N66-15394
(ACCESSION NUMBER) (THRU)
54 1
(PAGES) (CODE)
CR 69356 04
(NASA CR OR TMX OR AD NUMBER) (CATEGORY)
FACILITY FORM 602

EXPERIMENTAL STUDIES OF PERCEPTUAL PROCESSES

Progress Report, September 1965

SECTION ONE

Covering Section, Budgets

SECTION TWO

I. Complex discriminative behavior in chimpanzees.

II. Fixed ratio reinforcement of large units of behavior.

III. Deferred reinforcement.

SECTION THREE

I. Signal detection psychophysical research.

II. Signal detection in the design of operant experiments.

III. Other research.

SECTION FOUR

Neurobiological program.

Performed under Grant NsG-450 from the

National Aeronautics and Space Administration

INSTITUTE FOR BEHAVIORAL RESEARCH

SILVER SPRING, MARYLAND

GPO PRICE $ _____

CFSTI PRICE(S) $ _____

Hard copy (HC) 3.00

Microfiche (MF) .50

ff 653 July 65

EXPERIMENTAL STUDIES

OF PERCEPTUAL PROCESSES

Progress Report, September 1965

SECTION TWO

Prepared by:

CHARLES B. FERSTER, Ph.D.
JOHN RANDOLPH, Ph.D.
CLIFFORD HAMMER, B.A.

I. Complex discrimination behavior in chimpanzees 1

II. Fixed ratio reinforcement of large units of behavior 25

III. Deferred reinforcement 28

Performed under Grant NsG-450 from the

National Aeronautics and Space Administration

INSTITUTE FOR BEHAVIORAL RESEARCH

SILVER SPRING, MARYLAND

I. COMPLEX DISCRIMINATIVE BEHAVIOR IN CHIMPANZEES

A. Introduction

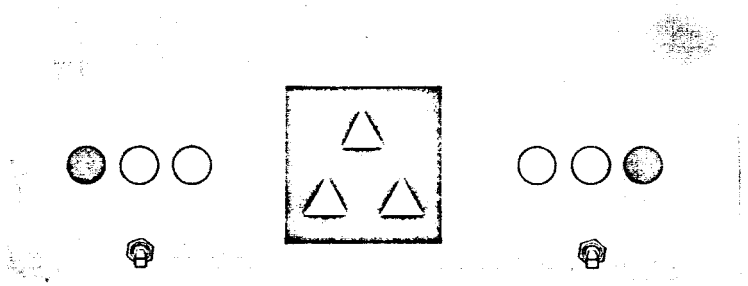
The experiments described in this section were performed by Charles B. Ferster, Ph.D., and Clifford E. Hammer, Jr., B.A.

The work described in last year's progress report consisted of the development of a special kind of environment and concomitant behavioral techniques necessary for experiments on complex repertoires. In our last report we described the kinds of complex cognitive repertoires which had been developed, as well as the techniques and methods for producing them. The work accomplished this year has consisted largely of experiments which take their start from these cognitive performances. We have used the complex performances developed over the past three years as tools for evaluating variables such as schedules and kinds of reinforcement, and also special contingencies to control error rates and the overall rate of responding.

B. New Kinds of Stimulus Control

Some of the experiments performed this year were obvious extensions of our earlier work (Ferster, 1960; Ferster & Hammer, 1964). The later experiments were specifically designed to study the developments and maintenance of relationships between the chimpanzees' behavior, the varying numbers of geometric forms and the binary numbers up to 7. The most complex matching-to-sample task that we developed was one in which the chimp matched binary numbers to a stimulus consisting of a display of geometric figures, varying in form and spacial distribution. The

chimpanzee selected one of two binary numbers which corresponded to the number of items in the sample. Figure 1 is a diagram of the apparatus for the matching-to-sample experiment.



A number of geometric forms of varying shapes and spacial distributions are projected in the large window. Each of the groups of three lights on the sides represents a binary number, the unlighted lamp is a zero and the lighted lamp a one. Reinforcement occurred on one of the two keys below. When the animal pressed the key under the binary stimulus which corresponded, via the rules of binary arithmetic, to the number of geometric forms in the window above the response was reinforced.

Our experiments discovered techniques for the development of the complex performance and we described stable and accurate performances (baselines) with problems consisting of all combinations of the stimuli.

The acquisition of new stimulus control was orderly and predictable after we had discovered effective ways of generating the performance. Consequently, even the performance, as the chimp acquired the new discriminations, could be used as a baseline.

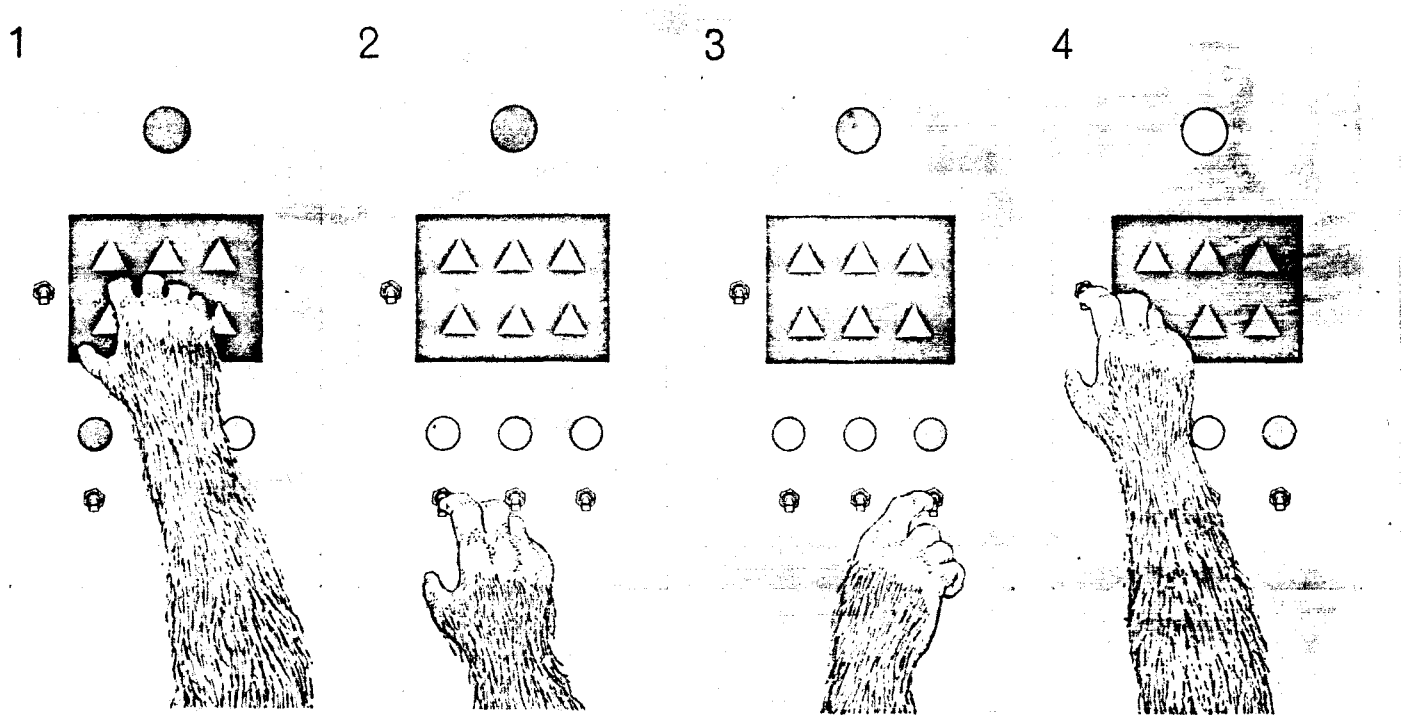
New kinds of stimulus control were developed in other experiments in which the animal actually produced the binary stimulus, as opposed to simply choosing one of two as in the matching-to-sample procedure described above. This repertoire is functionally analogous to writing, since the animal's behavior leaves a visual record. In these experiments, the chimpanzee was presented a sample consisting of geometric forms, as with the matching to sample. Reinforcement occurred for responses on a separate key, the register key, only if the animal had first "written" the binary number which corresponded with the number of geometric forms according to the rules of binary arithmetic. The animal "wrote" the binary number by pressing three lever switches, each under a light. He could adjust the "on/off" pattern of the three lights indefinitely. The relationships between the decimal numbers 0 through 7, the corresponding binary numbers, and the lights are presented in Figure 2. The majcr contingency in the writing experiment occurred when the chimp pressed the register key on the left side of the panel. If the pattern of light on the binary stimulus display corresponded to the number of geometric forms in the window, pressing the register key produced a secondary reinforcer, a tone or a pellet of food and a tone depending upon the particular conditions of

reinforcement. If the binary stimulus which the chimpanzee wrote did not correspond with the number of geometric forms, the register response produced a time out (a period during which the keys were inoperative and no stimuli were presented), or some other aspect of nonreinforcement.

DECIMAL	BINARY	LIGHTS
0	000	○ ○ ●
1	001	○ ● ○
2	010	○ ○ ○
3	011	● ○ ○
4	100	○ ● ●
5	101	○ ● ○
6	110	○ ● ●
7	111	○ ○ ○

Binary system, in which numbers are represented by only two digits, 0 and 1, is used for chimpanzee arithmetic. The digits are presented as lights turned off (0) or on (1). The table shows the decimal numbers 1 through 7 with binary and "chimpanzee" equivalents.

A representative procedure and illustration of the response panel is given in Figure 3.



A typical procedure is shown for a numerosity problem. The animal touches the numerosity panel to turn on the binary-number lights below it (1). Then it presses keys that adjust the number to match the number of triangles in the sample (2 and 3). When the number appears to be correct, the animal presses the upper key and the reinforcement light comes on (4) for a correct solution.

In one procedure, for example, when reinforcement occurred after a certain number of correct responses, each error reset the count to the beginning so that reinforcement could not occur unless responses were correct successively. Thus, an "error" increased the amount of behavior required per reinforcement.

The procedures for developing the counting performance were quite involved and difficult. At first, we attempted a fading procedure (Terrace, 1963a) starting with a "writing repertoire" in which the chimpanzee copied another binary stimulus located just above the one he had to write. We then tried to fade from the sample binary stimulus to the numerosity stimulus, but these procedures were in general unsuccessful. We finally developed the required performance by first carefully programming the simpler parts of the repertoire, increasing the complexity of the stimuli only when the current performance was well maintained and under close control of the binary stimuli. Otherwise, the possibilities for error would have been so large that the animal would soon lose the repertoire because of too infrequent reinforcement. Had we at this time discovered the procedures by which we built the intra-behavioral chain (described below) the experiment would have been simpler.

C. The Development of an Intra-Behavioral Chain

This experiment developed an analog of the discriminations of verbal behavior. We produced a sequence of responses functionally like the intra-verbal chain which we see when a child counts, as in saying the numbers one to ten in order. In the experiment, the chimp's task was to write, in order, the binary numbers 1-7 (001-111). Beginning with 000 the subject's task was to change the light to 1 (001) and press the register key. The binary number 1 (001) then served as the occasion for writing the number 2 (010) and pressing the register key. This

performance continued until the chimp wrote and registered the number 7 (111).

The degrees of freedom open to the chimpanzee at each stage of the intra-behavioral chain were extremely large. Not only were there eight possible numbers that could be written at any time, the immediate reinforcement history might produce a strong tendency to persist on one or more incorrect numbers. Therefore, very special procedures were necessary to discover how to build this very complex chain.

Many procedures were ineffective, including all variations of fading techniques, where the chimps remained under very close control of even the most minimal exteroceptive cue until the moment that it faded completely. Then the performance broke down. At this stage of the experiment we tried to use the "writing repertoire" under the control of the geometric forms. The animal could write the binary numbers in order so long as we cued this behavior with the appropriate number of triangles or circles. However, again the performance always broke down when the numerosity sample was no longer available. The ineffectiveness of the fading procedures in this, as opposed to other experiments, e.g., Terrace (1963b), is probably a function of the spacial separation of the two sets of stimuli. To transfer control from one stimulus to another may require that they be superimposed on each other.

We were able to build the intra-behavioral chain when we removed the visual numerosity sample altogether, and controlled the animal's behavior only through the binary stimuli. The principle we finally

evolved was to restrict the number of errors as naturally as possible by excluding some kinds of responses.

We started with the final member of the sequence (chain), that is, writing the number 7 (111) from the number 6 (110). For example, in this first step the stimulus confronting the animal was the number 6 (110) and the response required of him before pressing the register key to produce food was a single response on the lever under the right digit, thereby changing the 0 to a 1, and the whole number to 7 (111). The reinforcement, by food, of the register response on the occasion of 7 (111) on the binary stimulus display established it as a discriminative stimulus which provided conditioned reinforcement for appropriately adjusting the keys under the binary stimuli. But we restricted the possible errors which could be made. The first approximation to the final repertoire, guaranteeing a reinforcement of this last member of the chain, was to disconnect the center and left keys so that these lights would not go out even if the keys were pressed. Thus, there was only one possible error, failing to press the third key, and reinforcement always occurred on the occasion of 7 (111). This number as a consistent occasion for reinforcement would differentially reinforce the response required to change 6 (110) to 7 (111). Then we progressively increased the number of errors possible in writing the number 7 from a 6 by allowing more errors in successive stages. After there was a perfect performance in the first stage, a second key was unlocked, and so on. During the approximation to the final performance a food pellet was given each time the chimp got to

the end of the chain. As the chimp completed the chain with fewer unreinforced responses, we shifted to intermittent reinforcement where each completion of the chain produced a tone (conditioned reinforcer) and food was delivered every so many times the chain was completed successfully. The secondary reinforcer, the tone following each correct register response, provides immediate feedback to the chimpanzee, related to the accuracy of his performance, and critical in establishing the repertoire.

In successive stages we then added earlier members to the sequence and it became possible to develop the entire intra-behavioral chain rapidly and in an orderly way. In the final performance, the chimpanzee changed the three digit number beginning with 0 (000) to 7 (111) in successive steps in the manner of counting. Now, as we increased the length of the sequence the results were comparable at each stage and completely orderly. Once we discovered the appropriate method, the entire chain was developed smoothly and with a low error rate,in approximately 78 days compared with some 4 months of unsuccessful attempts to fade from the numerosity control to intra-behavioral chain.

This general method of producing an intra-behavioral chain, in which the animal's own performance provides the discriminative stimuli for the next behavior, appears to us to be quite general and capable of extension to many situations. We judge fading to be a special case of a class of procedures for strongly determining a response either by

the physical environment or the discriminative stimuli already in the animal's repertoire. This is accomplished by using well-established cues to permit the gradual approximation of the terminal performance, while limiting the opportunity for error.

D. The Effect of Schedule Contingencies on the Accuracy of Behavior

Introduction

One of the advantages of the present experiment is that we can measure the frequency of the behavior as well as how closely the behavior conforms to its controlling stimuli (accuracy). In other words, we can simultaneously measure how disposed the animal is to engage in the behavior (overall rate of responding) and how accurately he performs (error rate), while we manipulate the reinforcement variables. Some of these, such as a fixed-ratio or a preceding fixed-interval schedule of reinforcement (Ferster, 1960) have been shown to have critical effects on both the overall rate of responding and the error rate in matching-to-sample experiments. In the initial development of the chimpanzee repertoires we early discovered that continuous reinforcement, food reinforcement for each correct response, was not a sufficient condition for generating these complex behaviors. In every case we found that we needed special reinforcement procedures, such as requiring a certain number of coorect responses, or delivering reinforcement only if a certain number of responses were emitted successively without error. While we used procedures of this kind and were convinced of their necessity, we needed to assess their effects both quantitatively

and systematically. Our major experiment, following the final development of the repertoire described above, was therefore to systematically evaluate the reinforcement variables which were important for developing complex and highly accurate stimulus control.

Experiment 1

Procedure. In our first experiment we measured the effects of schedules of reinforcement on discriminations (problems) which were relatively easy for the chimpanzees. The general procedure for the experiment was as follows.

The baseline performance consisted of the matching-to-sample procedure in which the samples were a series of two numbers of geometric forms and the chimpanzee's task was to choose which of two binary stimuli corresponded with the numerosity of the sample. In this experiment the sample consisted of either 7 or 1 geometric figures and the correct binary stimulus below was either 111 (7) or 001 (1). Correct responses were reinforced on a fixed-ratio schedule. Each time the animal matched a sample correctly the reinforcement was a tone. A food pellet was delivered following every 15th tone (FR-15). At this stage of the experiment the error level was of the order of 5 percent. Most of the errors occurred at the start of the fixed ratio. When the performance was stable at this final error level with the fixed-ratio schedule of reinforcement we changed the procedure so the specified number of correct responses had to be emitted successively without error. If an error occurred the animal was required to start the fixed ratio all

over again. In other words, each incorrect response reset the ratio counter. We increased the number of successively correct responses required of the animal in stages, until food reinforcement occurred only after 50 successive correct responses (FR-50 with reset). We had to approach the final schedule very gradually, increasing the number of responses required as the number of errors decreased, to maintain an adequate reinforcement density. After we achieved a stable performance on FR-50 with the reset contingency, we then returned to a fixed ratio of 15 without the reset to see if the original error level could be recovered.

Results. The reset contingencies progressively decreased the error level until it reached values between 1 error per 100 responses and 1 error per 1,000 responses, but the performances were not reversible. Once an accurate performance developed, it remained accurate even under the original schedule values. When we returned to requiring simply 15 correct responses the performance remained nearly as accurate as it was with the reset contingency. We then tested the hypothesis that the amount of differential reinforcement needed to produce a low error level depended on how difficult the discriminations were. Full quantitative details will be given when this experiment is published.

Experiment 2: Complex Discriminations

Procedure. The experiment was then performed with other pairs of stimuli, more easily confused with each other, such as 7 and 2, 7 and 3, 7 and 4, 7 and 5, and 7 and 6. The procedures were essentially

the same as those in Experiment 1.

Results. The main result of this experiment was that it took more reinforcements (trials) to develop the performance as the stimuli became more similar and the final error level, without the special reset contingency, was also higher. When 50 successive correct responses were required for each pellet of food, however, the error level for all of the stimuli reached the same low levels as with the simple discrimination of Experiment 1. Figure 4 gives some of the detail of how reset contingency affected the animal's accuracy, when the stimuli were 6 or 7 geometric figures.

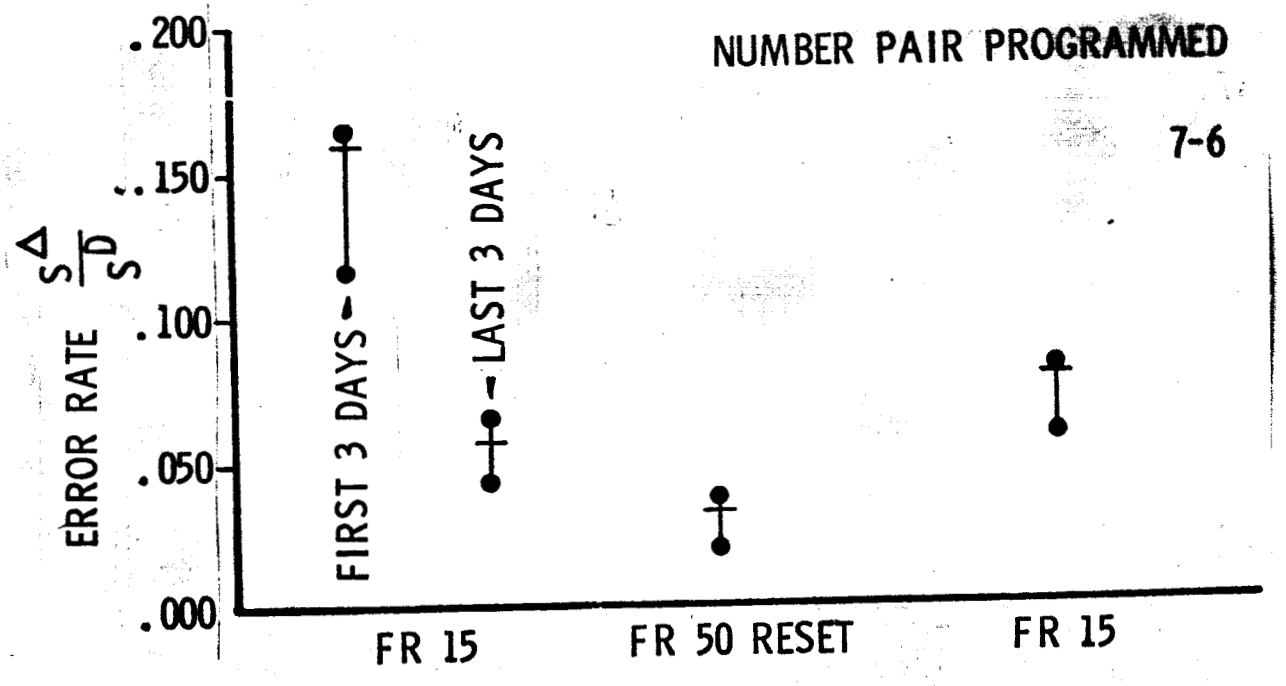


FIGURE 4

This figure shows the error rate for problems involving seven or six geometric figures. Three days' data are given for each point on the abscissa.

Because the stimuli are so similar, induction between them gives a high initial error level, 12 to 16 errors per 100 correct responses. After 33 sessions of further reinforcement on this same schedule, the error level falls to approximately 5 errors per 100. When 50 consecutively correct responses were required the error level falls even further to 2 or 3 errors per 100 correct responses. The return to the original reinforcement schedule increases the error level to between 6 and 7 responses per 100, approximately the same value as at the start of this experiment. The experimental control over the error level throughout the experiment was as precise as in this figure. Throughout all of the experiments, the distinctions between error levels of 2 errors per 100 and 5 errors per 100 are completely reliable from hour to hour and session to session. Although the error level on FR-15 was low to begin with (5 percent) the reset contingency reduced it to half the original value (approximately 2.5 percent) and the return to FR-15 without the reset increased the error level to slightly more than the original value. Full quantitative details will be given when this experiment is published.

Experiment 3: All Pairs of Numbers

Procedure. In the final experiments the performance we developed involved the presentation of any of 7 numbers of geometric forms on a given trial. We varied the schedules of reinforcement as before, to evaluate the fixed ratio and reset contingency. The experiments entailed very large and stable samples of the animal's behavior since the daily session consisted of some 1 - 3,000 trials. The performances were maintained at each value of the variables until a stable error level was achieved.

Results. The difficult discriminations of this experiment produced much higher error levels than in the previous experiments. Continued reinforcement on simply the fixed-ratio schedule reduced the error level to 1 error in 10 correct responses. The results of these procedures for one animal are summarized briefly in Figure 5 and show that the performance improved simply as a result of continued reinforcement until 50 consecutively correct responses were required. Then the reset contingency was responsible for a gain in accuracy of about 2.5 percent. The initial performance, after extended reinforcement after each 15th correct response, still produced an error level of 30 to 36 errors for each 100 correct responses. This value fell to 13 after more exposure at this point. Requiring 25 consecutively correct responses has little additional effect and the error level continues to decrease even with the return to the original FR-15 schedule of reinforcement. Returning to

a requirement of 25 consecutively correct responses decreases the error level even further and the shift to 50 consecutively correct responses decreases it even still further reaching approximately 3 errors per 100.

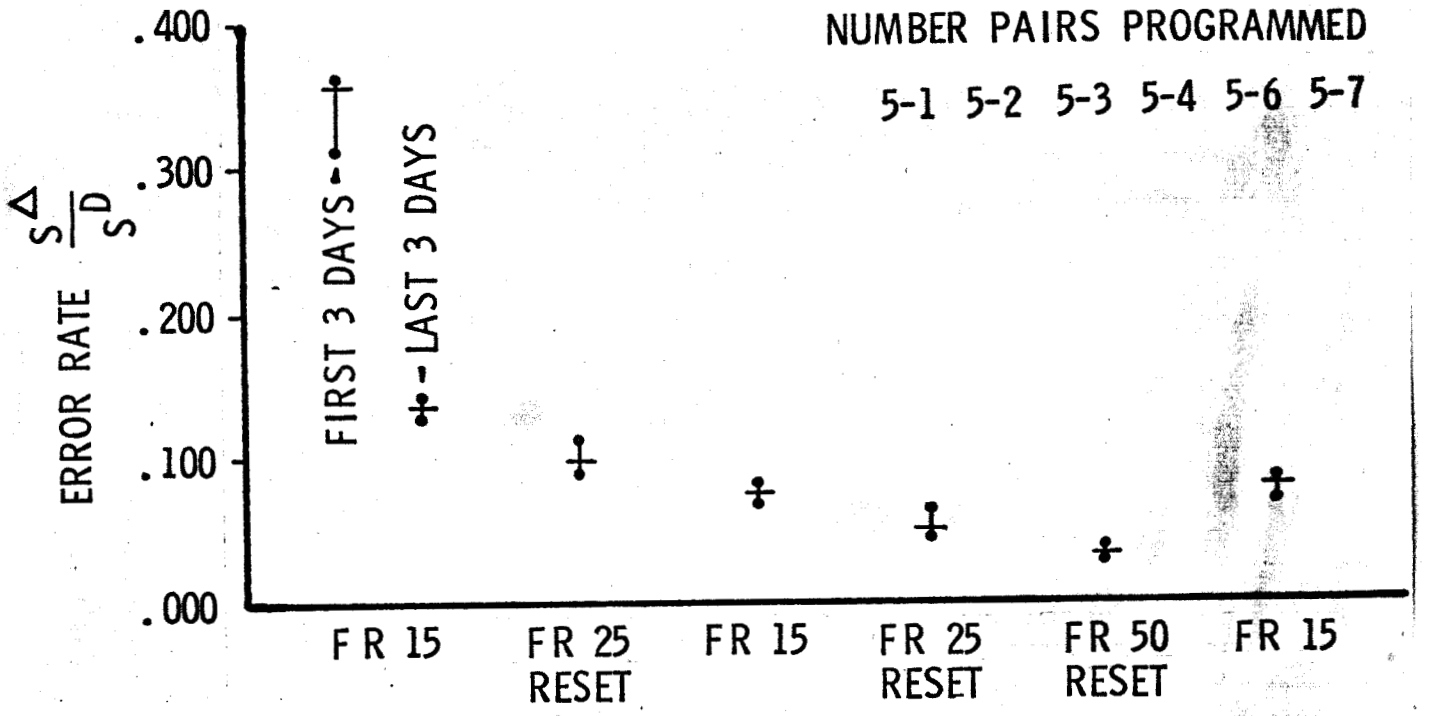


FIGURE 5

This figure gives error rates over three days for each of several contingencies, shown on the abscissa. The first two points show the fall in error rate associated with continuing exposure to FR-15.

The final procedure, the return to simply FR-15 shows that the requirement of 50 consecutively correct responses was in fact responsible for the change from 5 to 2.5 errors per 100 correct responses.

These results illustrate the difficulty of experiments with a changing baseline. We cannot always be certain in this experiment that the continued exposure to the reinforcement procedures produce the

decline in the error rate regardless of the schedule of reinforcement. Only the 50-response requirement clearly ocntrolled an accuracy level that was reversible. We will show evidence later, however, that even though the chimp would acquire the discriminations without the resetting contingencies, the acquisition is more accurate with them. Thus, it would seem that the main effects of these special conditions of reinforcement are to get the animals to attend to the stimuli. At high error levels the contingencies produce stable discriminations with moderate error levels that can be maintained under most conditions. Once the error levels reached 5 errors per 100, it goes to lower values only by requiring large numbers of correct responses to be emitted consecutively. Unlike the intermediate performances, the error levels immediately return to higher, although still moderate, levels when the schedule is changed to a moderate one. Full quantitative details will be given when this experiment is published.

Experiment 4: A Replication of Experiment 3

Procedure and Results. We replicated the effect of requiring 50 consecutively correct responses in still another experiment using 5, 6, and 7 geometric figures in the sample window which were the occasions on which the animal chose the appropriate binary stimulus below. An accuracy level of approximately 4 errors per 100 correct responses was developed quickly and the experiment shows how requiring 50 consecutively correct responses reduces the error level so long as the contingency is maintained. The results are shown in Figure 6. The first section of the

graph showing the mean values when 15 correct responses were required shows values of .03, .036, .045, and .046.

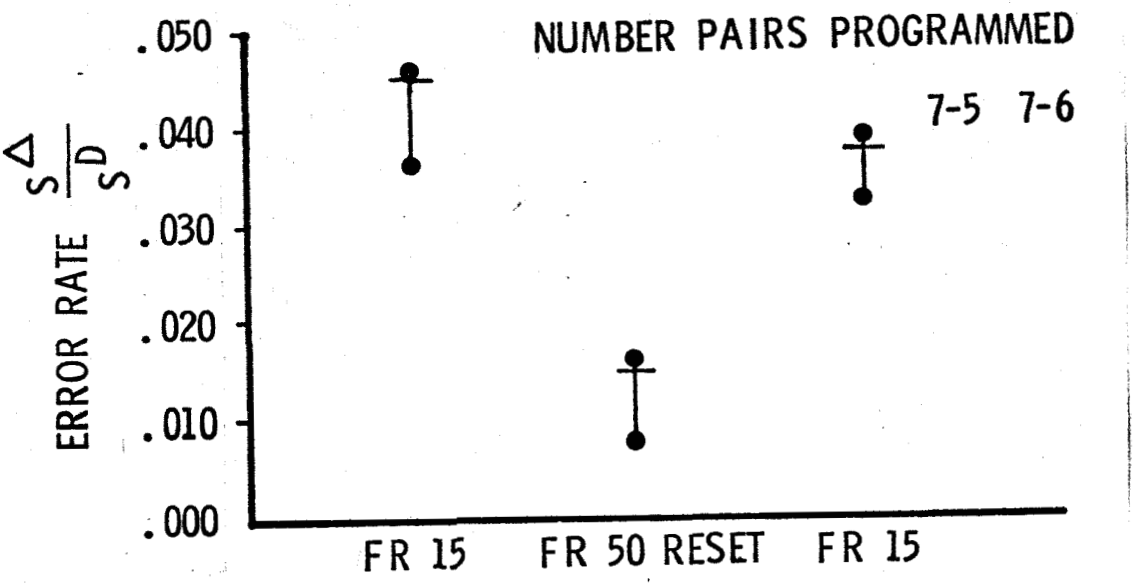


FIGURE 6

This figure presents error rates over three-day periods for each of the contingencies described on the abscissa.

The final performances after 5 sessions in which 50 consecutively correct responses were required were approximately 1 error per 100 correct responses, and the return to the original condition reinstates approximately the corresponding error levels. Full quantitative details will be given when this experiment is published.

Discussion. This experiment confirms all of our previous results, showing that extinction is the main mechanism for producing close control of the animal's behavior by these complex stimuli. The common factor among the procedures which produced stimulus control (accuracy) is that the amount of unreinforced behavior increases markedly when the animal's behavior does not conform closely to the stimuli. In contrast, punishment of the responses by time out has proven ineffective, except where extinction procedures have already brought the behavior closely under the control of the relevant stimuli. In this case they function,,in a manner of speaking, to make the animal more careful.

An interesting parallel exists between these data and the early work on the reinforcement of a simple response, e.g., bar-pressing or key-pecking, on fixed-ratio schedules. Reinforcement for each response produces relatively low and sometimes erratic rates of responding. However, if several responses are required for each reinforcement we

observe a transition to a relatively high and even rate of responding. If we think of this form of responding as analogous to the control by the stimuli in the matching-to-sample procedures, the common factor in the two procedures is the increased rate of the reinforced response as a result of its fixed-ratio reinforcement. If in the matching-to-sample procedure we had recorded only correct responses on a cumulative recorder, we might have obtained records which were similar to those obtained with simple responses maintained by continuous reinforcement or on fixed-ratio schedules.

The reset contingency might be considered as a special case of fixed-ratio reinforcement. The reset procedure guarantees a high rate of correct responding immediately preceding the food reinforcement. Consequently the reinforcement occurs selectively after "runs" of correct responses and never occurs soon after an incorrect response.

An additional parallel between matching to sample and a simpler response also exists with the conditioned reinforcer following each correct response. Most investigators have found that it is necessary to provide some feedback (frequently a relay click) to the animal from the operation of the manipulandum. This is done to maintain the topology of the response. The tone following each correct response has an analogous effect because it occurs only after responses whose topology is correct. A response that produces the tone can also operate the feeder.

Boren (1953) was unable to recover low rates of responding associated with low fixed-ratio values just as in our experiments. These data are quite similar to the results we obtained during fixed-ratio reinforcement of simple problem solutions (simple discriminations).

E. Maintaining Complex Behavior under Deferred Reinforcement

Procedure. In another experiment, while we were synthesizing the complex repertoire, we measured the chimpanzee's behavior as it was maintained by conditioned reinforcers leading to food only after a long sequence of various activities. The procedure was as follows. When the animal entered the first chamber each correct matching-to-sample response was reinforced by a tone. Every 50 correct responses produced the sound of the feeder mechanism but no pellets were delivered. After the feeder sounded a certain number of times (equivalent to the number of pellets, had we been delivering food), the door opened and the animal could leave the chamber into the second experiment where the same procedure was repeated with the "writing" performance. Only after an amount of behavior in the second chamber about the same as required in the first did the exit door to the second chamber open, and the animal was admitted into the third chamber where he received biscuits, fruit juice and food pellets in an amount equivalent to what he had been obtaining under more direct reinforcement. Thus, each chimpanzee engaged in some 2 - 400 trials before receiving any food.

This final performance was produced by a set of approximate procedures. In the first procedure the food pellets were delivered on a

percentage schedule (for example, every other time) so that every so often the feeder would make its customary sound but no food would be dropped. The percentage was continuously reduced until all the food was delivered at once in each of the two chambers. Then we no longer delivered food at the completion of the performance in the first chamber, but continued to deliver it in the second chamber. Finally, food was delivered only in the third chamber.

Results. Under these deferred reinforcement conditions, consisting essentially of a long chain of complex behavior in which food occurs only at the very end, not only did we sustain the animal's performance but it continued with a high degree of accuracy. In the final performance we maintained the chimp's performances under the deferred (or delayed) reinforcement conditions approximately two weeks. The major result was that the total amount of behavior emitted decreased somewhat (hence the food intake) although the animals still obtained enough food to maintain their weight. The error level increased slightly, but this was not the major effect of the experiment. When we returned to the original conditions (receiving food in each chamber) the food intake and amount of behavior returned to the baseline levels. The experiment is confirmation of the general principle we observed elsewhere, such as in the delayed reinforcement experiment (see below). Conditioned reinforcers conform the behavior to the relevant stimuli as their major effect in these repertoires. The conditioned reinforcers may at one time have the effect of increasing the accuracy of the performance by its differential

reinforcement of the animal's behavior and at the same time decrease the frequency of occurrence of the response. Full quantitative details will be given when this experiment is published.

The decreased food intake when we deferred reinforcement to the end of a long chain is not consistent with many of the earlier experiments which recorded increased levels of performance as the body weight of the animal falls (Ferster & Skinner, 1957; Clark, 1958). Our results confirm those of Findley (1962), who discovered that he could vary the animal's food intake in the total environment by increasing or decreasing the amount of behavior required for the delivery of each unit of food. The difference between Findley's and our results and those of earlier experiments come about from the total environment and the natural method of maintaining the animal's behavior on a 24-hour basis in which the amount of food received is limited only by the amount of behavior it can support.

The effects of the conditioned reinforcement in this study parallel the results of experiments in human verbal behavior in which the variable "knowledge of results" was investigated. In these studies, "knowledge of results" might be considered to be a conditioned reinforcer, both in terms of the immediate experimental history of the subject and in terms of his history prior to his participation in the experiment.

F. Future Plans

The next stage of our plans is a shift toward a social experiment in which one animal will present complex stimuli which will support behavior in the second animal in the manner of a true verbal interchange. We have, therefore, redesigned the living and experimental chambers including the spacial arrangmment as well as the design of the doors and automatic control systems. A new design will allow us to identify, control and

guarantee automatically (behaviorally) the location of any animal in any portion of the experimental environment. This will give us the capability of using access of one animal to another as a reinforcer in a social experiment. Our experience in building and maintaining these devices has produced a series of new designs which we think will solve many of the deficiencies of the early chambers, and make it easier to produce copies of the present one. The present chamber has been in operation for two years during which the door mechanisms have been progressively deteriorating while the animals have progressively acquired new skills for bypassing their protective and controlling functions. We have completed detailed plans for the new chamber which will be built according to a modular design from "Unistrut" components. The new cage will be built with most of the components under tortion from a rigid frame rather than as in the chamber we now have in operation in which the panels are under compression making it difficult to modify, remove, and rebuild. The new chamber should make it possible to design environments much as with an erector set.

Our first experiment in the chimp program will be designed to produce a generalized reinforcer functionally analogous to money, attention, or power in the natural human ecology, a major theoretical problem and an important technical advantage in producing the repertoires we require. We take some encouragement in this goal from our experiments reported below under Deferred Reinforcement which suggest that the technical problems of constructing a generalized reinforcer and measuring its properties is feasible within the limits of our techniques and equipment.

II. FIXED-RATIO REINFORCEMENT OF LARGE UNITS OF BEHAVIOR

Introduction

Our experience with fixed-ratio reinforcement of complex repertoires in the chimpanzee experiment raised the question of the generality of these effects. The analysis offered above suggests that this principle might be extended to a large number of situations, providing a more highly concise control of behavior than is currently possible. Also if our analysis is correct it is not unlikely that many highly accurate complex performances seen in the normal ecology are maintained in a similar way. The experiments described below are designed as an initial attempt to investigate this possibility.

It has been well established that a broad range of subjects may be taught to temporally space their responses in a relatively accurate manner. One schedule of reinforcement used for this purpose is the differential reinforcement of low rates (DRL). For example, on a DRL-20 sec. schedule of reinforcement all responses which follow the previous response by 20 sec. or more are reinforced. In other words, if the subject spaces a response (specified by the experimenter) so that each emission follows by 20 sec. or more the previous emission, the response is reinforced. Responses having a latency from the last response of less than 20 sec. are not reinforced, and began a new 20 sec. "waiting period". In most organisms the maximum probability of the experimentally specified response, for example, bar-pressing, occurs about 20 sec. after the last response. That is, a distribution

of interresponse times (IRT's) shows its highest relative frequency on or about the reinforced interresponse time. If the unit of behavior on this schedule is considered to be pausing for 20 sec. (doing something else) followed by the emission of the experimentally specified response, then the schedule would be analogous to the simple reinforcement described above in Section D of the chimpanzee experiment. If our analysis of the chimpanzee data is correct then the requirement of several of these DRL responses for each reinforcement, fixed-ratio reinforcement of DRL performance, should increase the temporal control exercised by the schedule. Reset requirements should provide additional increments in accuracy.

We believe pigeons to be particularly good subjects for the initial stages of this work. Their behavior is markedly poor at even moderate values of the schedule, for example, DRL-30 sec., and it is not unusual for them to receive fewer than 5 reinforcements per hour after several hundred hours exposure to the schedule. This kind of performance will give us a baseline which is both well known and which permits large increments in the accuracy of the temporal spacing of the response.

Experiment 1

Procedure. The subjects, pigeons, will be exposed to a DRL-30 sec. schedule of reinforcement until the performance is stable. We will then systematically manipulate the number of "correct" responses necessary to produce each food reinforcement. Our primary interest will be how

closely the behavior conforms to the temporal requirements of the schedule of reinforcement. Overall rates of responding, rates of reinforcement, as well as the IRT distributions will be obtained.

Experiment 2

Procedure. Procedures similar to those described above will be followed with the addition of the reset contingency.

Additional Experiments

Additional experiments will be performed, if they seem warranted, to assess the generality of this phenomena over a relatively wide range of complex performance requirements and several additional species.

III. DELAY IN REINFORCEMENT (DEFERRED REINFORCEMENT)

A. Introduction

The experiments in this section were performed by Charles B. Ferster, Ph.D., and Clifford E. Hammer, Jr., B.A. John J. Randolph, Ph.D., who recently joined our staff, has also been participating in these investigations since his arrival at the Institute for Behavioral Research.

In our early experiments on deferred reinforcement we developed conditions for maintaining behavior reinforced with long delays (24 hours). We measured the behavior as a function of some of the variables judged to be critical. The results have been published (Ferster & Hammer, 1965). In these experiments, we discovered conditions under which we could maintain small fixed-ratio performances in baboons when food was delivered some 18 to 24 hours following the performance. Our analyses of delayed or deferred reinforcement performances emphasize that we are developing chains of behavior, and that deferring reinforcement by a blank interval is a special case of long chains of behavior. While we do not specify a chain of responses in many delayed reinforcement paradigms such chains are likely to occur by accidental reinforcement (Kelleher & Gollub, 1962; Skinner, 1948; Ferster, 1952). Both the classical delayed-reinforcement procedures and a chain specify that a given response is maintained by a stimulus which, in turn, controls some other kind of behavior. In the chain some response is reinforced in the presence of the stimulus and in the delayed reinforcement case the

stimulus is simply present for a period of time followed by reinforcement. The two cases differ in how closely the animal's behavior is specified during the interval. In the chain, the stimulus is an occasion on which a specified response is reinforced by the next stimulus. In the delayed reinforcement case, the stimulus is an occasion on which the next stimulus occurs after a fixed period of time. Even though no behavior is required during the delay stimulus, responses may be reinforced accidentally if they happen to occur just before reinforcement. Another possibility during the delay stimulus is that the animal may engage in other behaviors maintained by other reinforcers.

Our experiments this past year have stressed measurements of the role of other performances which can be emitted during the delay in reinforcement. The procedures have in common that the animal acquires some concurrent performance maintained independently of the response whose reinforcement is deferred. Thus, for example, the baboon enters a chamber and locks the door behind him. When he presses a key 75 times, a stimulus comes on for 6 hours. Then a stimulus on a second key comes on; the second stimulus is an occasion when pressing the food key produces food pellets to a preset total. Still another key, however, reinforced on a fixed-ratio schedule, opens the door to the main chamber where he can socialize with other animals, exercise, and return to the experiment at will. Thus, the delay stimulus may control two behaviors. One is "waiting" for the stimulus in whose presence

food is delivered. The second is operating a key whose reinforcement is exit to the rest of the cage where the animal can exercise and socialize with the other animals. The exit stimulus is functionally a generalized reinforcer, since its reinforcing effect derives from a wide range of behaviors supported outside of the experimental chamber. These might be a chance to exercise, grooming with other animals, sexual activity, a chance to act aggressively, or play with the physical aspects of the larger external environment. Our results here serve as a pilot data for the proposed chimp generalized reinforcement experiments. We noted that the exit behavior is a function of the dominance hierarchy and other relations among the animals. Still we measured the overall of effect of "exit" as a reinforcer. The redesigned chimp environment will be a better situation to analyze the component variables of a generalized reinforcer.

B. Current Experiments

Experiment 1

Procedure. In our first experiment, we introduced the baboon to the final procedure: In the presence of a red light, 75 responses produced the delay stimulus, a flashing green light. At the end of the delay period the green light went off and in the presence of the food-key light, each press of the food key delivered a pellet, to a preset total. Five responses on an exit key opened the exit door at any time. The baboon could return to the work chamber at any time. The elapsed

time counted toward reinforcement whether or not he was in the chamber. Colloquially speaking, he could "pass the time" during the delay interval by joining the other animals.

Results. The major result was a complete loss of the performances reinforced with food that we had generated with delayed reinforcement during our previous experiments. The baboon continued to go in and out of the chamber frequently even though the food behavior broke down to the point where the animal would have starved to death, had we not intervened. The animal simply went in and out of the experimental chamber without ever operating the key whose reinforcement by food was deferred. Full quantitative details will be given of this experiment when it is published.

Experiment 2: Measuring the Effect of Exit During the Various Parts of the Chain

Procedure. We returned to the conditions of our first experiment without exit from the chamber to reinstate the behavior under the deferred reinforcement. When we recovered a stable performance under delayed reinforcement, we began a set of approximations to the complex procedure we had attempted earlier. In this series of experiments we first limited the exit behavior to a 50-minute time out after food reinforcement. This was done by a light over the exit key which controlled it in the manner of a discriminative stimulus. When the light was on, pressing the exit key opened the door; when it was off, pressing the exit key had no effect. We then permitted the animal to exit during

the delay stimulus. In the final performance the exit option was made concurrent with the food option and a number of responses required to exit was added as a parameter to this experiment.

Results. With exit possible only during time out after reinforcement the food performance under deferred reinforcement continued normally and the baboon left the experimental chamber during most of the time outs. When we allowed the baboon to exit during the delay stimulus, the result was an increase in the amount and frequency of the food-reinforced behavior. Now the stimulus controlling the delay interval was a more effective reinforcer because it allowed the second repertoire, playing outside. When we allowed the baboon to leave at the start of the sequence, when key presses produced the delay stimulus, the food behavior was weakened - the baboon pressed the food key less frequently.

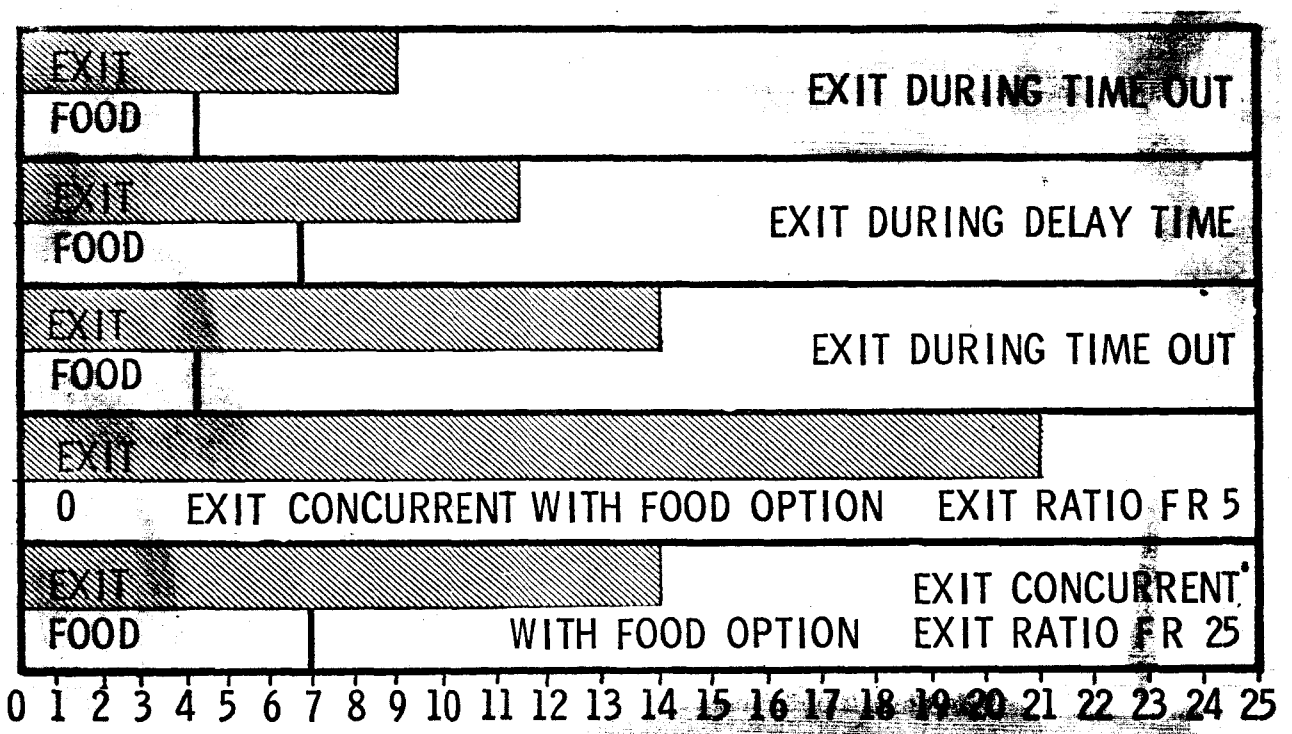


FIGURE 7

Figure 7 shows the number of food chains and the number of exit chains completed, as a function of the location of the time-out option. It also shows the effect of an increase in the number of responses required to exit, when the exit option was concurrent with the food option.

The exit behavior weakened the food performance especially when only a small number of responses was required on the exit key. With a small exit requirement, the animal simply went in and out of the chamber, only occasionally pressing the food key. When we increased the number of responses required to exit from the chamber sufficiently to weaken this behavior the baboon stayed in the chamber long enough to press the food key enough to produce a delay stimulus, and hence food. Full quantitative details will be given when this experiment is published.

Discussion. We interpreted this experiment to mean that the exit behavior strengthened the food-key behavior only if it occurred in the exact temporal relation to the food-maintained behavior (that is, if it followed it). When the exit behavior could occur concurrently with the food behavior, it was so pre-potent over it that the food behavior hardly occurred at all. We saw the unusual condition that when the animal could leave the chamber at any time by pressing the exit key five times or less, he seldom stayed in the experimental chamber beyond a few minutes and was unable to produce even the smallest amount of food. This phase of the experiment had to be interrupted as the animal began to starve to death, running in and out of the experimental chamber repeatedly. When we increased the amount of behavior required before the animal could exit, we once more weakened the exit performance sufficiently so that it was not pre-potent over the food behavior and the animal began feeding himself once more.

This aspect of the experiment is relevant to the problems in self-control, where there is a performance (the exit performance), which is

immediately and strongly reinforced and another performance (the food performance) which is weak because its reinforcement is deferred; but the weakly reinforced behavior is required for the animal's survival. Many analogs could be constructed here of performances in the normal ecology important to the individual which do not occur in competition with less biologically important but more immediately occurring reinforcers.

Experiment 3

Procedure. In this experiment we made systematic and parametric measurement of the animal's performances as a function of the number of responses we required on the exit key to leave the chamber. The duration of the delay by which food reinforcement was deferred was a parameter of these measurements. Thus, we fixed the delay in reinforcement at 30 minutes and systematically varied the number of responses required to exit. We then repeated the entire experiment at a 6-hour delay in reinforcement.

Results. We discovered values of the exit FR which were sufficiently high that the animal continued to engage in the food performance, despite the possibility of exit. Further increases in the number of responses required to exit had no effect on the food behavior once the exit behavior was weakened enough that it was no longer pre-potent over the food behavior under delayed reinforcement. Even at large values of the exit fixed ratio, the baboon continued to exit from the experimental chamber. The baboon exited from the experimental chamber three or four times per day even when 1,200 responses were required for each exit. This experiment

gives us a direct measurement of exit as a generalized reinforcer, controlling activities in the larger open part of the cage with the other animals.

In summary, the exit to the outside chamber is a powerful reinforcer, and its concurrent availability with the food-reinforced behavior has little effect on the food behavior except when it is completely pre-potent over it, as for example, when only a few responses are required to exit.

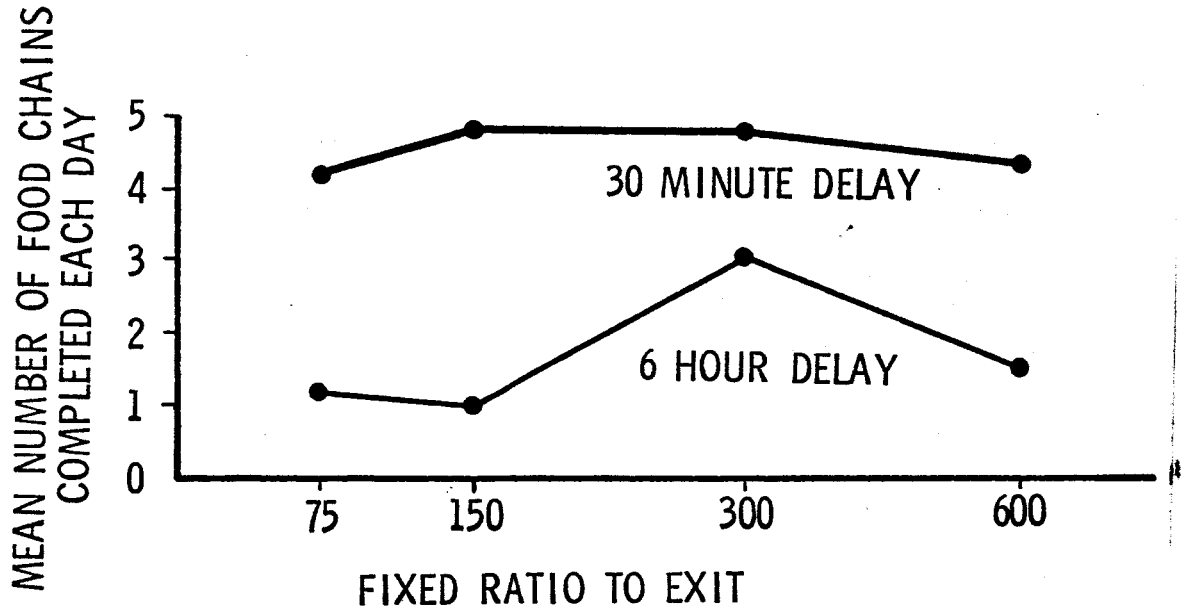


FIGURE 8

Figure 8 shows the number of food chains completed as a function of the number of responses required to exit, and the length of the food chain delay.

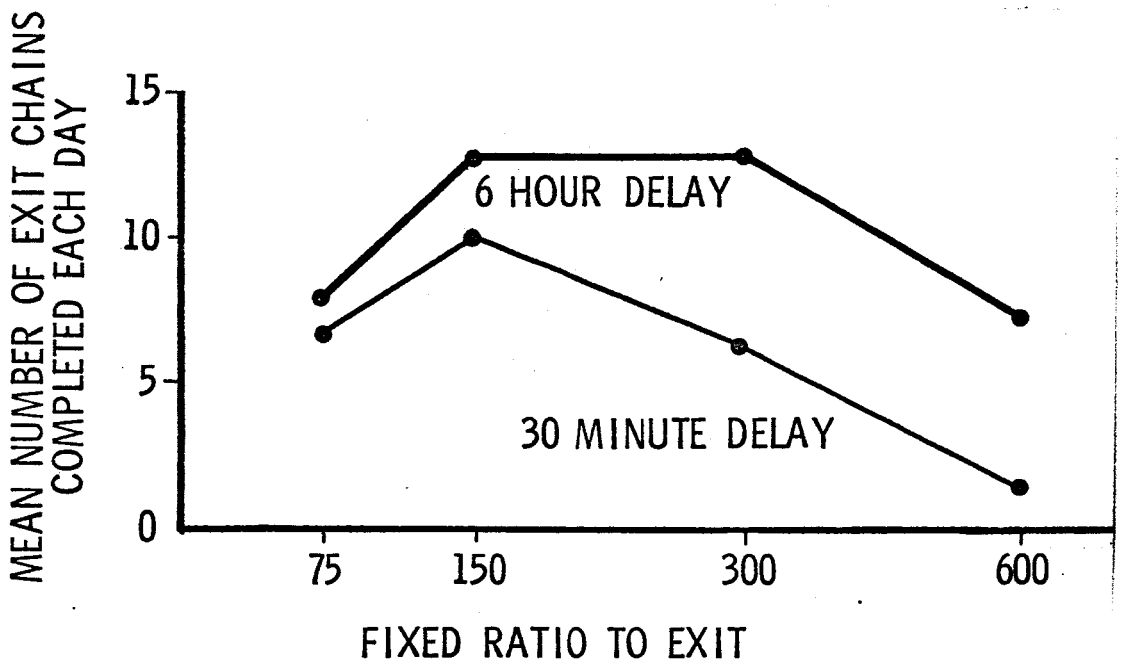


FIGURE 9

Figure 9 shows the number of exit chains completed as a function of the number of responses required to exit, and the length of the food chain delay.

Following the parametric study of the number of responses required to exit and the interval by which the food reinforcement was deferred, we reduced the exit requirement to five responses and maintained the delay at six hours. At these values of the parameter the animal stopped responding on the food chain. We returned to the previous conditions, 75 responses on the exit key and a 6-hour delay, but were unable to recover the baseline. We then removed the exit option and the number of completed food chains increased slightly. However, the amount of food the animal received each day did not constitute a subsistence diet.

Full quantitative details will be given when this experiment is published.

Experiment 4

Procedure. In our next experiment we set out to discover conditions for maintaining the food-reinforced behavior (under deferred reinforcement) despite the concurrent strength of the exit behavior. Our general plan was to strengthen the food behavior by reducing the length of the delay in reinforcement to five minutes, a value at which strength of the food behavior would be sufficient to compete with the exit behavior. We first varied the number of responses required for food, to recover this baseline at five minutes delay. We then progressively increased the length of the delay, keeping the exit behavior strongly maintained under virtually continuous reinforcement.

Results. With the careful adjustment described above we were able to increase the delay to 30 minutes and at the same time maintain the food behavior strongly enough, despite the delay, so the animal continued to work for food. At these same values we had been unable to maintain the animal's food performance prior to the special history. Thus, a gradual progression in the conditions of the delay-reinforced response proved to be a way of developing the equivalent of self-control in the baboon. Full quantitative details will be given when this experiment is published.

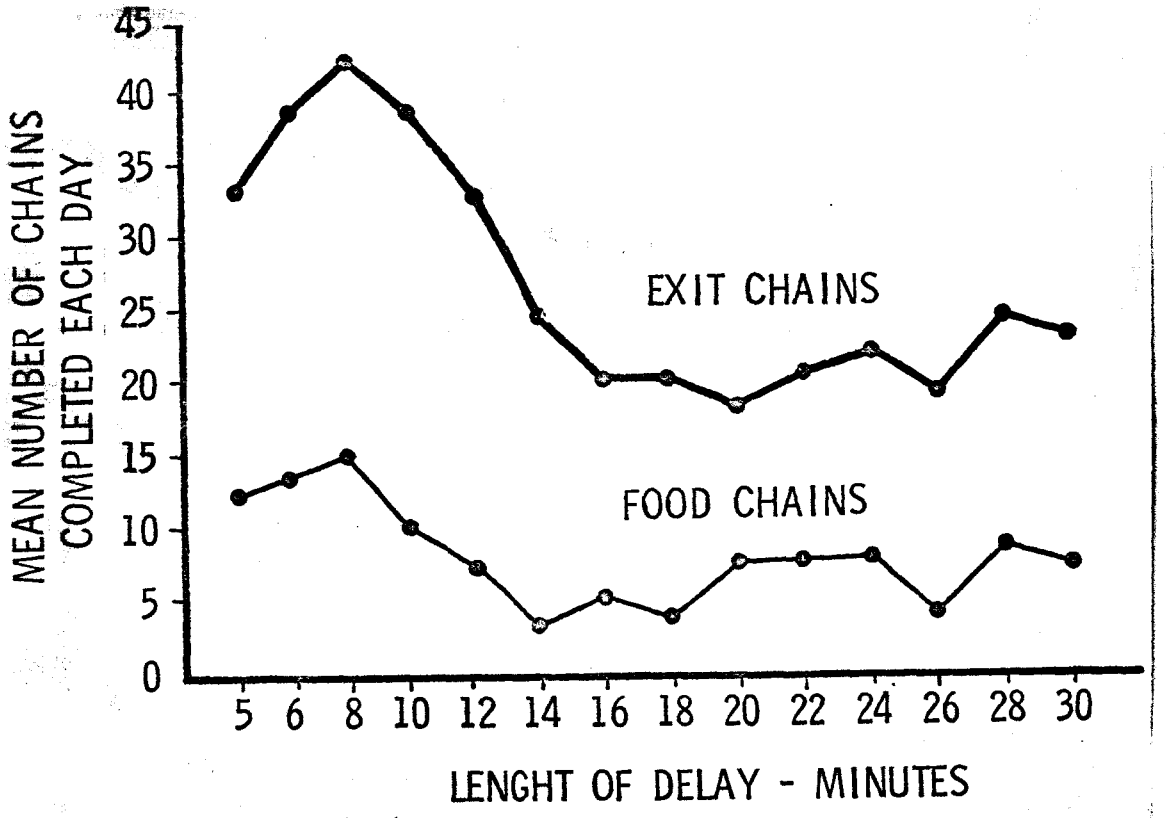


FIGURE 10

Figure 10 gives the number of food chains and the number of exit chains completed as a function of the length of the food chain delay.

The complex interaction among the baboons in the living area of the chamber is a complication in these experiments. The tendency to exit is clearly a function of the interpersonal relationships among the animals. This variable is an important factor in these experiments as well as in the experiments which we plan with the chimpanzees. Our continuation experiments, therefore, will focus on direct measurement of the social interactions among the animals and effects of these social interactions on the reinforcing properties of access to the living area. We are developing instrumentation toward this purpose.

C. Concurrent Chains of Behavior Each Maintained with Deferred Reinforcement

Introduction

In this experiment, we developed an experimental paradigm for independently maintaining three responses, each with its reinforcement deferred. The rationale of the experiment was to give the animal an opportunity to work on a second chain while he was waiting in the delay stimulus of the first, and in the third chain while he was waiting in the delay stimulus of the other chains.

Procedure and Results. This experiment has only been in progress six months, and we have just developed an effective experimental paradigm. At the present time we are maintaining behaviors on the three chains which we have demonstrated to be independent of each other. The experiment is being carried out in the same total environment as we described above. The intelligence panel contains three keys each above the other, with a projection unit to provide the stimuli which define the chains and the delay stimuli. Reinforcements occur on a neighboring panel controlled by appropriate lights and colors. Initially, the behaviors on each of the keys was controlled by interactions from other keys. The experiments during the first six months of this program were characterized by a series of procedures designed to reduce the inductive effects among the various chains. This was accomplished by adding stimuli correlated with differential contingencies, and by programming two chains at a time with schedules and amounts of reinforcement with large differences to produce maximum contrast by differential

reinforcement. Currently we are able to maintain behavior with three different magnitudes of delay (6 hour, 30 minute, and 5 minute) and three magnitudes of reward (100 pellets, 20 pellets, and 2 pellets). Each of these six conditions is associated with a different discriminative stimulus, all of which control the behavior of the subject.

Now that we have achieved a workable experimental paradigm for studying these concurrent repertoires our plan will be to measure the strength of the behavior on one chain as a function of the performances on the other chains which are available to the animal. Thus the issue here is self-control as in the experiments described above. Thereafter, we can use the experimental paradigm to make measurements of the effects of the critical variables determining the influence of deferred reinforcement. The three concurrent chains may allow us to set different values of the parameters in each chain, so the entire experiment can be carried out within a session and within an animal. With such a procedure we will be able to study the magnitude of reinforcement as a variable, without being limited by frequencies of reinforcement which will starve an animal. In some of the previous experiments, for example, the animals did not receive sufficient food to maintain body weight and we had to interrupt the experiment lest the animal starve to death. With three concurrent chains we should be able to make a systematic description of the amount of food reinforcement as it maintains behavior. Full quantitative details will be given when this experiment is published.

D. Exit and Social Contact as Reinforcers

Introduction

We decided to set up an experiment specifically designed to measure how the performance in the experimental chamber might be related to social behavior outside. Like the previous experiments, the animal entered an experimental chamber (3' x 3' x 3') whose door locked behind him while he engaged in the experimental procedures. He could return to the larger outside space with the other baboons by pressing the exit key. Our experiments above showed that exit to the social situation was a durable reinforcer and the unique environment we had developed provided a convenient method for its investigation. These experiments continued these investigations in more detail.

Experiment 1

Procedure. In the first experiment, the animal received 15 pellets as soon as he closed the door behind him. However, a fixed-ratio performance on the exit key was required in order to open the exit door. Food would then be delivered when he again entered the chamber. This is a functional analog of "eat now and pay later" reminiscent of the installment plan in the normal human ecology. We systematically varied the size of the fixed ratio required to exit, measuring how long the animal remained outside, how long it took to emit the fixed-ratio performance, and how long it took to eat his food. Following this series of measurements, we then kept the exit performance constant and systematically varied the amount of behavior we required on the food key.

For example, the animal would enter the chamber, 200 responses would lead to 15 pellets of food following which a single response would open the door letting him into the social area. The fixed-ratio values used were 100, 200, 400, and 800 responses both on the food and exit keys.

Results. At first we thought we had designed a circular chain in which each performance is reinforced by every other performance. Our results indicate, however, that the place in the chain where the fixed ratio was largest, controlled a pause or a weakened behavior immediately preceding it. For example, when 400 responses were required to produce the food and only 5 responses to exit to the social area, the longest time was spent outside the experimental chamber just before entering, and inside the chamber before completing the fixed ratio. After the food was delivered, the subject ate it quickly and left immediately. When the fixed ratio on the food key was only 5, and 400 responses were required on the exit key, the longest latencies occurred in the experimental chamber following the delivery of food just before leaving for the social area. It took the animal longer to eat his food under these conditions, longer to operate the exit key following the food, but he stayed outside in the living area a shorter period of time. All of these measurements were systematically correlated with the magnitudes of the relevant parameters, giving smooth functions.

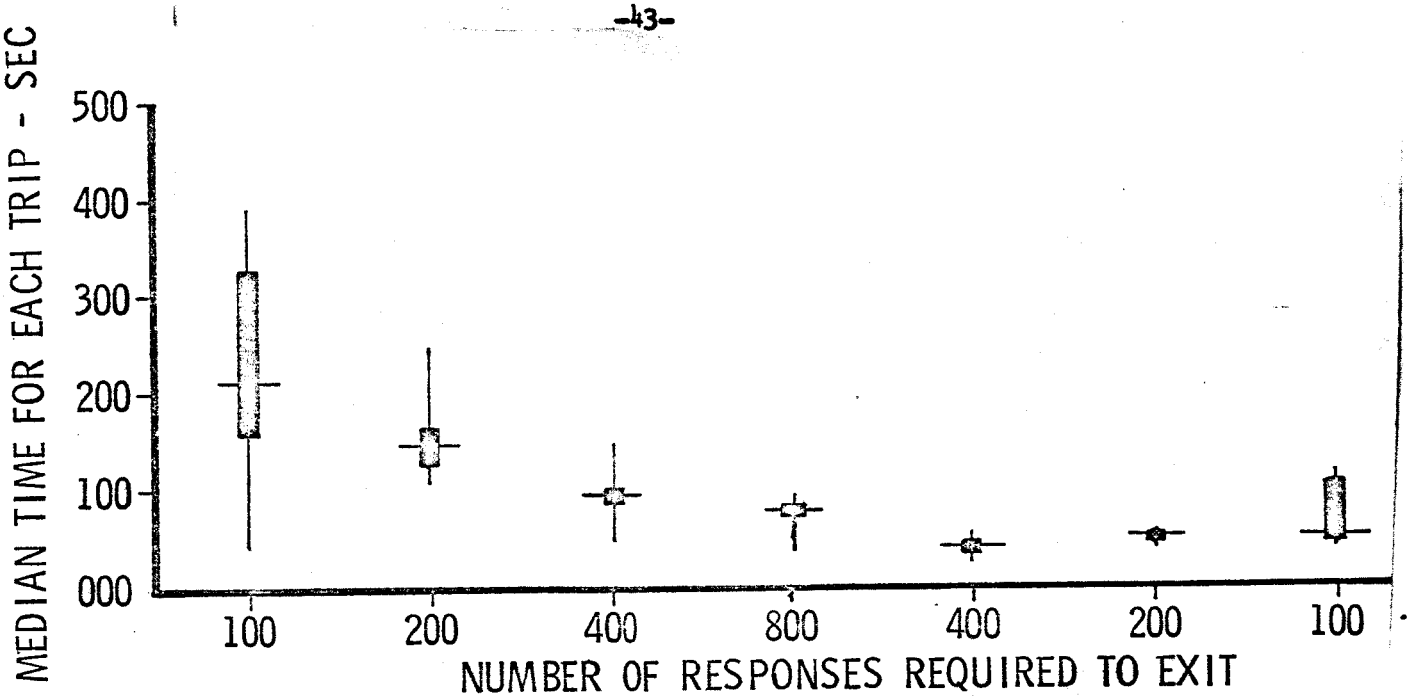


FIGURE 11

Figure 11 shows the range, interquartile range, and median time spent in the social area as a function of the number of responses required to exit.

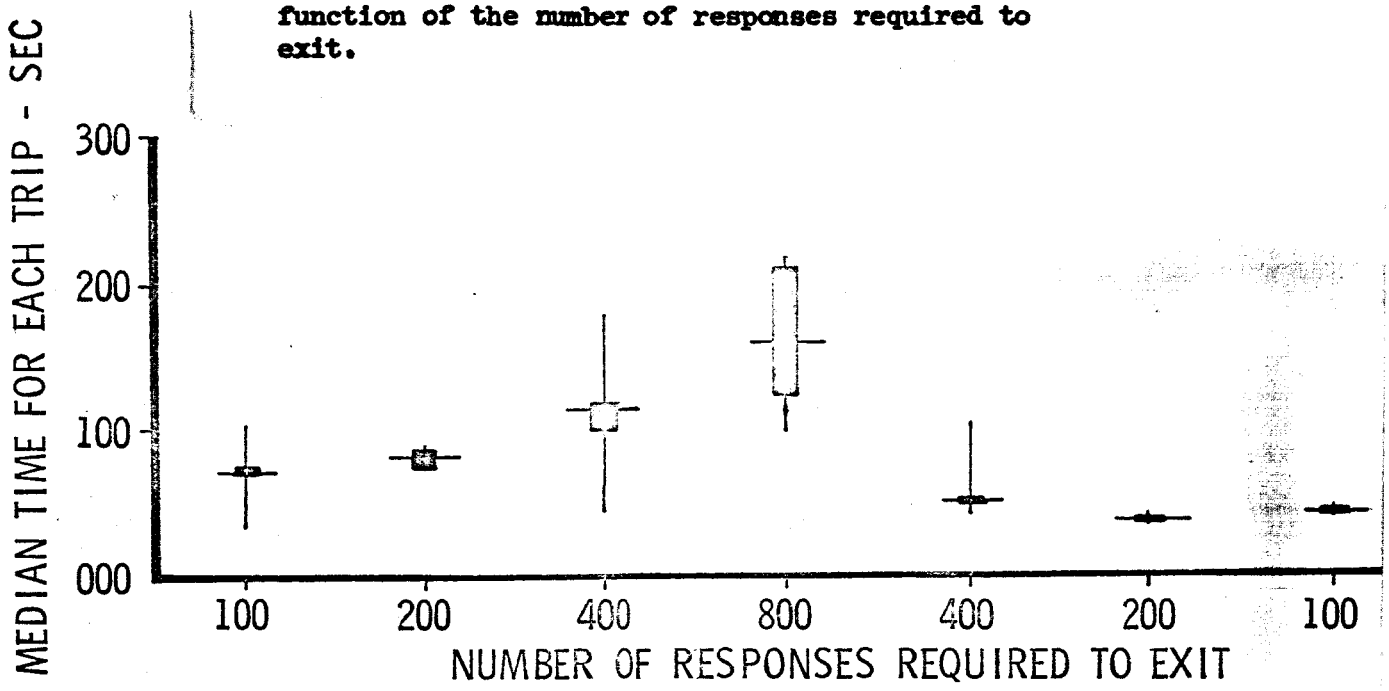


FIGURE 12

Figure 12 shows the range, interquartile range, and median time the subject spent eating as a function of the number of responses required to exit.

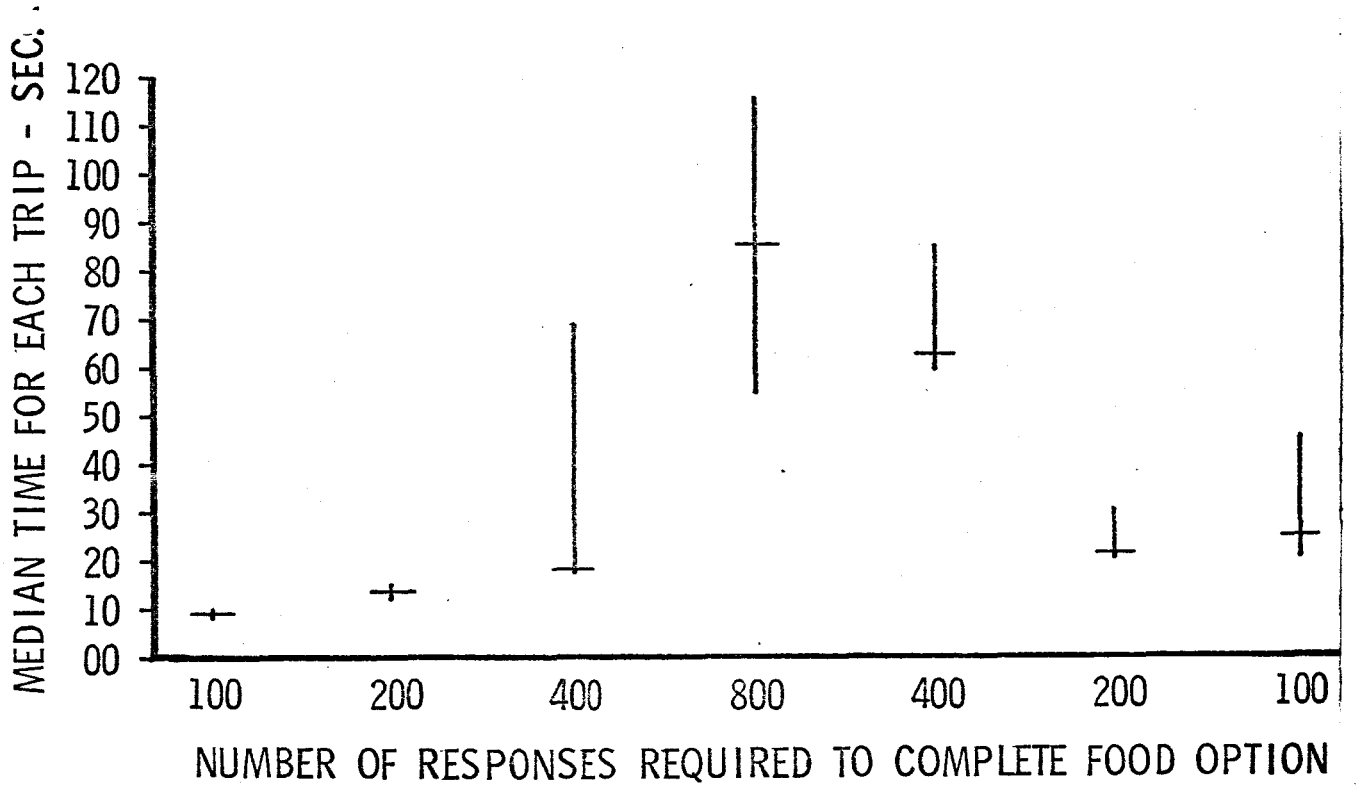


FIGURE 13

Figure 13 shows the range and median time that it took the animal to enter the cubicle as a function of the number of responses required to complete the food option.

Our inability to control the nature of the social interaction when the animal leaves the experimental chamber into the living area confounds our results slightly, as with the previous experiments. Our plans for direct measurement of the social interaction should solve this problem by documenting it and leading to methods of manipulating the relevant variables. Full quantitative details will be given when this experiment is published.

Experiment 2

Procedures. We made the opportunity to exit to the social area concurrent with the food-reinforced behavior. This experiment was designed to compare these behaviors when they are not serially dependent

upon each other. For example, the food can be available under very optimal schedules of reinforcement so that the animal is essentially on free feeding while we systematically vary the reinforcement schedules of the exit behavior and conversely.

Results. This experiment confirmed the findings in the delayed reinforcement experiment where the immediate reinforcement of the exit behavior interfered with and was pre-potent over the weaker food behavior. In this case, the conditions were reversed. The exit behavior reinforced on FR-400 was weak and the food behavior reinforced continuously was strong. The result was that the animal remained in the chamber operating the food key at a sufficiently high rate so that the floor of the cage was continuously littered with food pellets which the animal did not eat. At the same time the behavior on the exit key was so weak that he did not leave the experimental chamber once in ten days. This occurred in spite of the fact that the exit behavior could be maintained on schedules of reinforcement as high as 800 if the food schedules were sufficiently high. Thus the issue here is self-control as in the experiments described above. The total amount of behavior the animal emitted reinforced by food which he did not eat was sufficiently large to have produced an exit reinforcement had he pressed the exit key this number of times. The result is even more striking given the fact that this animal is "boss monkey" for whom exit to the living area is maximally reinforcing under all conditions, in contrast with other animals whose exit behavior interacts with the dominance conditions operative in the social area.

It is interesting to note here that when the discriminative stimulus indicating that the food key was operative was turned off, the subject almost immediately worked off his exit ratio. This latency was about 15 minutes, as opposed to failure to operate the key over the past ten days. We believe that this is particularly compelling evidence for the interdependence of these two performances. Full quantitative details will be given when this experiment is published.

Current Experiments and Future Plans

We are presently completing a parametric study of all combinations of several fixed-ratio requirements on both the food key and the exit key. The results to date are quite promising.

At the completion of this experiment we plan to do further experiments to systematically evaluate the interdependence of these two performances. These experiments would utilize the probe technique described above. The baselines used for the experiments would be generated by values of the ratio parameters on both the exit key and the food key which had been shown in the current experiment to produce strong preferences for one of the two alternatives. We would be specifically interested in immediate changes in the disposition of the animal to engage in a nonpreferred alternative, and the effects of the removal of the preferred alternative on our other behavioral measures.

In addition to these experiments we plan to study the effects of delayed reinforcement of the exit option. This work will extend our investigations of deferred reinforcement to a generalized reinforcer (see the introduction to this section). We should then be able to make direct comparisons between these data and the effects of delay in reinforcement on behavior maintained by a conventional reward (food).

We will also attempt to establish stimulus control for the various values of both the food fixed ratio and the exit fixed ratio. As in the experiment on concurrent chains, Section C, this will permit us to manipulate critical parameters of these variables in a single session. This is particularly important for conditions which result in large decrements in food intake, since these values of the parameter cannot be investigated in a steady state experiment without endangering the animal's health.

We also intend to do experiments manipulating the social variables considered to be critical for the interpretation of some phases of our results. For example, it has been noted that an animal's disposition to exit from his work cubicle may frequently be controlled by his interpersonal relationships with the subjects who are currently in the social area. Conversely it is also believed that frequently his disposition to enter his work cubicle will be controlled by these interpersonal relationships. We plan to study these variables by systematically manipulating them. This will be accomplished via the rigid control of the social environment.

References

Boren, J. J. Response rate and resistance to extinction as functions of the fixed ratio. Unpublished Ph.D. dissertation, Columbia University, 1953.

Clark, F. C. The effects of deprivation and frequency of reinforcement on variable-interval responding. J. exp. Anal. Behav., 1958, 1, 221-228.

Ferster, C. B. Intermittent reinforcement of matching to sample in the pigeon. J. exp. Anal. Behav., 1960, 3, 259-272.

Ferster, C. B. Sustained behavior under delayed reinforcement. J. exp. Psychol., 1952, 4, 27-45.

Ferster, C. B. and Hammer, C. Variables determining the effects of delay of reinforcement. J. exp. Anal. Behav., 1965, 8, 243-254.

Ferster, C. B. and Hammer, C. Synthesizing the components of arithmetic behavior. In Honig, W. (Ed.) Operant Behavior and Psychology, New York: Appleton-Century-Crofts, 1964, in press.

Ferster, C. B. and Skinner, B. F. Schedules of Reinforcement. New York: Appleton-Century-Crofts, 1957.

Findley, J. D. An experimental outline for building and exploring multi-operant repertoires. J. exp. Anal. Behav., 1962, 5, 113-166.

Kelleher, R. T. and Gollub, L. R. A review of positive conditioned reinforcement. J. exp. Anal. Behav., 1962, 4, 543-597.

Skinner, B. F. Superstition in the pigeon. J. exp. Psychol., 1948, 38, 168-172.

Terrace, H. S. Discrimination learning with and without "errors". J. exp. Anal. Behav., 1963a, 6, 1-27.

Terrace, H. S. Errorless transfer of a discrimination across two continua. J. exp. Anal. Behav., 1963b, 6, 223-232.