a 35 mm DuKane rearprojection filmstrip viewer (Model 576 65) to alternately display questions and answers. A first advance of the filmstrip brought a question into view; a second advance caused the correct answer to appear, etc. During baseline, and while the experimental treatment was in effect, the subject responded at a full-alphabet keyboard. The keyboard keys were alphabetically arranged in rows, unlike the standard typewriter keyboard. Each response to a question item, whether correct or not, caused the correct answer to appear. The subject operated the same key a second time in order to bring the next question into view. In responding to posttest items each subject was instructed to write complete answers to all parts of each item. He wrote on paper tape controlled by a DuKane write-in answer unit (Model 99A230). With the exception of three groups, the keyboard was no longer available, and subjects used a handswitch to simultaneously advance the filmstrip and answer tape. Subjects of Groups 7, 8, and 9 were instructed to respond to the keyboard as they had been responding during treatment after writing the answer to each posttest item. By substituting the keyboard for the handswitch both write-answer responses and keyboard responses were obtained from these subjects.

Recording and switching-circuit control equipment were located outside the subject's room, in two other rooms. A speaker in the subject's room provided a masking noise to attenuate sound from equipment. The subject sat with his back to a one-way viewing window. (The experimenters let the subject know that he would be observed from time to time.) The experimenter had access to recording and control equipment while observing the subject and could thus note portions of data rendered invalid by equipment malfunctions. Cumulative recorders and printing counters served as backup recording devices. Other apparatus details will be given along with the experimental treatments to which they relate. (Also, see APPENDIX, page 50.)

Instructions to the Student

In using the keyboard, which made possible an abbreviated overt response derived from a larger-scale covert response, the subject followed general instructions which were modified, when necessary, by specific instructions which explained exactly how he was to respond to a particular item.

Conventions followed by the program authors in specifying how the student was to answer were explained, and a summary of these was posted on the viewer. For example, TT called for a technical term, a single blank meant a one-word answer, etc. Then the subject was told to follow these conventions but respond with a key of the keyboard according to a general rule. He was to press the lettered key which corresponded to the third letter (counted from the beginning) of his answer. If the item called for more than one answer word, the subject was to respond only with the key corresponding to the third letter of the last word.

The subject was told the general rule sometimes would be superseded by specific instructions, which would appear beneath certain questions. These would aid the subject when an item (1) had more than one correct answer, (2) required an answer of fewer than three letters, and (3) could be "improved" by requiring a word other than the last. When several correct answers were possible, all were arbitrarily ruled out except the preferred or most common answer by listing their third letters as incorrect. For example, "Not e or o." This specific instruction would be used to rule out "frequency" and "probability" as responses to the item for which "rate" was the preferred answer. Other special instructions, "Second letter," superseded the general rule for yes-or-no items. The subject was also occasionally asked to respond according to a particular letter of some answer word other than the last.

The experimenters emphasized to the subject that he was to write complete answers to each posttest item.

EXPERIMENTAL TREATMENTS

The dependent variables were: 1) time spent on each item, and 2) the correctness of each response. With the exception of Baseline, the independent variables were reinforcement contingencies. These will be described in general, followed by details, and a summary.

Nine independent groups of subjects each participated under a different treatment condition while they completed Sets 5 through 16. Subjects participated individually and at their own rates, with two exceptions. Each subject completed the last part of the baseline phase (Set 4) and the beginning of his experimental treatment (Set 5) during the same session. He also completed the last set under treatment (Set 16) and the posttest (Set 17) during a single session.

During Sets 5-16 the condition in effect for all subjects during Sets 1-4 continued for one group: the Baseline Group, whose subjects did not receive reinforcement (points worth money). Another group received points contingent upon correct keyboard responding, regardless of their response times. The other seven groups obtained points by responding correctly within the time allowed, which typically differed from item to item. In addition to their speed-and-correctness contingency, one of these seven groups had an additional correctness contingency--a concurrent correctness contingency. Whenever one of these subjects responded correctly within the time allowed one point registered on his fast-correct counter and another on his correct counter, both located above and slightly left of viewer screen. Contingent upon their slow correct response, these subjects received weak reinforcement (a point on the latter counter which was worth half as much as points awarded to subjects of other groups or as a pair of points derived from both counters.).

Of the six groups whose reinforcement depended entirely upon fast correct responses, one group received reinforcement intermittently. The fast-correct-response contingency was in effect for a random two-thirds of their responses. The other five groups of subjects differed in procedural details ranging from one to several of the following: revision of estimates of average response time for establishing the fast-response criterion for each treatment item, revision of instructions concerning the subject's use of the keyboard, the location of conditioned reinforcing stimuli, and whether time criteria were set at estimated average response times for all items (fixed contingency) or adjusted from these values to allow less time on the item following a fast-correct response or more time following each non-reinforced response (adjusting contingency).

At the time subjects were recruited, they were told they would receive $1.00 as transportation allowance for each trip to the laboratory and $2.00 per hour. Half of the hourly earnings would be paid at the end of each session and the other half at the end of the experiment. Subjects who were assigned treatment conditions which permitted them to earn more than $2.00 per hour had the additional amount credited to them at the time of the session but received it at the end of the experiment. These subjects were told they would earn $1.50 per hour plus the amount earned on the basis of their performance. If the total was less that $2.00 they would merely continue to earn $2.00 for each hour. However, if the amount was greater they would receive the greater amount. These subjects were told how they needed to perform in order to increase their earnings.

Baseline (Group 1)

For these subjects bonus earnings of more that $2.00 per hour were not possible. Speed or correctness of responses was not indicated by a stimulus change such as lights or counters changing. However, these subjects, as well as other subjects, saw correct answers immediately after responding.

Correctness (Group 2)

Regardless of how long the subject spent in answering, each correct response resulted in reinforcement. As the subject pressed the correct keyboard key, he lighted an amber light for a fraction of a second and incremented a counter. The points counter and light were located above the viewer screen. Only one point was possible for each item answered correctly. Each point was worth two cents.

Speed and Correctness with Adjusted Response-Time Criteria (Group 3)

A response-time code was programmed for each instructional item during the experimental treatment. This code determined how many timing cycles were to be allowed to elapse before a correct response was no longer considered fast. Each session began with the cycle set at five seconds duration. The number of cycles programmed for each item resulted from dividing by five its estimated (in secs.) average response time, and rounding. (Estimates were based on data from previous studies.) Response-time criteria were adjusted by increasing or decreasing the cycle in half-second increments. Each fast and correct response reduced the cycle one half second. Each slow response, whether or not it was correct, increased the cycle duration one half second. The programmed time code which represented the number of cycles for each item, when combined with a variable-speed clock switching device, made possible an increase or decrease in the time allowed for "fast" responding while maintaining initial proportionality among item times. The points counter and light were the same as was used by subjects of Group 2, and each point was the result of a fast and correct response. Again each was worth two cents.

Speed and Correctness with Fixed Response-Time Criteria (Group 4)

The cycle size was fixed at five seconds duration and unchanged by performance. Of course, the total number of five second cycles for each item was determined by the estimated averages, as above.

Intermittent Speed and Correctness Reinforcement with Adjusted Response-Time Criteria and Added Stimulus (Group 5)

A stimulus (red light) was installed near the keyboard. It served as a conditioned stimulus which provided continuous reinforcement contingent upon fast-correct responses. According to a random pattern, the circuits of the amber light and points counter were inoperative while subjects responded to one-third of the items. Backup reinforcement (amber light flash and point on counter) which determined the subject's earnings was scheduled intermittently. As in the case of Groups 1 through 4, counter points were each worth two cents. These were absent on Sets 10 and 11, as were flashes of the amber light. The red light continued to flash each time the subject gave a fast correct response.

Continuous Speed and Correctness Reinforcement with Adjusted Response-Time Criteria and Added Stimulus (Group 6)

This treatment was the same as for subjects of Group 5 except that backup reinforcement was scheduled for each fast and correct response.

Revised Procedures for Speed and Correctness Reinforcement with Adjusted Response-Time Criteria and Added Stimuli (Group 7)

On the basis of response-time data, obtained from subjects in Groups 1 through 6, the number of five-second timing cycles for each item was revised. If the new data suggested a change of only one cycle for an item, its cycle value was not revised. Approximately one-third of earlier cycle values were changed.

Increasingly detailed instructions to subjects were prepared concerning the mechanics of their answering, for this group and the other two which follow. These instructions were followed by asking the subject how he would operate the keyboard if his answer were to be "is not." If needed, the experimenter supplemented written instructions by emphasizing a portion which the subject seemed to have misunderstood.

The red light near the keyboard was replaced by two lights in the base of the keyboard. The keyboard consisted of snap-action switches attached to clear plastic, and had a clear plastic base. The green light came on whenever the subject answered correctly . The red light came on whenever the correct answer occurred before the last timing cycle had ended. Simultaneously the points counter was incremented and the amber light associated with it came on. When the green light, or all lights, came on they remained on until the subject had time to begin reading the next question, in contrast with the briefer flash of lights provided previously.

Only fast-correct or slow-incorrect responses adjusted timing cycle duration, the former shortening it and the latter increasing it. Previously, slow-correct as well as incorrect responses increased it.

Revised Procedures for Speed and Correctness Reinforcement with Fixed Response-Time Criteria and Added Stimuli (Group 8)

Conditions were the same as for subjects of Group 7 except that for subjects of Group 8 cycle size was not changed by performance.

Revised Procedures for Speed and Correctness Reinforcement with Fixed Response-Time Criteria, Added Stimuli, and a Separate, Concurrent Correctness Contingency (Group 9)

This condition was the same as for subjects of Group 8 except that points which were each worth one cent registered on two counters. Fast, correct responses incremented one counter, and correct responses, regardless of response times, incremented the other.

The summary of the nine conditions also shows the three phases of the experiment: baseline, experimental treatment, and posttest. Subjects merely wrote their answers to items of Set 17 during Posttest_A. During Posttest_B subjects first wrote their answers and then responded to the keyboard as they had been responding. There were 142 instructional items in Sets 1 through 4 of Phase I, 469 items in Sets 5 through 16 of Phase II, and 39 test items in Set 17 of Phase III. There were 39 errors possible when items of Set 17 were answered by subjects using the keyboard and 69 possible when subjects wrote their answers.

A total of eighty-eight subjects participated with 21, 10, 9, 10, 7, 6,

SUMMARY OF THE NINE EXPERIMENTAL TREATMENTS

INDEPENDENT GROUP	PHASE I (Pre-treatment)	PHASE II (Treatment)	PHASE III (Post-treatment)	N*
1	Base line	Base line continued	Posttest_A	21
2	Base line	Correct responses reinforced	Posttest_A	10
3	Base line	Fast, correct responses reinforced, cycles based on first estimate, size of cycle adjusted, increased by slow response	Posttest_A	9
4	Base line	Fast, correct responses reinforced, cycles based on first estimate, size of cycle fixed	Posttest_A	10
5	Base line	Fast, correct responses reinforced intermittently, cycles based on first estimate, size of cycle adjusted, increased by slow response, added stimulus near keyboard	Posttest_A	7
6	Base line	Fast, correct responses reinforced, cycles based on first estimate, size of cycle adjusted, increased by slow response, added stimulus near keyboard	Posttest_A	6
7	Base line	Fast, correct responses reinforced, cycles based on second estimate, size of cycle adjusted, increased by incorrect slow response, added stimuli in keyboard	Posttest_B	10
8	Base line	Fast, correct responses reinforced, cycles based on second estimate, size of cycle fixed, added stimuli in keyboard	Posttest_B	10
9	Base line	Correct responses reinforced and also fast, correct responses; cycles based on second estimate, size of cycle fixed, added stimuli in keyboard	Posttest_B	5

*Data from some of these subjects were incomplete.

10, 10, and 5 subjects in Groups 1 through 9. Exceptions to explicit random assignment were the first 10 subjects of Baseline, who were randomly selected from a pool of 30 as the first subjects to participate. Similarly, about one-third of the subjects were assigned to the correctness-only contingency condition (Group 2), before experiments began with Groups 3 through 6.

RESULTS

Figure 1 shows changes in time for the nine groups during the course of their experimental treatment. For each subject the time spent reading and answering each question item plus the time spent viewing each correct answer which followed his answering response was automatically recorded. From each of the three average total item times based upon performances under the experimental treatment, in turn, the baseline average was subtracted. Data from Baseline subjects (Group 1) suggest that, on the average, subjects performed more slowly on baseline items than any of the three blocks of items, even in the absence of an experimental treatment designed to cause them to perform faster. Average item times are nearly the same for each third of treatment, increasing slightly from first third to second third. Averages were based on the following numbers of items for successive thirds: 149, 151, and 189.

When reinforcement was exclusively contingent upon correctness (Group 2) subjects came to spend more time than they had during their own baseline phase. They also increased their time more, relative to baseline, than did the Baseline Group. Subjects of Group 9, who worked under a dual contingency (correctness

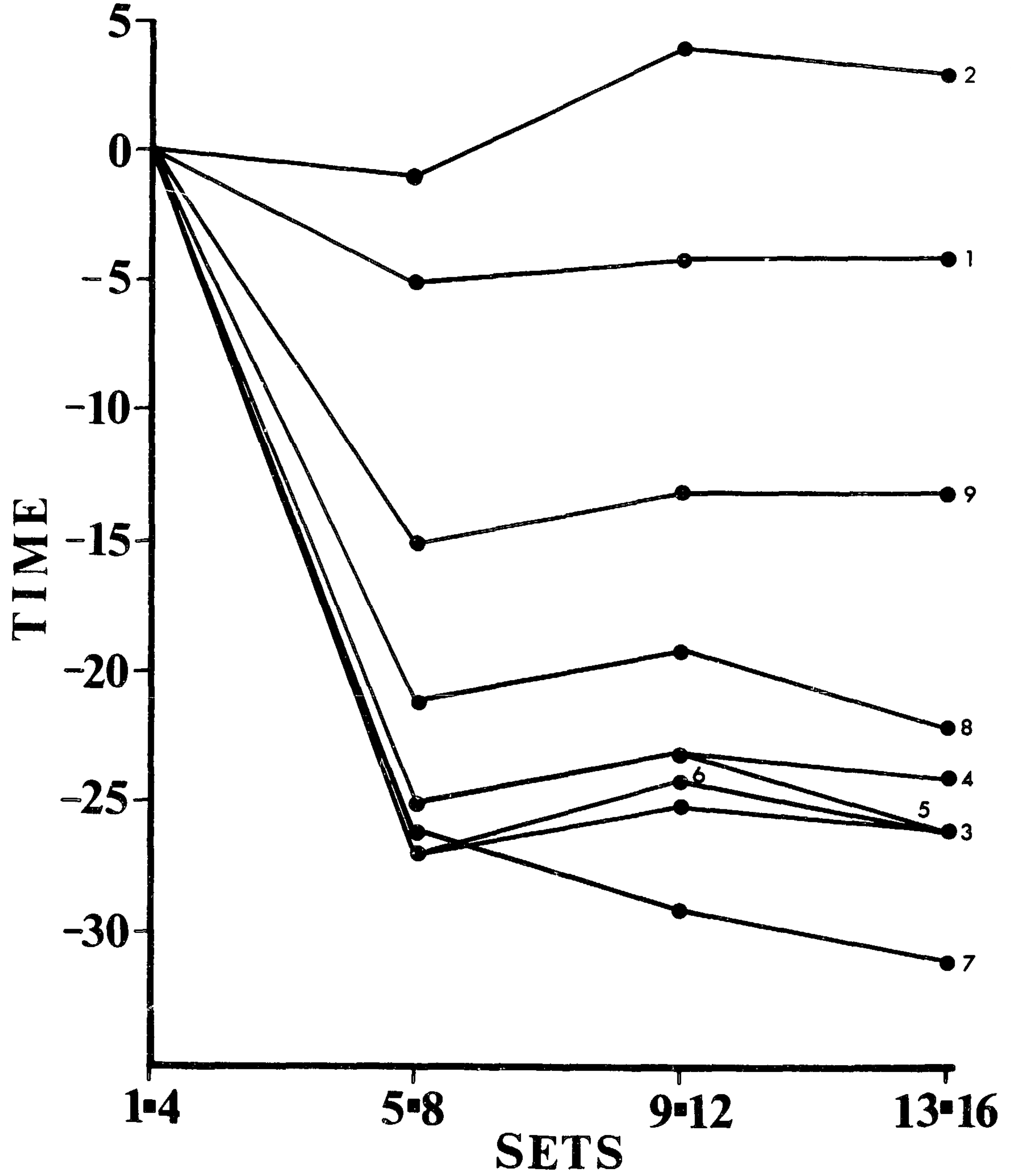


Fig. 1. Changes in average time per item (in half seconds) from baseline during the experimental treatment. Points plotted are for averages for differences between baseline (Sets 1-4) and Sets 5-8, 9-12, and 13-16.

only and speed and correctness) performed faster than subjects under the single correctness contingency (Group 2) and also faster than Baseline subjects (Group 1) but slower relative to their own baseline than any group who performed under a single speed-and-correctness contingency. The continued decrease in time throughout their treatment seems noteworthy for subjects of Group 7 (for whom instructions in their use of the keyboard and their response-time criteria had been revised, stimuli added to the keyboard and the duration of these stimuli prolonged, and time criteria adjusted to be less by fast correct responding and greater by slow incorrect responding). Adjustment of time criteria, primarily downward, resulted in fewer reinforcements for subjects of Group 7 (26 per cent) than for subjects of Group 8 (41 per cent), a difference which was statistically significant at p less than 0.001, according to a t-test. Per reinforcement, subjects of Group 7 responded faster than subjects of Group 8 (whose time criteria were not adjusted) and continued longer during treatment their increase in speed.

Two other ways of presenting relative performance speeds of the nine groups are not shown but will be described briefly. Figure 1 shows in half-seconds how performance changed, as compared with the baseline phase, on the average during the three portions of the experimental treatment. If, instead of differences between baseline and treatment, average times during both baseline and treatment had been shown, it would have been apparent that subjects of Group 7 were slower than those of the other groups during the baseline phase and also slower than any of the groups whose treatment included a fast correct response requirement, including Group 9. The last comment is true because, on the average, subjects of Group 9 performed faster than those of any other group during the baseline phase.

If the percentage change in time per item from baseline during the experimental treatment, instead of change in half seconds, had been shown, essentially the same relationships among groups would have been evident as appear in Figure 1, with one exception. The difference between performance curves for Groups 7 and 8 would be greatly reduced, primarily since that of Group 7 would lie in the midst of curves for Groups 3, 4, 5, and 6. Percentage reduction in time might have aided interpretation of the practical significance of these results. Subjects of all treatment conditions which included a response-time contingency (Groups 3 through 9) reduced their time under treatment by at least 25 per cent (when measured early, middle, or late, as in Figure 1). The greatest reduction, 45 per cent, was by subjects of Group 3 and occurred early, on Sets 5-8.

For statistical analyses, consisting of an F-test and t-tests, the three average differences between early, middle, and late treatment were combined into a single value. The F-test showed a significant overall difference, p less than 0.001. According to t-tests, which were single-tailed tests with p at 5 per cent, Group 1 differed significantly from Group 2. These both differed from all other groups. Group 9 differed significantly from all except Group 8. Groups 3 through 8 did not differ significantly among one another.

Figure 2 shows changes in average percentage of correct responses from baseline during the experimental treatment. There was a slight decrease in percentage correct which did not result from experimental treatment (Group 1).

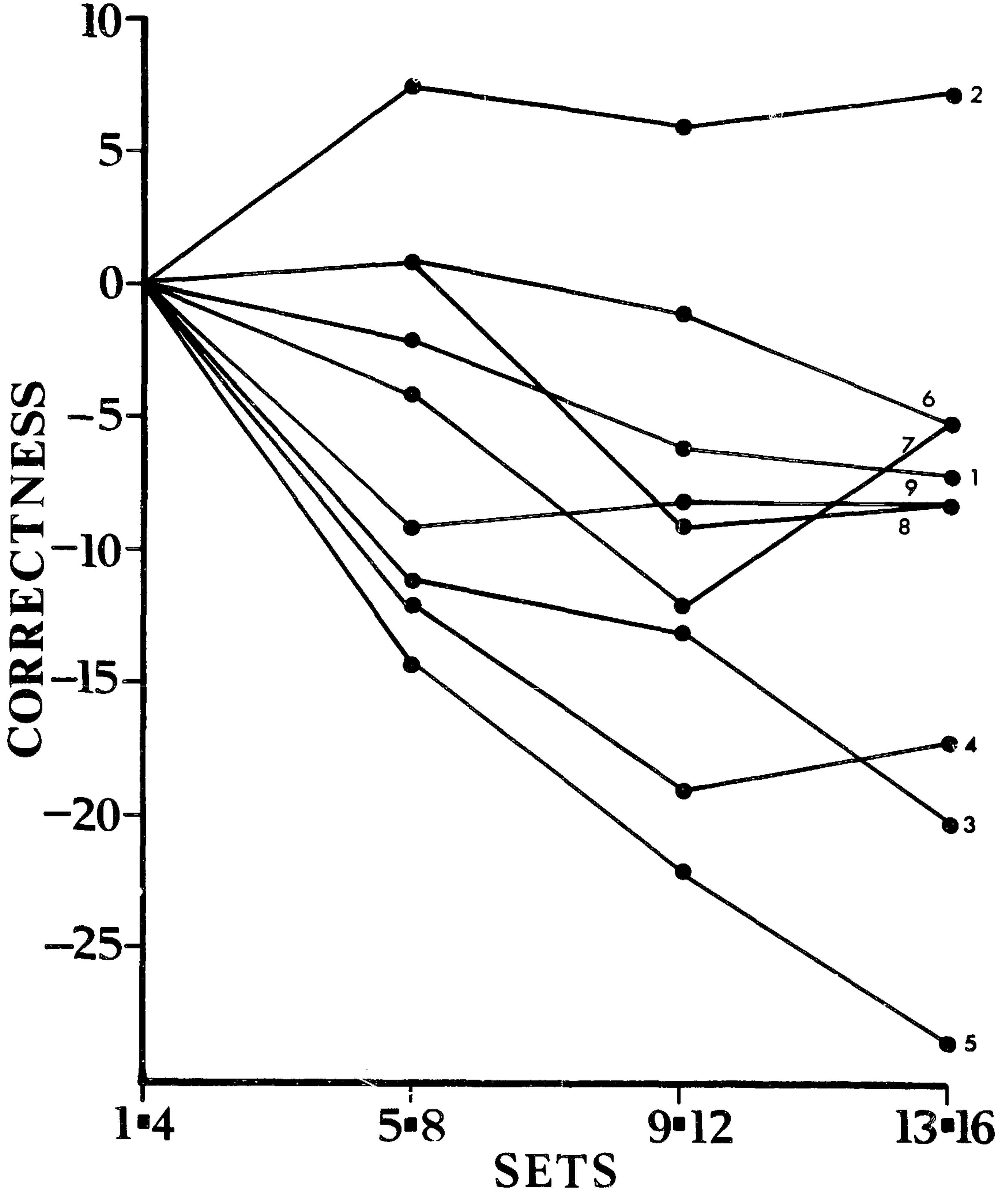


Fig. 2. Changes in average percentage of correct responses from baseline during the experimental treatment. Points plotted are for averages for differences between baseline (Sets 1-4) and Sets 5-8, 9-12, and 13-16.

Curves of Group 1 and 2 show increasing separation throughout treatment, with that of Group 2 rising. This reflects a continued increase in correctness under the correctness contingency. Group 5, whose experimental treatment consisted of intermittent reinforcement of fast-correct responses, is at an opposite extreme from Group 2. The curve from Group 5 reflects a continued decrease in correctness. Groups 3 and 4 both lacked added stimuli near or in the keyboard. Their decrease in percentage correct is only slightly less extreme than that of Group 5. Groups 6 through 9 did not differ appreciably from Baseline (Group 1).

Statistical analyses of percentage correct parallel those of Figure 1. The early, middle, and late treatment average differences from baseline were combined into single values. The t-test was again statistically significant at p less than 0.001. According to t-tests, Group 2 differed significantly from Group 1. Both differed significantly from Groups 3, 4, and 5. In percentage correct, however, Groups 6, 7, 8, and 9 did not differ significantly among themselves or from Group 1.

DISCUSSION OF SPEED AND CORRECTNESS DURING LEARNING

The findings suggest that reinforcement contingencies can be used to affect both speed and correctness of subjects while they perform one kind of complex task. However, the results suggest that the details of how contingencies are implemented may be remarkably important. For example, in the classroom it may seem to make little difference whether answers to test items are made available to students at the end of a test session or a week later. In these laboratory studies, however, the characteristic of having a conditioned stimulus onset at the subject's finger tip, almost unavoidably viewed at the instant of his correct response, may largely explain the superiority of Group 6 over Group 3 and contribute to the superiority of Group 8 over Group 4 in percentage of correct responses. Here, more may be involved than the slight delay in the subject's looking from the keyboard back to the viewer to a counter. In an earlier study, some subjects indicated it was aversive to be speeded up too much. They resorted to avoiding the positive reinforcing stimulus to limit the control it exerted. When points registered only on a counter, the subject sometimes reported that he quit looking at the counter since it sometimes made him nervous. In order to ensure that the onset of a positive conditioned reinforcing stimulus promptly followed each response which satisified the reinforcement contingency, a light was connected with the points counter. Apparently this was an imperfect precaution. The subject could largely miss the flash of the light by a blink or averted glance away from the light and toward the keyboard. In the present study, an additional light in unavoidable view at the moment of the subject's response was associated with a greater percentage of correct responses. Prolonging the duration of these lights may also have been of benefit.

The incompatibility between speed and correctness, which casual experience may lead one to expect, was evident in the performance of subjects of Group 2. Compared with all other groups, these subjects performed slowest and showed the greatest increase in percentage of correct responses, relative to their own baseline. For them, reinforcement depended solely on correctness. That incompatibility between speed and correctness is not inevitable was suggested by performance, shown in Fig. 2, of subjects in Groups 6, 7, 8, and 9. For these groups, positive conditioned reinforcing stimuli were located near, or in, the keyboard. This reinforcement was contingent upon both speed and correctness. While the same was true for subjects of Group 5, for these, backup reinforcement (points worth money) was intermittent. Intermittency evidently increased the incompatibility between speed and correctness.

Incidentally, when backup reinforcement was omitted, on Sets 10 and 11, (for Group 5) the conditioned stimulus near the keyboard continued to be contingent upon fast correct responding. The experimenters' concern was that, if all sources of conditioned reinforcement were abruptly removed, this would be likely to have varied effects on subjects. Some might react as if the equipment were not operating properly; of these some might leave the room seeking assistance, etc. The error in design of the experiment seems to have been in the opposite direction. Extinction, the discontinuation of points reinforcement, did not have an obvious effect, leaving the question open as to whether it would have eventually been different in its effect on subjects of Groups 5 and 6 if it had been prolonged. Intermittency was expected to be slower in establishing a performance change. However, once the change had been established, the new characteristic should be lost more slowly as a result of intermittency.

Performances of Groups 6, 7, 8, and 9 show that, with properly implemented procedures, the time spent on a learning task can be decreased without concomitant decrease in the percentage of correct responses. There was one other source of savings in time which was not related to the reinforcement contingency treatments of these groups. Instead it was related to the abbreviation of response requirements. The procedure of requiring only a single response to the keyboard made it possible for subjects to respond faster than they would have had they written their answers. Even if a computer were used to evaluate answer correctness, requiring the student to type out complete answers would be costly in terms of the student's time, the complexity of computer programming, and the equipment required.

In a later section, performance on the posttest, which followed the experimental treatment, will be considered. There were large differences among subjects in the number of their posttest errors. A satisfactory analysis required that influences on posttest performances which were not related to experimental treatments be taken into account.

POSTTEST ERRORS

Predictor variates which had high validity for the criterion of number of posttest errors were needed. While any high-validity variates might be useful, it best served the purposes of this study to consider those which would be generally available. Variates were needed which would be available under conditions in which findings would be applied. Those derived from performances of subjects early during their work on the learning task would seem especially useful, particularly if these were not tied to unique characteristics of the task or of subjects.

Previous experiments had shown several statistically significant relationships between response times on instructional items and the number of later posttest errors, i.e. errors on Set 17. This had been true of response times on 6 or so items within the first set at the beginning of baseline. In multiple regression analyses response times from items 2, 3, 22, and 48 (Set 1) were used to predict the number of posttest errors. With complete data available from 70 subjects, from within Groups 1 through 8, a multiple R of 0.61 was obtained. Its p-value was less than 0.01. Subjects who spent more time on these four items made greater numbers of posttest errors.

The predicted number of errors was established for each subject. The difference between predicted number and actual number was obtained for each of the 70 subjects. A simple one-way analysis of variance indicated statistically significant differences among treatment groups: $F = 6.17$, $df_1 = 7$, $df_2 = 62$, and p was less than 0.01. All subjects within Groups 7 and 8 made fewer errors than had been predicted. This was not true of any of the other six groups. Also, in actual number of posttest errors, each subject in these treatment groups made fewer than the average number of errors based upon the total number of subjects.

The multiple regression analysis was carried a few steps further. Response times on item 22 of Set 1 were dropped from the analysis, with little reduction in the value of the multiple R. The data of the 70 subjects were divided into two samples based on 35 subjects each. The value of the multiple R for one sample was 0.61, and for the other it was 0.52. Both values were statistically significant at p less than 0.05. Multiple R's remained significant even when recalculated with interchanged beta weights.

Predicted numbers of posttest errors based upon the times subjects spent on items 2, 3, and 48 of Set 1 was one of nine predictor variables in a further analysis. Other variables were membership in treatment group and the sex of the subject. Of all of these variables, response times on the three items was the most powerful predictor of the number of posttest errors. Presence or absence of membership in treatment Group 8, then Group 7, came next, followed by the sex of the subject. For all nine variables the value of the multiple R was 0.78, based on the data of the 70 subjects.

KEYBOARD RESPONDING COMPARED WITH ANSWER WRITING

Procedures of the present study made possible a one-letter keyboard response to items which originally had been designed to occasion the writing of at least one word, and often several. In the present experiments, keyboard responding made possible the evaluation of both the speed and correctness of responses. Some subjects first wrote their answers, then responded using the keyboard. Before considering how well the two modes of responding agreed, some of the practical and theoretical issues will be considered.

Testing, or a mixture of testing and instruction, may be accomplished with multiple-choice items for the sake of automatic evaluation of answer correctness. However, Skinner (1958) pointed out the hazards of using such items. The student may forget the correct answer but remember one of the plausible wrong answers. In the sense of Skinner's concern, the housewife, instructed by multiple-choice first-aid items, might remember flour, instead of soda, for putting out a stovetop grease fire. Skinner offered a theoretical objection to the use of multiple-choice items.

He could also offer a practical objection. It may be more difficult in time and effort for the instructor to formulate good multiple-choice items. Implausible alternatives are not true alternatives. Items may be constructed more on the basis of the availability of a sufficient number of plausible alternatives than on the basis of the item's bearing on important instructional objectives.

In classroom applications, machine scoring of student quizzes may make it possible for the instructor to be able to quiz more often, for an instructional as well as evaluational purpose. The advantage of machine scoring may be attained in ways which do not require the typical form of multiple-choice items. For example, students may follow coding rules and compose their answers on cards used with IBM Port-A-Punch Boards. Another example would make use of a mark-sense answer sheet, say an eight-column sheet. Students would be provided lists of answer words, arranged in alphabetical order. Each answer word would have one of eight column numbers associated with it. In answering a given item the student would decide upon a particular word, find it in the list, then darken the column corresponding to the number associated with that word.

A common objection to the last procedure is that the student may work with an incorrect answer but be credited with a correct answer because his incorrect choice had the same column number as the correct choice. The instructor who regularly uses four-alternative multiple-choice items with his students may resist using the eight-column procedure, although his objection may be overcome by pointing out he could take pains to avoid assigning the same column number to answer words of near-opposite meaning.

The procedure of requiring the third letter of the last answer word could also be used in the classroom. Perhaps the instructor should ask the student to record his printed letter answers first, then later code them onto a card or a machine scorable answer sheet. A trial use of the procedure would be needed to acquaint students with it. Even with college-level students, careful explanation would be needed. The possibility of a student's receiving correct-answer credit for a wrong answer with the same third letter would need to be discussed, but put in perspective by comparing chance expectancy here with chance expectancy for multiple-choice items. Precise expectancy value would depend on the particular application. Of course, not all 26 letters are equally likely to be the third letter of most words. Therefore, chance expectancy, while lower than for traditional multiple-choice items, would not reach its lowest possible value, of one twenty-sixth.

Degree of agreement between the number of errors students made when writing their answers compared with the number of errors during keyboard responding (with keyboard responses given immediately after writing) was considered for its bearing on the validity of keyboard responding.

Relationship between Write-Answer and Keyboard Errors

Subjects of Groups 7, 8, and 9 first wrote their answers to the posttest items of Set 17. Then, instead of using a handswitch to advance the filmstrip, they responded to the lettered key corresponding to the third letter of their answer, or according to the special instructions of a given item.

Results were considered in terms of rank-order correlation. For each subject the x-value was the number of errors he made on the written posttest, for which the maximum was 69. The y-value was the number of errors made during his 39 keyboard responses. Of the 25 subjects who responded both ways, valid data of both kinds were obtained from 19. The rank-order correlation was positive and substantial: 0.84, p less than 0.01.

CONCLUSIONS

It was possible to decrease speed of responding without increasing errors significantly by adding speed-and-correctness contingencies of reinforcement to an ongoing task. The onset of an unavoidable conditioned reinforcing stimulus at the instant of correct response seemed important in preventing a decrement in correct responding. Possibly prolonging its duration was of benefit. The two groups of subjects for whom the duration of this stimulus was prolonged, time criteria were revised, and initial instructions in the task were improved, made fewest posttest errors and fewer than were predicted for each of these subjects. Procedures were automated, and the procedure for abbreviating the overt response requirement was investigated. Subjects who most often wrote correct answers, most often pressed the correct key of the keyboard. This demonstrated the validity of a procedure which spared the subject time and effort in responding, while permitting effective monitoring of the speed and correctness of his responses.

The experimental treatment which adjusted response-time criteria according to the subject's success in satisfying each criterion provided for prolonged performance improvement.

THE CUEING OF SPEED REQUIREMENTS

Lloyd O. Brooks

ABSTRACT

Previous experiments demonstrated increases in speed without impairment of correctness when reinforcement was made contingent upon speed and correctness and when conditioned reinforcement was immediate, prolonged, and unavoidable. Whether response-time criteria remained fixed or were adjusted by performance was not critical to previous findings, although adjusted contingencies seemed to prolong gains in speed. In these earlier experiments subjects had no indication of how long they would be allowed for fast responding. Time criteria typically differed from item to item.

In the present experiment, the same items and posttest were used, the same baseline-treatment-posttest design was followed, and two ways of cueing time requirements were added to the superior previous conditions. Since fixed or adjusted time criteria had both been effective, and a single procedure was needed for investigating cueing, adjustment of time criteria by performance was minimal. Every fifth item was an opportunity for adjustment (and the sixth, in case of two consecutive fast correct reponses).

Cueing resulted in performance significantly faster than for Baseline subjects but not as fast as obtained previously without cueing. Correctness during learning was reduced under cueing. Posttest errors were slightly, but, statistically, insignificantly, greater than for Baseline subjects.

There was no appreciable difference in any performance measure between cueing which indicated the remaining last unit of time and cueing which indicated at the beginning of each item how many time units would be allowed as well as units remaining throughout the interval.

THE CUEING OF SPEED REQUIREMENTS

Lloyd O. Brooks

American Institutes for Research, Palo Alto, California

The experimental treatments which were described earlier, in the main section of this final report, did not provide subjects any indication of time criteria. Correct responses to some instructional items were reinforced as fast even when response times were several times as great as they had been on other items. Time criteria were derived from average response times, and these typically differed from item to item. It seemed possible that subjects would more often respond correctly if the time to be allowed them were somehow indicated. Stimuli which indicated time criteria values would serve as prompts, or cues. A stimulus which indicated that a brief interval was to be allowed was expected to occasion a faster response than one which indicated a greater interval. Informed more precisely of each momentary temporal requirement, the subject might avoid any excessively fast responding which came at the expense of response correctness. Through an improved distribution of his time on components of the task the subject might also come to make fewer posttest errors.

As considered up to this point, the cueing of speed requirements would oppose the development of faster performance. Such cueing would guide responding toward the slower limits of speed contingencies. However, response-time criteria could be adjusted to become more demanding on the basis of fast correct responding. As a concomitant of a pace which was more appropriate to his task, the subject could acquire a quickened pace.

Cues could indicate the time to be allowed at the beginning of each new item and indicate how many units of time remained as each unit elapsed. How effective such cueing would be in terms of the subject's speed and correctness during learning and in terms of his posttest errors would be essentially an empirical question. Possibly cueing could indicate too much. Cueing could compete with important charecterisitics of the task for the subject's attention. For example, residual sequential effects of a series of instructional items dealing with a common topic might be reduced, or lost, as the subject became preoccupied with the "clock." On the other hand, the subject might develop skill in making use of cues and learn to ignore them when they were likely to interfere.

Cueing which served to warn the subject when the end of allowed time was near, by indicating each final portion of time criteria, would seem to compete less with the task for the subject's attention, since cue onset would be limited to once per item and cue duration would be brief relative to total performance. However, emotional or other effects would be possible. For example, the subject who was unable to respond correctly before the warning cue could be punished by its onset. If, on the other hand, he was ready at that moment to respond correctly, the onset could act as a positively reinforcing stimulus, through indicating not only that the subject's correct response met the time criterion but, also, that it did so with margin to spare.

In addition to the empirical question concerning the relatively direct effects of cues was the question of whether early speed requirement cues would benefit later performance which took place after these cues were no longer used. Possibly cueing would have a lasting benefit even if the immediate direct effect was somewhat disruptive.

PROCEDURE

Design

Subjects were randomly assigned to two independent experimental treatment groups. They participated in the experiment individually. Each worked under baseline while completing the first four sets of items. Materials were those used in the previous work, described earlier in this final report. Again, the experimental treatments were in effect on Sets 5-16. Set 17 served as the posttest.

The experiment was designed primarily for a comparison of two cueing procedures. However, the performances obtained under these cueing procedures could also be contrasted with performance under each of the nine early treatment and baseline conditions. In order to provide for these possibilities, the cueing experiment followed the same base-line-treatment-posttest sequence as earlier treatments, as suggested above. In addition, subjects were recruited in the same manner from the same population.

Subjects

Subjects were recruited as paid volunteers from Foothill College. As initially formed, each of the two treatment groups consisted of five male and five female subjects. Complete data were obtained from nine subjects of one group and eight of the other.

Experimental Treatments

The two procedures for cueing speed requirements were explained to both groups of subjects by a single set of instructions. These informed the subject that a white light would come on near the end of the time he was to be allowed in which he could earn a point (worth two cents) by responding correctly. After the white light went out the opportunity for earning a point would be over. As in previous experiments, the subject worked at an hourly rate of $1.50 plus earnings from points. Again he was assured of $2.00 per hour, if points and the lower hourly rate together amounted to less than $2.00 per hour.

Multiple-light cueing. Two horizontal rows of lights, consisting of 17 each, were located above the viewer screen. A white light was below and to the right of the last light of the bottom row. Lights of the two rows formed pairs: each light of the top row was directly above one in the bottom row.

As the subject advanced the filmstrip, bringing a question into view, one of the 17 lights in the top row became illuminated and remained on during that question, until the subject responded. This light in the top row corresponded to the number of timing cycles the subject was to be allowed for that question. For example, if the third light from the left were on, the subject would be allowed 3 timing cycles in which to answer correctly and have his response reinforced. After a first timing cycle, which was of five seconds duration (at the beginning of each session of the experimental treatment but could become adjusted larger or smaller beginning with the subject's response to the first item, as will be explained later in this treatments section), the first light on the left of the lower row came on during one cycle. This first light then went out. Subsequently the second light came on for the next cycle, went out, was followed by the third light, etc. During the last timing cycle, before the time allowed for fast correct responding elapsed, the light on in the bottom row would match the light on in the top row directly above it and, at the same time, the white light at the lower right also came on. The white light went off as soon as the time allowed ended.

Single-light cueing. For this second experimental treatment, only the white light at the lower right continued to come on during the last timing cycle. When only the white light was used, the two rows of lights above the viewer were covered and out of view.

Cycle duration adjustment. Every fifth instructional item was an opportunity for the subject to increase or decrease the duration of the timing cycle. It was decreased on the basis of a fast correct response and increased when the subject responded slowly and incorrectly. When the subject decreased cycle duration it was possible to adjust the cycle downward once more on the next item. Thus, there were two opportunities for adjusting cycle duration to require faster responding for every opportunity for adjusting it to permit slower responding.

Duration of cueing. For six subjects in each treatment group the cueing of speed requirements continued throughout Sets 5-16. For other subjects cues were removed after Set 12.

Treatment duration was also limited somewhat by the necessity for exploratory investigation of how the subject's performance was to be used to adjust response-time criteria. The beginning portion of treatment for two or three subjects was complicated by brief trials of adjusting procedures before the procedure finally selected was identified. Both fixed and adjusted contingencies had been effective previously and a single procedure was needed. Therefore, a procedure which was not completely fixed, but which did not adjust time criteria as greatly as before, seemed the best choice. Some different ways of attaining this choice were tested briefly.

RESULTS

Figure 1 shows average time per item under the experimental treatment relative to baseline. Curves are for (1) six subjects under single-light cueing throughout, (2) six subjects under multiple-light cueing throughout, and (3) five subjects (two from single-light and three from multiple-light cueing conditons) who had cues removed during the last third of their treatment. There was no apparent change resulting from cue removal, and the differences among the three curves are not of practical or statistical significance. A comparison of Figure 1 of this section with the Figure 1 of the earlier, main section of this report shows the subjects who had cueing of speed requirements to be among the slowest of all speed-requirement treatment groups. However, the difference between speed-cued subjects, who are faster, and Baseline subjects appears to be of practical significance.

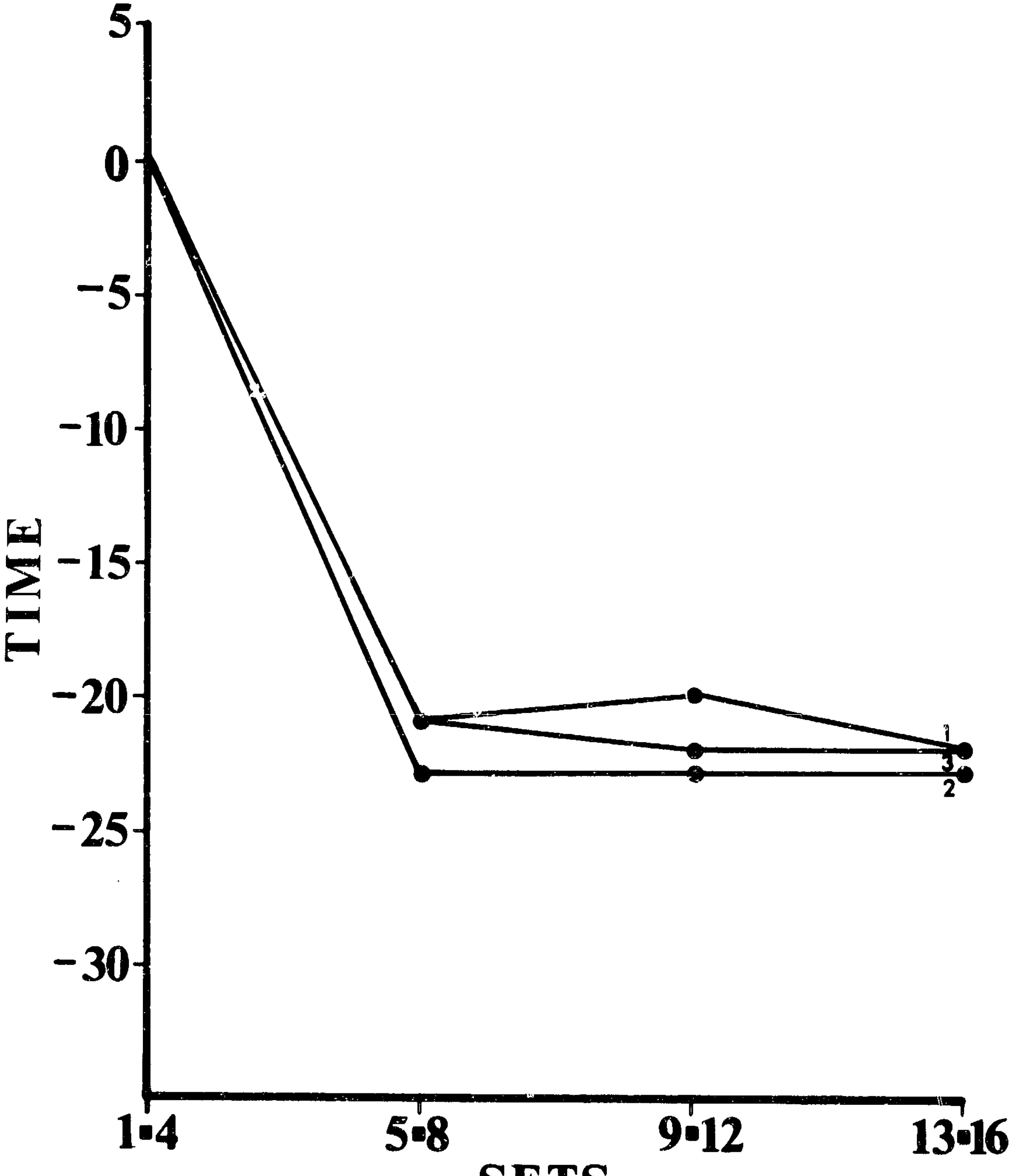


Fig. 1. Changes in average time per item (in half seconds) from baseline during the experimental treatment. Points plotted are for averages for differences between baseline (Sets 1-4) and Sets 5-6, 9-12, and 13-16.

Figure 2 of the present section shows fewer correct responses under cueing treatments. A comparison of these results with Figure 2 of the earlier section shows only one experimental treatment, the one involving intermittent reinforcement, to have reduced correctness to a greater degree. The differences among patterns in the three curves, or their separations, are slight enough to be attributable to chance.

Single-light compared with multiple-light cueing did not appear to make a difference in terms of the number of reinforced fast correct responses. For example, the approximate median value was 26 per cent of responses reinforced as fast and correct for the eight subjects of the single-light cue group and 23 per cent for the nine subjects who had multiple-light cues.

Subjects of the single-light cue condition averaged 31.16 posttest errors (based upon written answers on Set 17, with 69 errors possible) while those of the multiple-light cue condition averaged 38.16. This difference was not statistically significant. Respective variances were 151 and 37. The difference in variances, while appreciable, was not statistically significant. These results were based on the six subjects of each group. When subjects were included who had cueing discontinued during the last third of their treatment, differences were even less between cueing treatment groups for averages (33.12 compared with 36.77) and variances (124 compared with 46).

The average number of posttest errors for subjects of the earlier Baseline group was 30, the variance 132, and 19 subjects provided these data. For a pooling of the two full-treatment cueing groups the mean for these 12 subjects

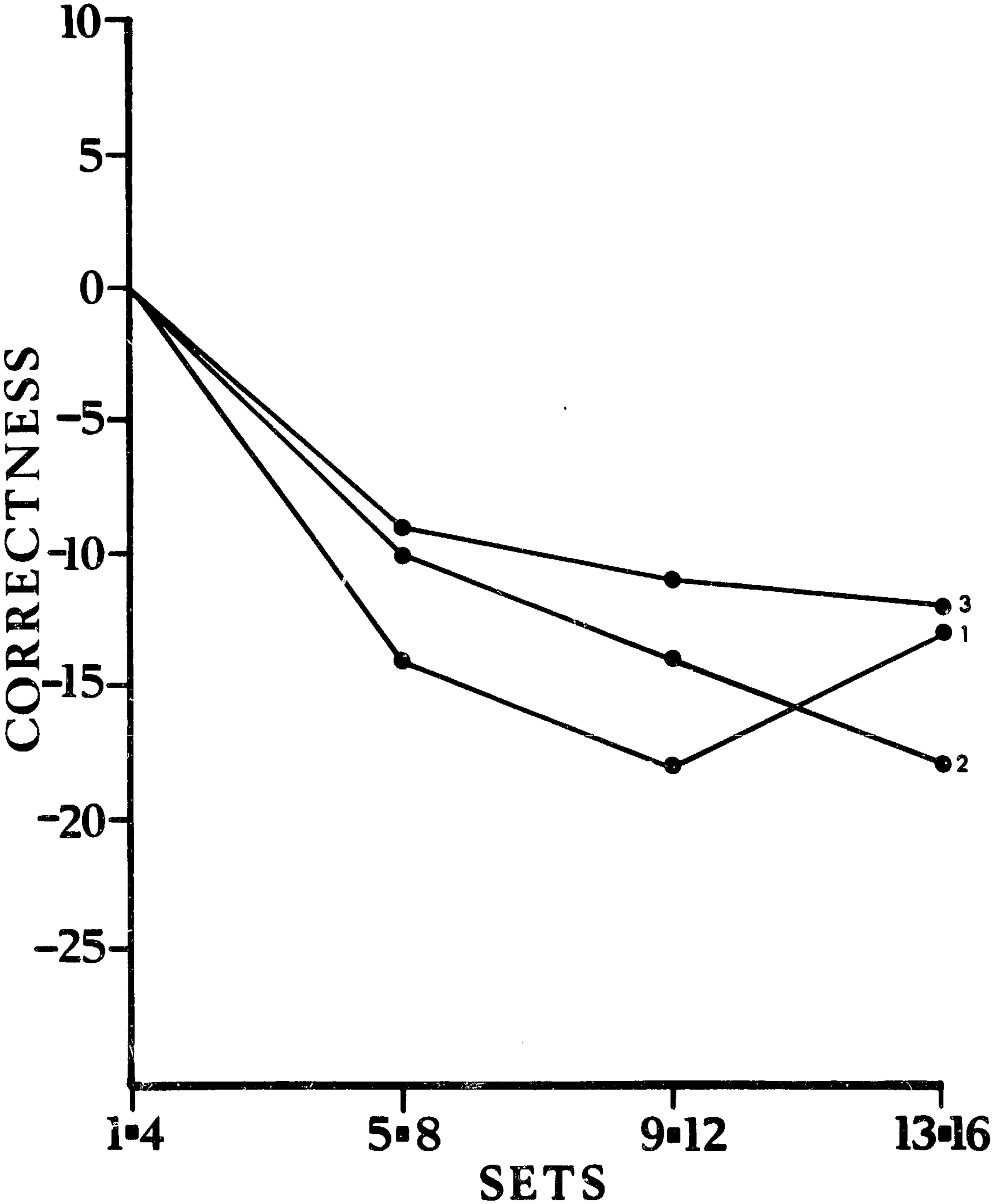


Fig. 2. Changes in average percentage of correct responses from baseline during the experimental treatment. Points plotted are for averages for differences between baseline (Sets 1-4) and Sets 5-8, 9-12, 13-16.

was 35 and the variance 98. While the subjects whose speed requirements were cued made more errors, their average did not differ significantly from that of Baseline subjects.

The number of posttest errors was predicted for each subject on the basis of response times on items 2, 3, and 48 of Set 1. The weighting procedure was approximately the same as used in the multiple-regression analysis, reported earlier. All 9 of the subjects under the multiple-light cueing condition made a greater number of errors than was predicted, while 6 of the 8 subjects in the single-light cueing condition made more errors than were predicted. The overall average for the two cueing groups combined was 25 errors. Hence, the subjects of this experiment before treatment seemed destined to perform slightly better than Baseline subjects.

Informal res[illegible] were the comments of subjects. Some volunteered comments which indicated that they found time requirements to be unpleasant. One subject said the points worth money upset her by reminding her of a time pressure. Others reported trying to ignore light cues in order to concentrate fully on subject matter and on answering correctly. One subject mentioned a strong urge to give an answer, even a wild guess, when the white light came on. For another subject, outspoken objection to a time requirement was associated with particularly fast correct responding and with few posttest errors.

DISCUSSION

None of the results show the cueing of speed requirements to be a promising avenue for establishing greater speed and correctness. Preliminary observations which led to the plan of this study were somewhat misleading. These observations were based on the response to single-light cueing by subjects who were more familiar with the learning task. On an entirely new task responses to cueing were more varied.

Under the cueing conditons of the present study, subjects did show an increase in speed, relative to their own baselines. Also, following their cueing conditions subjects did not make a significantly greater number of posttest errors than subjects of the Baseline group. However, the increase in speed under cueing conditions was less than obtained under most experimental treatment conditions which involved speed and correctness reinforcement contingencies. Only the experimental treatment which included intermittent reinforcement was associated with a greater decrement in correct responding during learning.

Speed requirement cues were added to conditions which had been especially effective. The limitation to effectiveness imposed by cueing may be tied in with the importance of conditioned reinforcement, near or in the keyboard, and with the comments volunteered by subjects, in this and in previous experiments. Reinforcement based on speed was partly responsible for subjects responding faster than was pleasant to them. One way the subject could control his own behavior was by avoiding stimuli associated with reinforcement based on speed. Conditioned positively reinforcing stimuli near or in the keyboard were not easily ignored, since the subject was forced to look toward these visual stimuli in order to operate the keyboard.

To a greater extent speed requirement cues could be disregarded. Cues could further load the task. They could influence the subject to make an (unpleasant) effort to be fast, an effort he would not make in the absence of such stimuli. At the same time, they could add complexity, causing him to be less likely to respond correctly. A mixture of disregarding speed requirement cues and of using them to the detriment of performance seems to have occurred.

During earlier experiments, some subjects often missed meeting time requirements by a small margin. It seemed possible that had time requirements been indicated they would have managed the extra effort to meet requirements. From reports of subjects in the present experiment, however, it is likely that such subjects would have attempted to take speed cues into account before performance stabilized. Largely unsuccessful early use of speed cues probably would have hindered them in their later attempts, when successful use would have been more likely. Perhaps an experimental treatment group should have been included that worked under a speed and correctness contingency a while before speed-requirement cues were introduced.

REFERENCES

Berman, M.L. Experimental explorations in programmed instruction and objective testing measures. V.I.B. Project, Paper No. 2, Department of Anthropology, Arizona State University, August 1966, Tempe, Arizona.

Berman, M.L. Shaping performance on programmed materials. *Psychological Reports,* 1967, 21, 29-32.

Berman, M.L. Simultaneous maximization of speed and accuracy on programmed materials through the use of adjusting conjunctive schedules of reinforcement. *National Society for Programmed Instruction Journal* (NSPI), May 1968, 7, 15-18.

Blough, D.S. & Millward, R.B. Learning: Operant conditioning and verbal learning. *Annual Review of Psychology,* Palo Alto, California: Annual Reviews Inc., 1965, 16, 63-94.

Brooks, L.O. Shaping faster question answering. Palo Alto, California: American Institutes for Research, June 1965, Technical Report AIR-C28-6/65-TR No. 10.

Brooks, L.O. Response time during instruction. *Perceptual and Motor Skills,* 1967, 25, 200-204.

Ferster, C.B. & Skinner, B.F. *Schedules of reinforcement.* New York: Appleton-Century-Crofts, 1957.

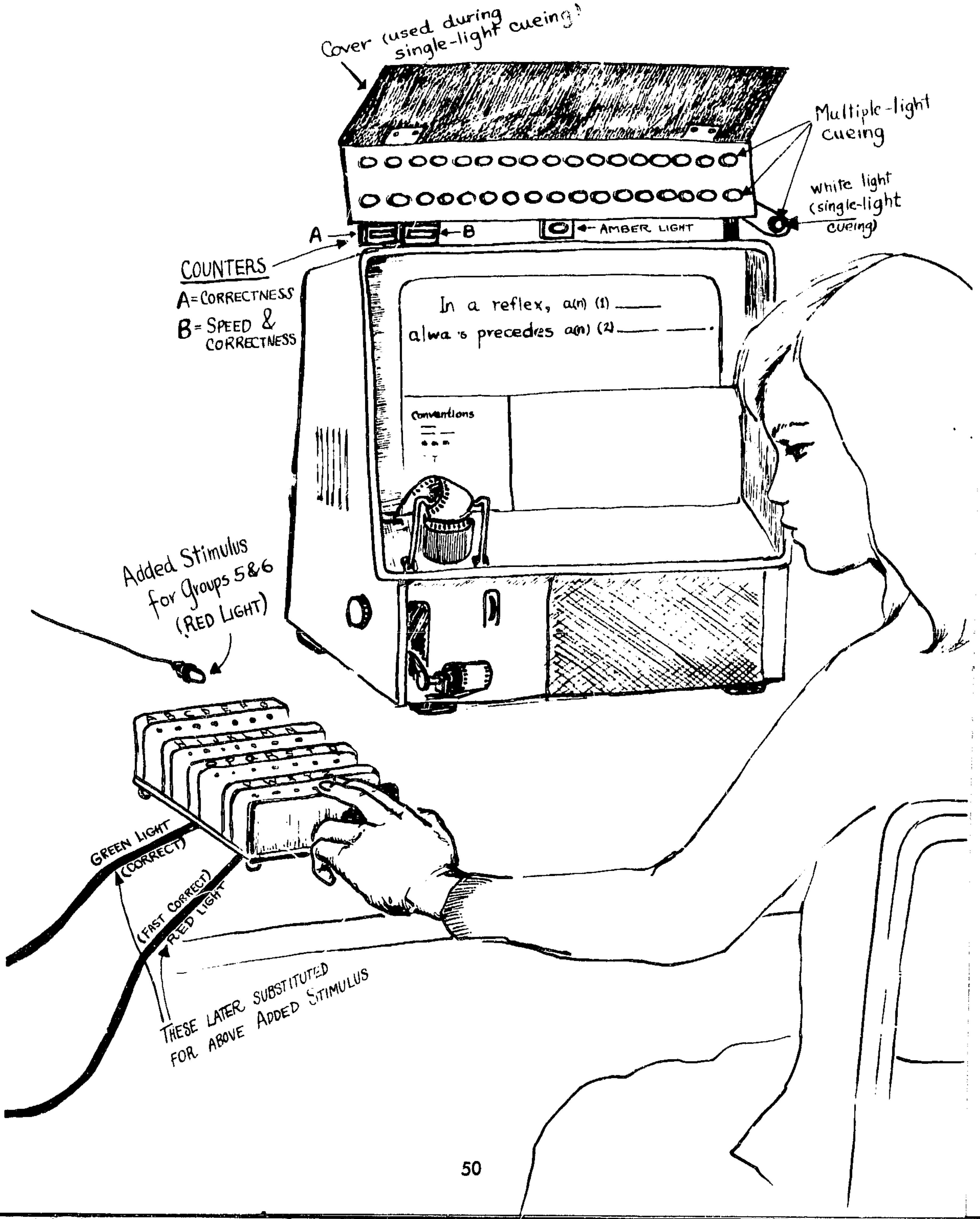
Holland, J.G. & Skinner, B.F. A program for self-instruction: the analysis of behavior. New York: McGraw-Hill, 1961.

Logan, F.A. Micromolar behavior theory and performance speed in education. Harvard Educational Review, Spring 1963, 33, 178-185.

Skinner, B.F Science and human behavior. New York: Macmillan Co., 1953.

Skinner, B.F. Teaching machines. Science 128: 969-77; October 24, 1958.

Appendix



Cover (used during single-light cueing)
Multiple-light cueing
White light (single-light cueing)
A
B
AMBER LIGHT
COUNTERS
A = CORRECTNESS
B = SPEED & CORRECTNESS
In a reflex, a(n) (1) ____
alwa s precedes a(n) (2) ____ ____.
Conventions
Added Stimulus for Groups 5&6 (RED LIGHT)
GREEN LIGHT (CORRECT)
RED LIGHT (FAST CORRECT)
THESE LATER SUBSTITUTED FOR ABOVE ADDED STIMULUS