National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

Human Performance Differences Following the Learning of Behavior Sequences by Chaining Procedures

Janny M. Thompson

University of Ottawa

A thesis submitted to the School of Graduate Studies of the University of Ottawa as partial fulfilment of the requirements for the degree of Doctor of Philosophy

ISBN 0-315-85861-3

Canada

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

Dedication

This thesis is dedicated to my family in Whaletown, Heriot Bay, and North Vancouver; to Kathy Cairns; and to two teachers who made significant early contributions to my interest in and love of learning - Bev Mathews and Ed Piggott.

Acknowledgements

I would like to thank a number of people for their contributions to this thesis. Bob Watters was a tremendous teacher and supervisor - he oversaw the production of this document with skill and patience. Rob Cairns provided emotional (and financial) support as well as spending hours writing the computer programs for the experiments. Andrea Boucher-Shaw was an outstanding research assistant - generous with her time, fast, organized, and thorough. Vicki Douglas enthusiastically recruited subjects. Vicki Jenkins helped collect data. My thesis committee, Pierre Baron, Cathy Bielajew, and Pierre Mercier were prompt and helpful throughout. Hundreds of students volunteered to push buttons in the lab. Many friends listened to the story of this thesis for years (and also pushed buttons) - in particular, thank you to Andrea Nauf, Tamara Walsh, Kim Burgess, Mickie Blain, Melissa Mitchell, and Kathy Scott.

Table of Contents

Abstract

Chaining procedures (forward chaining, backward chaining, and whole task training) are used to teach sequences of behaviors to diverse populations. Recent research comparing the three procedures has revealed differences between the procedures in terms of acquisition variables (e.g., number of errors made during learning, time required to the learn a sequence). Much less research has examined whether the procedures produce differences in performance following acquisition.

A series of six experiments investigating performance differences following exposure to four chaining procedures are presented. The four chaining procedures included the three traditional chaining procedures - forward, backward, and whole task - as well as a random chaining technique.

Overall, results indicated that forward chaining led to the most accurate and fastest performance of responses in the absence of the external antecedent stimuli used during training. Forward chaining also led to more accurate ordering of these stimuli following training. Random chaining led to the most accurate and the fastest performance of responses following presentation of the antecedent external stimuli out of sequence order. The theoretical and practical implications of these findings are discussed.

Human Performance Differences Following the Learning of Behavior Sequences by Chaining Procedures

Dividing a behavior into a sequence of discrete responses can facilitate both teaching and learning of the behavior. For example, a dance routine, an assembly task, the operation of an automatic teller machine, or making a bed may be divided into a number of discrete responses. Behavioral sequences are usually conceptualized as series of stimuli and responses. A dance routine, for example, may be divided into a sequence of dance steps, each of which is performed in response to a particular musical stimulus. The routine would consist of stimulus 1 (bar of music) followed by dance step 1 followed by stimulus 2 (another musical stimulus) followed by dance step 2 followed by stimulus 3 and so on. Three techniques commonly used for teaching sequences of behavior are forward chaining, backward chaining, and whole task training.

In forward chaining, the subject is taught the first behavior in the sequence, then the first behavior followed by the second behavior, and so on, until the entire sequence is performed in order. In backward chaining, the subject is first taught the final behavior of the sequence. At the next step of backward chaining, the subject is taught the second last followed by the last behavior; then the subject is taught the third last followed by the second last, followed by the final behavior. This process is continued until the entire sequence is performed

in order. In whole task training, the subject attempts the entire sequence of behaviors on each learning trial until the sequence is mastered.

Chaining procedures have been used to teach a wide variety of behaviors to many populations. For example, chaining procedures have been used to teach retarded persons self-care skills. Azrin, Schaeffer, and Wesolowski (1976) used forward chaining to teach profoundly retarded persons to dress. Giles and Morgan (1989) taught a post herpes simplex encephalitis patient personal hygiene behavior using chaining. Chaining procedures are also commonly used in vocational training programs to teach mentally disabled persons complex assembly tasks (Bellamy, Horner, & Inman, 1979). For example, Martin, Koop, Turner, and Hanel (1981) used backward chaining and whole task training to teach severely retarded persons to assemble telephone operator's headsets and bicycle turn signals.

Chaining procedures have also been used to teach non-handicapped populations a variety of behaviors. Wightman and Sistrunk (1987) used a chaining procedure to teach college students aircraft carrier final approach skills in a flight simulator. Simek and O'Brien (1989) used chaining to teach little leaguers how to hit a baseball. Weiss (1978) and Watters (1989) taught university students sequences of button presses using chaining procedures.

Two lines of chaining research will be reviewed here, with a focus on human learning studies: literature on the acquisition of behavioral sequences by chaining procedures and literature on post-acquisition performance following training.

Acquisition of Sequences by Chaining Procedures

Although chaining procedures are a popular and effective set of behavior modification techniques, the choice among them has raised questions for many years.

Choosing a Chaining Procedure

Until recently, backward chaining was recommended as the chaining procedure of choice (e.g., Kazdin, 1989; Martin & Pear, 1983; Rachlin, 1976). This recommendation was based on a theoretical analysis of chaining that gave a critical role to conditioned reinforcers in learning a sequence of behaviors (e.g., Schwartz, 1984).

It was hypothesized that, during backward chaining, the stimulus between any two behaviors in the sequence was likely to become both a conditioned reinforcer for the preceding behavior and a discriminative stimulus for the following behavior. During backward chaining, the final behavior in the sequence is learned first and correct performance of this behavior is reinforced. Any stimulus antecedent to the correct performance of the behavior will theoretically become a conditioned reinforcer for the second last behavior, as well as a discriminative stimulus for the final behavior. At the next step in backward chaining,

the second last behavior followed by the final behavior is learned. The conditioned reinforcer established at the first step of training is now contingent upon the correct performance of the second last behavior. Therefore, any stimulus antecedent to the performance of the second last behavior followed by the last behavior, becomes a conditioned reinforcer for the third last behavior and a discriminative stimulus for the second last behavior.

This process continues as each behavior is added to the sequence, so that the power of the reinforcer at the end of the sequence filters down (into conditioned reinforcers) and so that discriminative stimuli are developed for each behavior. Theoretically then, backward chaining gives the learner two advantages: (1) a conditioned reinforcer following each behavior which should increase the probability of performing the behavior again in the future, and (2) a discriminative stimulus preceding each behavior which should set the occasion for performing the behavior in the future.

However, this theoretical advantage is based on higher order conditioning, a process that is noted for its weakness (Schwartz, 1984). Higher order conditioning is the process through which a stimulus becomes a conditioned reinforcer by virtue of a contingent relationship with another conditioned reinforcer. In backward chaining, assuming the reinforcer for the final behavior is a primary reinforcer, the conditioned reinforcers that follow

the third last and earlier responses in the sequence gain their reinforcing power through higher order conditioning. In cases other than higher order conditioning, a stimulus becomes a conditioned reinforcer by virtue of a contingent relationship with an unconditioned reinforcer. A contingent relationship with an unconditioned reinforcer is a more effective means of establishing a conditioned reinforcer than is a contingent relationship with a conditioned reinforcer.

More recently, for applied reasons, whole task training has been recommended as the procedure of choice for teaching retarded persons behavioral sequences (Martin & Pear, 1988). Whole task training has been recommended for this purpose based on two sets of considerations: ease of teaching and ease of learning.

Bellamy, Horner, and Inman (1979) suggest that whole task training requires teachers to spend less time assembling and disassembling when teaching assembly tasks and it teaches response topography and sequence simultaneously. In forward chaining, the teacher must disassemble the object following each performance of a subsequence (e.g., after parts 1 and 2 are assembled in order then after parts 1, 2, and 3 are assembled in order and so on). In backward chaining, the teacher must assemble the object except for the final part, then disassemble the object. At the next step, the object must be assembled except for the last two parts then disassembled and so on. Less time is spent assembling and disassembling during whole task

training because the learner attempts the entire sequence on each trial.

Attempting the entire sequence on each trial also leads to the second advantage, learning response topography and sequence simultaneously. The second advantage seems less unique to whole task training than does the first advantage because in forward and backward chaining response topography (overall form of the final behavior) and sequence (order of responses in the final behavior) are also learned simultaneously, although in smaller subsections. Bellamy, Horner, and Inman (1979) also state that whole task training should lead to faster learning because there is less repetition and should therefore maximize the learner's independence early in training.

Empirical evidence from human learning research has not supported the theoretical superiority of backward chaining but has partially supported the recommendation of whole task training. Empirical evidence has also suggested that forward chaining may have advantages over the other chaining procedures.

Empirical Evidence Regarding Acquisition by Chaining

Studies comparing the three training procedures on the number of errors made during acquisition have demonstrated that forward chaining is often the better procedure. Weiss (1978) taught university students sequences of six button presses by forward and backward chaining and found that all subjects made

fewer errors during forward chaining than during backward chaining.

Watters (1989) compared all three training procedures (forward, backward, and whole task) for teaching university students sequences of five button presses on a computer keyboard and found that subjects made more errors when taught by whole task training than when taught by forward or backward training. In a subsequent study (Watters, 1990) examining sequences of 8 and 10 button presses, subjects made the most errors during whole task, and the fewest during forward training.

Studies examining acquisition time have provided some support for the superiority of whole task training. Martin et al. (1981) taught severely retarded individuals vocational assembly tasks by backward chaining and whole task training and found that the tasks were mastered more quickly by whole task training. Spooner (1984) compared backward chaining and whole task training for teaching severely retarded persons assembly tasks. The rate of learning was faster for whole task training than for backward chaining. Similarly, Watters and Scott (in press) found that sequences of button presses were performed to criterion more quickly by university students when the sequences were taught by whole task training than when they were taught by forward or backward chaining. Results related to acquisition time support the recommendation of Bellamy, Inman, and Horner

(1979) to use whole task training because it should lead to faster learning.

Martin et al. (1981) also recommended the use of whole task training for teaching retarded persons. They provided an explanation for the fact that the superiority of backward chaining had not been demonstrated in the case of human learning. They state that a difference between the use of chaining in studies of human and animal learning is the nature and use of reinforcement.

In animal learning studies, where backward chaining has been the procedure of choice, a single primary reinforcer is presented after performance of the final response in the sequence. This method of reinforcer delivery allows for the "filtering down" of reinforcement strength that is important for the theoretical superiority of backward chaining. By contrast, in human learning studies of chaining, most of which are conducted in applied settings with retarded persons, praise is generally used as a reinforcer. Rather than one reinforcer being delivered at the end of training, praise is delivered as various behaviors are added to the sequence. They propose that this frequent use of reinforcement (praise) throughout each of the chaining procedures may account for the failure to find backward chaining the superior procedure with human subjects in applied settings.

Experiments with human subjects by Watters and colleagues have two characteristics that may account for the failure to find

backward chaining to be the superior procedure. The behaviors in the sequences (button presses) are not very distinct which may impede the development of discriminative stimuli. In backward chaining, the stimulation produced by performing the responses may become discriminative stimuli for subsequent behaviors. One of the key recommendations for the effective establishment of discriminative stimuli is the use of distinct stimuli (Martin & Pear, 1983, 1988). If the response produced stimulation for each response is very similar, as is the case with button presses, the development of this stimulation into discriminative stimuli is likely to be very slow. The reinforcers used in the Watters experiments (percent correct scores, proceeding to the next subsequence) are relatively weak and may impede the filtering down effect expected in backward chaining (Scott & Watters, 1992). Again, it should be noted that the filtering down of reinforcement strength in the theoretical model relies on higher order conditioning which is not a particularly reliable means of establishing conditioned stimuli, even with animals (Schwartz, 1984).

Other research on the acquisition of behavioral sequences has investigated the effect of procedures overlaid on the chaining procedures (e.g., time-out, prompts). Chartrand (1990) examined the impact of 0, 3, 7, and 12 s time-out periods following errors during the learning of an 8 button press sequence by university students. When 7 s and 12 s time-outs

were administered during whole task training, subjects made significantly fewer errors and total responses. Scott (1990) did not find any significant impact on errors when examining consequences for errors (auditory cues, visual cues, and time-out) during learning a sequence of button presses. Prompts that identified the correct button push after an error had been made reduced the number of random errors (errors made before a correct response had occurred) made during each training procedure.

The bulk of research comparing chaining methods has examined what can be referred to as acquisition variables (e.g., errors during learning, time to learn sequence, trials to criterion, total responses during learning) and the influence of additional procedures applied during training (e.g., varying consequences following errors, prompts). Little research has investigated what subjects learn or can do following training and whether this differs depending on the chaining procedure utilized during training. The present research was motivated by an interest in describing and explaining such post-acquisition differences and in answering the more general question: What can subjects do following learning by the various chaining procedures?

Post-Acquisition Performance Following Chaining Procedures

One means of addressing this question is to examine retention following exposure to the chaining procedures. Watters (in press) randomly assigned university students to forward chaining, backward chaining, and whole task training groups and

taught them eight button press sequences on a computer keyboard. During training, each button press was preceded by a particular external stimulus (a symbol on the monitor screen). Subjects were asked to repeat the sequence after a 20 minute delay with the aid of the external antecedent stimuli that had been used during acquisition. After the delay, virtually no errors were made following any of the chaining procedures. In a second experiment, subjects were asked to repeat the sequence after a delay of 5 days. In this experiment, the group that had learned the sequence by backward chaining made more errors than the other two groups. Also, the forward chaining group tended to make more errors near the end of the sequence and the backward chaining group tended to make more errors near the beginning of the sequence.

Nannay (1971) taught subjects to assemble a dado head and place it on a radial arm saw. Two weeks following training, subjects were asked to complete the task again. The forward and backward chaining groups performed better on the retention test than did a no treatment control group. There was no difference between the forward and backward chaining groups.

Wilcox (1974) taught college students two sequences, a sequence of origami folds and a sequence of arithmetic operations by forward, backward, and whole task training. Three lengths of sequence (short, medium, and long) were examined. Subjects learned the components of each task from a programmed instruction

book that presented the components individually. The retention task involved performing the origami task and the arithmetic operations to a criterion of one correct performance of the task without the use of prompts from the programmed instruction book. Retention was assessed one week later by measuring the time to criterion, trials to criterion, and the number of prompts received to criterion. Forward chaining led to significantly better retention of the origami tasks. There was a trend for the backward chaining group to perform better at retention on the arithmetic sequences, though the comparisons with other chaining procedures were not significant. Generally, then, forward chaining and backward chaining led to superior retention than whole task training.

Cornell (1975) used forward and backward chaining procedures to teach second grade children to spell groups of 12 words. Retention was measured four weeks later. No differences were found between the chaining groups at the follow-up.

Nettelbeck and Kirby (1976) used (a) whole task training, (b) a variation of forward chaining, (c) a progressive part, and (d) a pure part training procedure to teach mildly mentally retarded workers to thread an industrial sewing machine. The part training procedures were similar to chaining in that they involved dividing the task into component responses, teaching each component, and then chaining the components together. The threading task was divided into 4 components. One month later,

retention of the task was assessed by asking the workers to complete the threading task. There were no significant differences in retention among the treatment groups as measured by the number of trials to criterion, number of errors to criterion, and time to criterion.

Weber (1978) taught educable mentally retarded adults to assemble plastic gears and levers by either forward chaining or backward chaining. Retention of the tasks was assessed approximately 20 hours after training by measuring the number of steps that were performed correctly, in order, prior to making an error. No difference was found between the two chaining procedures.

Another means of investigating what is learned or can be done following learning a behavior sequence by a chaining procedure is to look at what subjects can do following training besides repeat the sequence. Thvedt, Zane, and Walls (1984) used this strategy to evaluate two positions on how sequences are learned. The positions are derived from Skinner's (1938) Law of Chaining which states that "the responses of one reflex may constitute or produce the eliciting or discriminative stimulus of another" (p. 32). The two positions Thvedt, Zane, and Walls derived from this law were: (a) Position 1, that each response in a behavioral sequence constitutes the discriminative stimulus for the next response in the sequence, and (b) Position 2, that each response <u>produces</u> the discriminative stimulus for the next

response. A discriminative stimulus for a particular response is an antecedent stimulus in the presence of which that response has been reinforced.

Thvedt, Zane, and Walls (1984) used a combination of modelling and whole task training to teach mentally retarded adults a sequence of behaviors for assembling a circuit board. The sequence involved placing six resistors on a circuit board and pressing a switch after each had been placed. The switches each illuminated a light located directly above them. The lights remained on for one second. The sequence of behaviors and stimuli was: place resistor 1, press switch 1, light 1 illuminates for 1 s; place resistor 2, press switch 2, light 2 illuminates for 1 s; and so on. The sequence of behaviors was random with respect to the location of the switches (i.e., the sequence did not proceed from left to right along the row of switches).

Following training to criterion, subjects' performance of the sequence was tested under three conditions: altered stimulus location, altered stimulus sequence, and missing stimulus. The altered stimulus location condition involved one of the lights being illuminated following a different switch press than had led to its illumination during training. The altered stimulus sequence condition involved one of the lights being illuminated following the placement of a resistor rather than the following a switch press, as had been the case during training. During the

missing stimulus condition, one of the lights did not illuminate following the usual switch press.

The authors suggested that, if a light (stimulus) was illuminated in an altered stimulus location or stimulus sequence, and subjects continued to emit responses in the correct order, this might support Position 1. Continuing to emit responses in sequence order may imply that the responses constitute discriminative stimuli for subsequent responses (Position 1) or that some response produced stimulation constitutes the discriminative stimuli for subsequent responses. Alternatively, if subjects emitted the response that had followed the out of order stimulus in the original sequence, this would clearly support Position 2.

The results did provide some support for the position that responses in a sequence produce stimuli that function as discriminative stimuli for subsequent responses (Position 2). In the altered stimulus condition, subjects continued to perform the sequence in correct order as if the sequence had not been disrupted. In the missing stimulus condition, only about half of subjects continued to perform the sequence in the correct order.

This study is of interest here because it investigated performance following chaining and because it suggests a means of investigating such performance (i.e., the administration of various tests following acquisition). However, it is imperfect in two regards. A pure chaining procedure was not used to teach

the sequence, and only one procedure for teaching a sequence was examined. A stronger test of the theoretical positions underlying chaining would have been to teach sequences by each of the three chaining procedures, without the aid of modelling, and then test performance. More information about both positions should be found comparing all three types of chaining. Investigating all three chaining procedures would allow for examination of differences in performance across types of tests as well as across types of training.

Aside from the Watters (in press) study of retention, little research has systematically investigated whether, given the characteristic differences among the three chaining procedures during acquisition, the procedures produce differences in performance at the end of training. One clear pattern that emerges from empirical research on the acquisition of behavioral chains is that forward chaining leads to fewer errors than the other procedures whereas whole task training leads to briefer acquisition time. It seems reasonable to expect that tests of various kinds of post-acquisition performance would yield differences depending upon training procedure. For example, if a training procedure produces more errors than another procedure during acquisition, this difference might be reflected in some measure of post-acquisition performance.

The research on retention of behavioral sequences and post-acquisition performance has not provided a consistent picture of

post-acquisition performances following chaining procedures. Researchers have seldom compared all three chaining procedures and the method of assessing post-acquisition performance has been variable. Some studies have required subjects to relearn the sequence to criterion with the aid of discriminative stimuli used during training whereas others have required subjects to perform the sequence without the aid of discriminative stimuli used during training. Performance has been assessed by measuring time to criterion, trials to criterion, and number of correct responses prior to an error.

Research Rationale

Given that research has revealed clear differences among chaining procedures on acquisition variables, it was hypothesized that there would be differences in performance <u>following</u> learning by the various chaining procedures. There are a number of behaviors that might be learned through exposure to chaining procedures. These behaviors include: (a) performing one response followed by another, in correct order, without the external antecedent stimuli, (b) ordering the antecedent external stimuli, and, (c) performing a correct response following the presentation of a particular antecedent external stimulus. The extent to which each of these behaviors can be performed following training will have both theoretical and practical implications. In the present series of six experiments, three tests were administered

following chaining in order to assess differences in post-acquisition performance of these behaviors.

Test 1 assessed the first type of behavior; performing one response followed by another, in the absence of the external antecedent stimuli. Test 1 required that subjects perform the sequence that they had just learned, but in the absence of the antecedent external stimuli. In theoretical terms, if the subject can perform the sequence under these conditions, response produced stimulation appears to constitute the discriminative stimulus for the next behavior. That is, stimulation produced by the performance of response 1 becomes the discriminative stimulus for response 2, stimulation produced by the performance of response 2 becomes the discriminative stimulus for response 3 and so on. The performance of this behavior following training would support the position that each response in a behavioral sequence produces proprioceptive stimulation that constitutes the discriminative stimulus for the next response in the sequence. It would also support the notion that other stimuli derived from performing the behavior (e.g., the sight of performing the behavior) may become discriminative stimuli.

In practical terms, performance of a sequence without the aid of external discriminative stimuli that were present during training will be desirable for some types of behaviors. Imagine an actor rehearsing a long soliloquy (a behavioral sequence) by using a series of flash cards (external discriminative stimuli),

on each of which is printed small segments of the speech. Two aspects of the sequence will be important once it has been acquired: the order in which the segments are performed and the conditions under which they are performed. The segments of the speech must be performed in the correct order and in the absence of the external discriminative stimuli (flash cards) used during rehearsals. These aspects of post-acquisition performance are assessed by Test 1.

A number of other characteristics of the performance of a sequence in the absence of external discriminative stimuli would be of interest: How accurately does a subject perform this behavior following training? How quickly does a subject perform this behavior following training? How many correct responses does a subject perform before making an error? What is the longest series of consecutive correct responses that a subject performs following training? Where in the sequence does this series begin? Where in the sequence are errors made? These characteristics were measured in the present experiments.

Test 2 assessed the second behavior; ordering antecedent external stimuli. For this test, subjects were instructed to arrange the antecedent external stimuli in the order in which they had occurred in the sequence. The characteristics of this behavior that were measured in the present series of experiments were accuracy (identification of the sequence position of each stimulus) and stimulus-stimulus pairs (identification of pairs of

stimuli that had followed one another in the sequence). From a theoretical perspective, accurate performance on Test 2 would suggest that the contingent relationship between the external discriminative stimuli had been learned and that they were salient at the end of training.

One reason for designing Test 2 was that, following the acquisition of some types of behavioral sequences in applied settings, it may be desirable that subjects are able to order the antecedent external stimuli. Imagine teaching a friend the behavioral sequence required to find a particular address. This sequence will consist of a series of discriminative stimuli (e.g., the corner with the grocery store on it, the street sign that says "Wilbrod") that set the occasion for particular responses (e.g., continuing for one block, turning right). Following acquisition, it may be desirable that the subject is able to order the antecedent stimuli to facilitate their own performance on future occasions or to communicate the directions to others.

Test 3 assessed the third behavior; performing the correct responses following the presentation of antecedent external stimuli that had been used during training, when the antecedent stimuli are presented in a different order than during training. From a theoretical perspective, if the correct responses are made following the presentation of the antecedent external stimuli, then these stimuli appear to play an important controlling role

for the responses. For Test 3, the antecedent external stimuli used during training were presented out of order and subjects were instructed to perform the component responses of the sequence in response to the antecedent external stimuli. Subjects' performance of this behavior was measured in terms of accuracy, speed, and sequence location of symbols on which errors were made.

An example of the type of practical situation in which Test 3 may be of interest is that of a student studying for an arithmetic test on the multiplication tables. Even though during training the tables will be run through in an orderly way by incrementing on of the two factors being multiplied, the student will wish to learn to respond correctly to the questions (discriminative stimuli), in whatever order they are presented subsequently.

It is likely that subjects learn some combination of the three behaviors that the Tests were designed to assess when exposed to each of the training procedures but that particular training procedures will lead to better learning of some behaviors than will other training procedures. What subjects learn may also depend on the stage of training. For example, the stimuli controlling performance of the sequence may change depending upon how much a subject has practiced the sequence, as well as the procedure that was used to teach the sequence. The stimuli controlling behavior may shift as training progresses.

For example, early in training, antecedent external stimuli may play an important controlling role. Later in training, stimulation produced by the performance of the responses themselves may play an important controlling role.

As well as being of theoretical interest, what is learned during each chaining procedure is of practical interest. Although the three procedures for teaching sequences of behavior are usually treated as three means of achieving the same goal (i.e., teaching a subject to perform a sequence in the correct order), this does not mean that they are similar when subjected to a more detailed analysis. A more detailed analysis of how each procedure works may reveal differences among them. Looking in more detail at what subjects can do following exposure to each procedure has practical implications. One procedure may be better for learning something in particular about the sequence. For example, one training procedure may lead to better performance of component responses following antecedent external stimuli when the stimuli are presented out of sequence order. Such a training procedure would allow subjects to adapt more easily to subsequent changes in the desired order of the component behaviors. A training procedure that allowed subjects to adapt more easily to changes in the desired order of component behaviors would be useful for teaching behaviors like pressing piano keys in response to sheet music or spelling words in response to spoken words.

Chaining has been used to teach a wide variety of behaviors. Depending on the behavior taught, the desirable post-acquisition performance may vary. Behaviors like assembly tasks (e.g., assembly of radial saw parts, gears and levers, telephone operator headset, bicycle turn signal, circuit board) consist of responses that must be performed in a particular order for the end product to be functional. Similarly, tasks such as threading a sewing machine, dressing, washing, origami, landing an air craft, consist of sequences of responses that must be performed in a particular order, or at least with little variation in this order. For these types of behaviors, in which the ordering of the responses in the sequence is relatively fixed, performance of the whole sequence in order, without the aid of instructions or prompts (antecedent external stimuli) that might have been used during learning, would likely be most desirable post-acquisition behavior.

There are other types of behaviors, however, for which performance of the whole sequence in order may not be the most desirable or the only desirable post-acquisition performance. Task such as solving series of multiplication table problems, . spelling series of words, and pressing piano keys in response to written musical notes do not have fixed orders. The multiplication problems can be solved in any order, the words can be presented in any order, and the order of the piano key presses will depend on the tune. For behaviors such as these, it may be

equally desirable or even more desirable that the subject be able to perform the sequence in some other order than the learning order.

Despite the fact that the nature of the desired post-acquisition performance likely differs as a function of the task taught, those studies of chaining that have included a follow-up have typically employed one measure of post-acquisition performance across tasks and chaining procedures (Cornell, 1975; Nannay, 1971; Nettelbeck & Kirby, 1976; Watters, in press; Weber, 1978; Wilcox, 1974).

The following series of six experiments were an investigation of post-acquisition differences following exposure to various chaining procedures. The experiments examined differences among chaining procedures on the three types of post-acquisition behaviors, using the three tests described above. Subjects were taught sequences of key presses on a computer keyboard. The stimulus-response sequences consisted of a series of external stimuli presented on the monitor, one at a time (e.g., *, @) each followed by a particular key press.

Experiments 1, 2, and 3 examined post-acquisition performance following forward chaining, backward chaining, and whole task training. Experiments 1 and 2 investigated differences in the performance of one response followed by another in the absence of antecedent external stimuli (Test 1) and the ordering of antecedent external stimuli (Test 2).

Experiment 3 used Test 3 to explore differences in the performance of responses following presentation of the antecedent external stimuli in an order other than that used during acquisition.

Experiments 4, 5, and 6 explored post-acquisition performance following forward, backward, and random chaining. Random chaining involves adding randomly selected component responses during training rather than systematically adding adjacent component responses as is the case for forward and backward chaining. The randomly chosen responses are ordered during training in the order in which they appear in the final sequence. For example, the subject might be taught response 4; then response 1 followed by response 4; then response 1 followed by response 4 followed by response 9 and so on.

Experiment 4 used Test 3 to examine performance following forward, random, and backward chaining. Test 3 was also used in Experiment 5, where post-acquisition performance was examined at two levels of proficiency criteria. Proficiency criteria referred to the number of consecutive correct performances of the sequence required to end training. For forward, backward, and random chaining the proficiency criterion also applied to the number of consecutive correct performances of each subsequence that was required before another response was added to the subsequence. Experiment 6 examined performance on all three tests following forward, backward, and random chaining.

The series of experiments was undertaken from a radical behavioral perspective which places an emphasis on the prediction and control of behavior through the application of experimentally derived principles of operant conditioning. In keeping with this perspective, the results of the experiments are described with reference only to those behaviors that were observed to change and not with reference to hypothetical intervening processes that might, in other perspectives, be seen as explanations of the observed behavior changes. The experiments are controlled laboratory experiments. Ross (1985), in his presidential address to the Association for the Advancement of Behavior Therapy, argued that there was a need to enhance both the theoretical and empirical soundness of behavior therapy techniques by returning to the laboratory. By this he meant that laboratory research should not be carried out in a clinical vacuum and that clinical research should not be carried out in isolation from laboratory derived principles. The purpose of returning to the laboratory for these experiments was to use a systematic and controlled experimental method to discover the lawful nature of post-acquisition performance following chaining procedures, as applied investigations had not yet produced consistent results.

Experiment 1

The purpose of Experiment 1 was to explore performance differences following forward chaining, backward chaining, and whole task training, using two types of post-acquisition tests. The tests were administered following training to criterion. Test 1 assessed whether subjects could perform the sequence of button pushes in the absence of the antecedent external stimuli provided during training. Test 2 assessed whether subjects could order the antecedent external stimuli when the symbols were presented in a mixed up order.

Based on the reasoning presented below, it was hypothesized that there would be post-acquisition performance differences and that forward chaining would lead to superior post-acquisition performance. More specifically, it was hypothesized that subjects would perform the sequence of button pushes more accurately and more quickly in the absence of antecedent stimuli (Test 1) following forward chaining than following backward chaining or whole task training. It was also hypothesized that subjects would order the antecedent external stimuli more accurately (Test 2) following forward chaining than following backward chaining or whole task training.

It was hypothesized that forward chaining would lead to superior performance on Test 1 because, during training by forward chaining, additional responses were added at the end of each subsequence until the complete sequence was performed to

criterion. The performance of each response in forward chaining should therefore become a reliable discriminative stimulus for the next response, thereby facilitating Test 1 performance. During backward chaining, additional responses were added at the beginning of the subsequence, making the screen symbols more reliable discriminative stimuli than the chaining sequence of responses themselves. Whole task training was hypothesized to lead to poorer performance because research on acquisition variables suggested that it led to more errors during acquisition. The relative roles of the responses and the screen symbols as discriminative stimuli during whole task training are not as clear as for forward and backward chaining. However, it was assumed that the high number of errors made during acquisition by whole task training might indicate slow development of effective stimulus control. The hypotheses regarding Test 2 performance were exploratory and based on the assumption that forward chaining would lead to the most repetitive experience of the ordering of the external antecedent stimuli.

Method

Subjects

Eighteen volunteers (17 females, 1 male), from an undergraduate psychology class served as subjects. The average age of subjects was 24.2 years (*SD* = 7.1, range 19 - 44).

Apparatus

The subject sat at a table, on the edge of which sat a standard Commodore 64 keyboard. Behind the keyboard was a 14 inch (36 cm) diagonal Commodore video monitor (model 1902A). A Commodore 64 computer counted the number of errors at each step of training, recorded which keys were pressed and when during the post-acquisition test, and timed the duration of training and test periods. This information was printed by a Commodore printer (model MP3803) after each subject had completed the post-acquisition test. A cardboard screen was placed over the keyboard, exposing only those ten keys the subject was required to use for each sequence (keys R, T, Y, U, I, F, G, H, J, K). These ten keys were covered with masking tape so that subjects could not see any of the letters or symbols on them.

Procedure

Subjects were run individually. Once the subject was seated in front of the computer monitor, instructions for the experiment were read aloud. The instructions explained that the subject would be learning three different sequences of ten button pushes, each button push preceded by a unique screen symbol, and that the computer would produce a "zit" sound when they pressed the wrong button (see Appendix A for complete instructions). Subjects were asked to read and sign an informed consent form before proceeding with the experiment (Appendix B).

The experimenter sat in an adjoining room while the subject was learning a sequence. Before the experimenter left the subject in the laboratory, she turned off the overhead lights, leaving a small lamp on in the corner of the room and started the computer program.

The procedure for teaching the sequences involved the presentation of symbols (e.g., *, :, +), one at a time, on the monitor screen to cue the subject to press buttons. Each symbol was associated with a particular button. If the subject pressed the correct button when a symbol appeared, the symbol would disappear, and the next symbol would be presented. If the subject pressed an incorrect button, the computer made a "zit" sound and the symbol remained on the screen until the correct response was made. When a subject completed a subsequence, the computer produced a "boing" sound and the percent correct score appeared on the monitor screen. The score then disappeared, and the computer produced a "boing" sound prior to presenting the first symbol of the next subsequence. During forward chaining, the first symbol presented was the first one of the sequence. During backward chaining, the first symbol presented was the last one of the sequence.

Each subject learned three sequences, one each by forward chaining, backward chaining, and whole task training. Each subject learned the same three sequences. The three sequences used the same ten keys in a different order and each involved

different screen symbols. The criterion to end training was three consecutive correct performances of the ten button press sequence. For forward and backward chaining, three consecutive correct performances of each subsequence was required before another key was added to the subsequence. For example, in forward chaining the subject was required to respond correctly to the subsequence consisting of the first symbol on three consecutive presentations before the computer initiated training of the subsequence consisting of the first symbol followed by the second symbol. After the performance of each subsequence or sequence, the percent correct was displayed on the screen. To control for order effects, the presentation order of the training procedures was counterbalanced such that three subjects learned the sequences by each of the six possible orders.

When the subject had performed a sequence to criterion, Test 1 was administered by the computer. The experimenter did not return to the laboratory to give instructions for Test 1. The monitor screen went blank and the computer produced the usual "boing" sound to signal to the subject to perform the sequence again. The subject was then required to complete the sequence they had just learned without the presence of the screen stimuli. During this test, if the subject made an error the computer made the "zit" sound and waited for the correct response just as it had during training. When Test 1 was completed, the computer displayed the message "Please tell the experimenter you are

finished" and the printer began printing the data. The experimenter returned to the laboratory and chatted with the subject until the printing was complete, restarted the program for the teaching of the second sequence, and left the laboratory. The same procedure was followed when the printer began printing the data following Test 1 for the second sequence. Following Test 1 for the third sequence, the experimenter returned to the laboratory and administered Test 2 (Appendix C). Test 2 presented the screen stimuli used in the final sequence in a mixed up order on a sheet of paper. The subject was requested to number the symbols so that the numbers represented the order in which the symbols had appeared in the final sequence. No feedback regarding accuracy was given to the subject during Test 2.

Variables and Design

The independent within subjects variable was training procedure (forward chaining, backward chaining, whole task training). Subjects completed Test 2 following the last sequence only so that six subjects completed Test 2 following each training procedure, hence Test 2 performance was a between subjects comparison.

Test 1 variables. Dependent variables related to Test 1 were percent correct, consecutive correct responses, completion time, and location of Test 1 errors.

Ten key presses were necessary to complete Test 1 correctly (i.e., consecutively pressing keys 1 to 10). The percent correct on Test 1 was calculated as follows: (10/total responses) * 100. Completion time was measured from a subject's first key press on Test 1 to a subject's last key press on Test 1.

Consecutive correct responses were examined using three variables. The number of correct responses prior to the first error on Test 1 referred to the number of correct responses made prior to the first error. The longest series of correct responses on Test 1 was number of consecutive correct responses in the longest series of such responses. The sequence position of the first response in the longest series of correct responses referred to the original sequence position of the response. For example, if a subject made the following key presses on Test 1: 1 - 2 - 3 - 4 - 6 - 8 - 5 - 1 - 6 - 7 - 8 - 9 - 10; the largest number of consecutive correct responses was 5 (responses 6 to 10) and the sequence position of the first response of the largest number of correct responses was 6. In cases where a subject made two equally long series of correct responses, the sequence position was the mean of the sequence positions of the first response in each series.

The location of errors on Test 1 was calculated in two steps. For each response in the sequence the number of errors made prior to the first correct response was counted. Then these errors were summed for the first 5 responses and for the last 5

responses in the sequence to provide a measure of errors made at the beginning and at the end of the sequence, respectively.

Test 2 variables. The dependent variables related to Test 2 were percent correct, correct pairs, and reverse pairs. In order to determine percent correct, Test 2 was marked as follows: if a subject correctly identified the sequence position of a symbol (e.g., responded that the 3rd symbol had been 3rd in the sequence), the response was considered correct. This method of marking Test 2 provided only an index of whether subjects could identify the exact sequence location of each symbol taken individually.

For this reason, the test was also checked for two types of pairs: correct pairs and reverse pairs. For example, assume that a subject completed a Test 2 after learning a sequence consisting of 6 symbols and responses and, on Test 2, responded that the first symbol had been 3rd, the second symbol 4th, the third symbol 5th, the fourth symbol 6th, the fifth symbol 1st, and the sixth symbol 2nd. The test would have been scored as containing 4 correct pairs (3-4, 4-5, 5-6, and 1-2) and 1 reverse pair (5-4).

This method of marking Test 2 allowed for assessment of whether subjects could identify whether pairs of symbols had been adjacent in the sequence even if they could not identify the actual sequence position of the symbols. For example, a subject

might report that a "*" had followed a ">" without reporting that these symbols were the 4th and 5th symbols in the sequence.

Acquisition variables. Dependent variables measured during acquisition were total errors, true errors, total responses, and learning time. True errors were errors made in response to a symbol to which the subject had already responded correctly. Total errors were the sum of true errors and errors made in response to a symbol to which the subject had not yet made a correct response. Total responses were defined as the number of responses, including errors, that were made during acquisition of the sequence. Learning time was measured from the presentation of the first screen symbol to the last button press at the end of the criterion performance that ended acquisition.

Data Analyses

Data analyses were performed using the Statistical Package for Social Science (SPSS^{x}). Descriptive statistics were calculated for each variable. Two repeated measures multivariate analyses of variance (MANOVAs) were used to examine overall difference among the chaining procedures on two sets of variables: (a) Test 1 variables (percent correct, correct responses prior to the first error, longest series of correct responses, sequence position of first response in the longest series of correct responses, and completion time) and (b) acquisition variables (total error, true errors, total responses, learning time). When a repeated measures MANOVA reached

multivariate significance, post hoc dependent pairs *t*-tests were performed to examine differences between chaining procedures on variables that reached univariate significance. In this and all subsequent experiments, the Bonferroni procedure was used to adjust the alpha level (.05) for the multiple *t*-tests when reporting the results. A repeated measures ANOVA was used to examine overall differences among the chaining procedures on the location of errors on Test 1 (beginning and end).

For between groups effects (i.e., Test 2 variables), a MANOVA was performed. Providing the MANOVA reached multivariate significance, one-way ANOVAs and the Student Newman Keuls procedure for multiple paired comparisons (with alpha protection) were used to follow-up variables that reached univariate significance on the MANOVA. Pillais trace was used as a conservative test of multivariate significance for all MANOVAs, including the repeated measures MANOVAs.

Exact probabilities for all inferential statistics are reported in Appendix D. Two-tailed significance levels are reported for the *t*-tests. Only significant results are reported. Any comparisons not reported did not reach significance.

Results

Test 1: Performance of the Sequence in the Absence of the Symbols

Descriptive statistics for Test 1 variables are presented in Table 1. A repeated measures MANOVA for the five Test 1 variables did not reach multivariate significance. For

descriptive purposes, means for the dependent variable are reported below.

Percent correct. Mean percent correct for forward, backward, and whole task training were 95.0%, 94.5%, and 82.3%, respectively.

Consecutive correct responses. The mean number of correct responses made prior to the first error on Test 1 for forward, backward, and whole task training were 8.2, 7.2, and 7.8, respectively. For the longest series of correct responses, the means for forward, backward, and whole task training were 9.0, 9.1, and 8.1, respectively. The mean sequence location of the first response in the longest series of correct responses on Test 1 for forward, backward, and whole task training were 1.6, 1.4, and 1.9, respectively.

Completion time. Figure 1 shows mean Test 1 completion times for the three training procedures. The mean completion times for Test 1 following forward, backward, and whole task training were 8.1 s, 15.1 s, and 21.9 s, respectively. These means did rank as hypothesized, and, although the multivariate test of significance was nonsignificant, the univariate test for completion time did reach significance ($F[2, 34] = 3.8$, $p < .05$). Follow-up dependent pairs t-tests revealed that whole task training ($M = 21.9$ s; $t[17] = -2.6$, $p < .05$) led to significantly longer Test 1 completion times than did forward chaining ($M = 8.1$ s).

Location of errors. Descriptive statistics for location of Test 1 errors are presented in Table 2. A 2 (Location, beginning or end) X 3 (Training Procedure) repeated measures ANOVA did not reveal any significant effects. The mean number of errors made at the beginning of Test 1 following forward, backward, and whole task training were 0.4, 0.3, and 2.1, respectively. The means for errors made at the end of Test 1 following forward, backward, and whole task training were 0.2, 0.5, and 1.8, respectively.

Test 2: Ordering of the Symbols

Descriptive statistics for Test 2 variables are presented in Table 3. Comparisons among the training procedures on Test 2 variables are between groups comparisons. Six subjects completed a Test 2 following each training procedure. A MANOVA performed on the three Test 2 variables (percent correct, correct pairs, and reverse pairs) did not reach significance.

Percent correct. The mean percent correct on Test 2 for forward, backward, and whole task training were 58.3%, 56.7%, and 38.3%, respectively.

Pairs. The mean number of correct pairs for forward, backward, and whole task training were 4.3, 4.5, and 4.0, respectively. The mean number of reverse pairs for forward, backward, and whole task training were 1.0, 0.8, and 0.5, respectively.

Acquisition Variables

Acquisition variables were not the primary focus of the experiment. For this reason, the results for acquisition variables are presented in less detail here. Descriptive statistics for acquisition variables can be found in Table 4.

A MANOVA performed on the four acquisition variables was significant ($F[8, 10] = 47.8$, $p < .001$). All four acquisition variables reached univariate significance.

Figure 2 shows the effect of training procedure on mean total errors during acquisition ($F[2, 34] = 26.5$, $p < .001$). Fewer total errors were made during forward chaining than during whole task training ($t[17] = -5.7$, $p < .05$). Fewer total errors were made during backward chaining than during whole task training ($t[17] = -4.8$, $p < .05$).

Figure 3 shows the effect of training procedure on mean true errors ($F[2, 34] = 22.3$, $p < .001$). Dependent pairs t-tests revealed that fewer true errors were made during forward chaining than during whole task training ($t[17] = -5.2$, $p < .05$). Significantly fewer true errors were made during backward chaining than during whole task training ($t[17] = -4.4$, $p < .05$).

Figure 4 shows the mean total responses during acquisition for the three training procedures. Training procedure had a significant effect on total responses during acquisition ($F[2, 34] = 4.7$, $p < .05$). Follow-up dependent pairs t-tests revealed that both backward chaining ($t[17] = 3.4$, $p < .05$) and whole task

training (t[17] = 2.6, p < .05) led to more responses during learning than did forward chaining.

Figure 5 shows the mean learning time for each training procedure. Training procedure had a significant effect on learning time (F[2, 34] = 3.4, p < .05). Subjects took significantly less time to learn a sequence by forward chaining than by backward chaining (t[17] = -3.0, p < .05).

Discussion

Only one dependent variable indicative of whether subjects could perform the sequence of button presses in the absence of antecedent external stimuli (completion time for Test 1) was by training procedure. Subjects completed Test 1 more quickly following forward chaining than following backward chaining or whole task training. This provides tentative support for the hypothesis that subjects would perform the sequence of button pushes more quickly in the absence of antecedent stimuli following forward chaining than following backward chaining or whole task training. This tentative evidence supports the theoretical position that, during forward chaining, response produced stimulation becomes a more reliable source of discriminative stimuli than it does during the other two training procedures.

Examination of the other Test 1 variables, all of which were directly related to the number of errors made on Test 1, suggest that the test may have been too easy to produce differences in

performance. The average percent correct scores for all three training procedures were high (>80%). Alternatively, the behaviors required by the tests may have been fully trained by the time the tests were presented.

The results for Test 2 variables did not provide support for the hypothesis that subjects would order the antecedent external stimuli more accurately following forward chaining than following backward chaining or whole task training. However, as only six subjects completed a Test 2 after each training procedure, the power to detect differences was low.

Experiment 2

The purpose of Experiment 2 was to continue the investigation of differences in post-acquisition performance following exposure to the three different training procedures. Three procedural changes were made to facilitate this goal. The sample size was increased to 24, Test 2 was administered after each sequence rather than after the final sequence only, and the criterion to end training was reduced to one correct performance of each subsequence or sequence. The last change was made in order to increase the number of errors, thereby allowing for a clearer pattern of differences to emerge between training procedures.

It was hypothesized that subjects would perform the sequence of button pushes more quickly and accurately in the absence of antecedent external stimuli following forward chaining than following backward chaining or whole task training. The rationale for these hypotheses was the same as that for Experiment 1 hypotheses. In addition, Experiment 1 findings provided tentative support for the hypothesis that subjects would complete Test 1 faster following forward chaining than following backward chaining or whole task training.

It was also hypothesized that subjects would order the screen symbols more accurately following forward chaining and backward chaining than following whole task training. This hypothesis was based on the ranking of the means for percent

correct on Test 2 in Experiment 1. Although the differences between the chaining procedures on percent correct on Test 2 were not significant, visual inspection suggested that scores were higher following forward and backward chaining (*M*s = 58.3% and 56.7%, respectively) than following whole task training (*M* = 38.3%).

Method

Subjects

Subjects were 24 university students (19 females, 5 males), primarily undergraduates. Mean age of subjects was 24.0 years (*SD* = 5.7, range 19 to 39).

Apparatus

The apparatus was the same as that used for Experiment 1.

Procedure

The procedure for this experiment was very similar to that for Experiment 1. Aspects of the procedure unique to Experiment 2 were: (a) the criterion to end training was one correct performance, (b) a message appeared on the screen prior to Test 1 ("Please perform the sequence you have just learned once more."), and, (c) subjects completed a Test 2 after each sequence.

Variables and Design

The variables and design were the same as for Experiment 1 with one exception: all variables, including Test 2 variables, were within subjects variables.

Data Analyses

Data were analyzed using the same procedures as those for Experiment 1. Repeated measures MANOVAs were used as overall tests for three groups of variables: Test 1 variables, Test 2 variables, and acquisition variables.

Results

Exact probabilities for all inferential statistics are reported in Appendix E. Only significant results are reported. Any comparisons not reported did not reach significance.

Test 1: Performance of the Sequence in the Absence of the Symbols

Medians, means, and standard deviations for Test 1 variables are presented in Table 5. A repeated measure MANOVA performed on the five Test 1 variables was significant ($F[10, 14] = 4.3$, $p < .01$).

Percent correct. Figure 6 shows the mean percent correct on Test 1 for each training procedure. The effect of training procedure on this variable approached univariate significance [$F[2, 46] = 3.2$, $p = .05$). Follow-up dependent pairs t-tests indicated that forward chaining ($M = 91.1\%$) led to higher percent correct scores than whole task training ($M = 76.7$; $t[23] = 3.1$, $p < .05$).

Consecutive correct responses. The mean number of correct responses prior to the first error on Test 1 for each of the training procedures is shown in Figure 7. Chaining procedure had a significant effect on correct responses prior to the first

error on Test 1 ($F[2, 46] = 4.4$, $p < .05$). Follow-up dependent pairs t-tests revealed that subjects made more correct responses prior to making an error following forward chaining ($M = 8.2$) than following whole task training ($M = 5.5$; $t[23] = 3.4$, $p < .05$). Following forward chaining then, subjects performed a larger section of the sequence prior to making an error than they did following whole task training.

Figure 8 shows the mean of the longest series of correct responses on Test 1 for each training procedure. Training procedure had a significant effect on the length of the longest series of correct responses ($F[2, 46] = 4.5$, $p < .05$). Dependent pairs t-tests revealed that subjects made longer series of correct responses following forward chaining ($M = 8.5$) than following whole task chaining ($M = 6.4$; $t[23] = 3.9$, $p < .05$). Training procedure did not have a significant effect on the sequence position of the first response in the longest series of correct responses. The means on this variable for forward, backward, and whole task training were 1.6, 1.8, and 2.6.

Completion time. Figure 9 shows the effect of training procedure on mean completion time ($F[2, 46] = 4.3$, $p < .05$). Dependent pairs t-tests showed that both backward ($M = 37.3$ s) and whole task training ($M = 27.2$) led to slower performance on Test 1 than did forward chaining ($M = 15.4$; $t[23] = -2.7$, $p < .05$ and $t[23] = -3.8$, $p < .05$ respectively).

Location of errors. Descriptive statistics for location of Test 1 errors are presented in Table 6. A 2 (Location of Errors) X 3 (Training Procedure) ANOVA for errors on Test 1 revealed a main effect for training procedure ($F[2, 22] = 6.3$, $p < .01$). Figure 10 shows the mean errors made on Test 1 following each chaining procedure. Follow-up dependent pairs t-tests revealed that fewer errors were made following forward chaining ($M = 1.4$) than following whole task training ($M = 4.8$; $t[23] = -2.9$, $p < .05$). No evidence was found to suggest that the training procedures made errors in different sequence locations.

Test 2: Ordering of the Symbols

Descriptive statistics for Test 2 variables are presented in Table 7. A repeated measures MANOVA for the three Test 2 variables was significant ($F[6, 18] = 3.1$, $p < .05$).

Percent correct. Figure 11 shows the mean percent correct on Test 2 for the three training procedures. The overall comparison among training procedures for percent correct on Test 2 reached univariate significance ($F[2, 46] = 8.4$, $p < .01$). Dependent pairs t-tests revealed that forward chaining (M = 64.6%) led to significantly higher percent correct scores than did backward chaining ($M = 33.8\%$; $t[23] = 4.2$, $p < .05$).

Pairs. Figure 12 shows the effect of training procedure on the number of correct pairs identified on Test 2. Training procedure had a significant effect on the number of correct pairs identified on Test 2 ($F[2, 46] = 3.2$, $p < .05$). Dependent pairs

t-tests indicated that fewer correct pairs occurred following backward chaining (M = 3.4) than following forward chaining (M = 5.1; $t[23] = 2.6$, $p < .05$). The mean number of correct pairs occurring following whole task training was 4.1.

Training procedure also had a significant effect on the mean number of reverse pairs reported on Test 2 ($F[2, 46] = 3.4$, $p < .05$). Figure 13 shows the mean reverse pairs identified following each training procedure. Dependent pairs *t*-tests revealed no significant differences. The means for forward chaining, backward chaining, and whole task training were 0.8, 1.8, and 1.7, respectively.

Acquisition Variables

Means, medians, and standard deviations for the acquisition variables are presented in Table 8. The repeated measures MANOVA for acquisition variables was significant ($F[8, 16] = 19.4$, $p < .001$).

Figure 14 shows the mean total errors for the three training procedures. The comparison among training procedures for total errors reached univariate significance ($F[2, 46] = 30.4$, $p < .001$). Follow-up dependent pairs *t*-tests revealed that fewer errors were made during forward chaining than during backward chaining ($t[23] = -3.3$, $p < .05$) and whole task training ($t[23] = -5.9$, $p < .05$). Fewer errors were made during backward chaining than during whole task training ($t[23] = -5.3$, $p < .05$).

Figure 15 shows mean true errors for each training procedure. The overall comparison of the training procedures for true errors also reached univariate significance ($F[2, 46] = 28.0$, $p < .001$). Follow-up dependent pairs t-tests revealed a similar pattern to those for total errors. Fewer true errors were made during forward chaining than during backward chaining ($t[23] = -3.7$, $p < .05$) or whole task training ($t[23] = -5.7$, $p < .05$). Fewer true errors were made during backward chaining than during whole task training ($t[23] = -5.0$, $p < .05$).

Figure 16 shows the effect of training procedure on total responses during acquisition. The overall comparison among trairing procedures on number of responses made during acquisition was significant ($F[2, 46] = 9.9$, $p < .001$). Subjects made fewer responses during forward chaining than during whole task training ($t[23] = -4.0$, $p < .05$). Subjects also made fewer responses during backward chaining than during whole task training ($t[23] = -2.2$, $p < .05$).

Figure 17 shows the mean learning time for each training procedure. Training procedure had a significant effect on learning time ($F[2, 46] = 8.3$, $p < .01$). Subjects learned the sequences more quickly by forward chaining than by backward chaining ($t[23] = -3.5$, $p < .05$) and whole task training ($t[23] = -3.9$, $p < .05$).

Discussion

The results of Experiment 2 provide support for the hypothesis that subjects can perform the sequence in the absence of antecedent external stimuli more accurately and quickly following forward chaining than following backward chaining or whole task training. Some evidence was found that backward chaining leads to performance superior to that of whole task training on Test 1. Subjects performed significantly longer series of correct responses on Test 1 after backward chaining than after whole task training. Forward chaining led to higher percent correct scores, greater numbers of consecutive correct responses, fewer errors, and faster completion times on Test 1 than backward chaining and whole task training.

Test 1 results from this experiment further support the idea that response produced stimulation plays an important role in forward chaining and that such stimulation likely becomes the basis of discriminative stimuli. Such stimulation apparently does not play as important a controlling role in backward chaining or whole task training.

The hypothesis that subjects could order the antecedent external stimuli more accurately following forward chaining also received some support. Higher percent correct scores were obtained on Test 2 following forward chaining than following backward chaining. More correct pairs occurred following forward chaining than following backward chaining whereas more reverse

pairs were produced following backward chaining than following the other two procedures of training. No support was found for the hypothesis that subjects would order the symbols more accurately following backward chaining than following whole task training. Rather, the findings suggested that subjects may order the antecedent external stimuli more accurately following whole task training than following backward chaining.

Test 2 results may reflect the order in which components are added to the sequence in the various training procedures. Forward chaining, in which behaviors are added to the sequence at the end of the subsequences and in order, led to the highest percent correct scores and the identification of the most correct pairs. Backward chaining led to the identification of more reverse pairs than did forward chaining, perhaps reflecting the fact that behaviors are added to the beginning of the subsequences and in order. Because behaviors are added at the beginning of the sequence during backward chaining, a subject sees symbol 10 for the first time before seeing symbol 9 for the first time. Seeing pairs of symbols in this reversed order is unique to backward chaining and may lead to the identification of reverse pairs on Test 2.

The superior performance on Tests 1 and 2 following forward chaining is interesting because it suggests that although the screen symbols are not necessary to guide sequence performance following training (Test 1 results), the symbols do remain

salient in that subjects are relatively good at ordering them following training.

Experiment 3

Experiment 3 examined performance of a third type of behavior that might be learned during chaining: the performance of the correct response following the presentation of an antecedent external stimulus, when the stimuli were presented out of sequence order. This behavior was examined for a number of reasons. First, it seemed likely that this behavior might be learned during the chaining procedures. Second, the performance of this behavior provided a measure of the controlling role of the antecedent external stimuli at the end of training. The correct performance of a response following the presentation of an antecedent external stimulus, when the stimulus was presented out of sequence order, would suggest that the antecedent external stimulus had become an effective discriminative stimulus for the response. Third, this behavior provided a measure of the flexibility of subjects' performance of component responses following training. It might be helpful to learn some behaviors (e.g., producing a musical note in response to sheet music) by a procedure that leads to ease in performing its component responses in other orders than those involved in the initial training.

No specific a priori hypotheses were made regarding differences among the three chaining procedures on the performance of the third type of post-acquisition behavior. However, the experiment was motivated by the general hypothesis

that the chaining procedures would lead to differences in performance.

Method

Subjects

Subjects were 24 volunteers (4 males, 20 females) from undergraduate psychology classes. The mean age of subjects was 22.5 years (*SD* = 3.6, range 19 - 32).

Procedure

The procedure followed was as for Experiment 2 except that a different type of post-acquisition test was administered following each training procedure. The post-acquisition test was designed to assess whether subjects could perform the response that corresponded to the antecedent external stimuli, when these stimuli were presented out of sequence order. Test 3 was preceded by the message on the monitor screen "Please press the button that goes with each of the symbols that are about to be presented. If you press the wrong button, the computer will wait for you to get it right". Following this message, the ten symbols used in the sequence that had just been learned were presented in two different orders. The two orders of ten symbols in the test were arranged such that no two symbols that had followed one another in the original sequence, followed one another in the test sequence. When a subject pressed the wrong button during Test 3, the computer made the same "zit" sound that it had during acquisition. When a subject pressed the correct

button during Test 3, the screen symbol changed to the next symbol of the test.

Variables and Design

The design consisted of one within subjects factor with 3 levels (Chaining Procedure). Two sets of dependent variables were examined: test performance variables and acquisition variables.

The variables related to Test 3 were percent correct, completion time, and location of errors. An error on Test 3 was a button press that did not correspond to the on screen symbol. The location of errors on Test 3 was calculated in a similar manner as for location of errors on the sequence test in Experiments 1 and 2. The number of errors made prior to correctly pressing the keys that had corresponded to each of the symbols were counted. The number of errors made prior to correctly pressing the keys that had corresponded to symbols 1 to 5 were then summed to provide a measure of the number of errors made prior to correct responses to the symbols from the beginning of the sequence. Similarly, the number of errors made prior to correctly pressing the keys that had corresponded to symbols 6 to 10 were summed to provide a measure of the number of errors made prior to correct responses to the symbols from the end of the sequence.

Data Analyses

The data analytic strategy was the same as for Experiment 2.

Results

Exact probabilities for all inferential statistics are reported in Appendix F. Only significant results are reported. Any comparisons not reported did not reach significance.

Test 3: Performance of Responses Following the Symbols

Descriptive statistics for Test 3 variables are presented in Table 9. The repeated measures MANOVA for Test 3 variables was significant ($F[4, 20] = 4.4$, $p < .05$).

Percent correct. Figure 18 shows the mean percent correct on Test 3 for each of the training procedures. The univariate comparison among training procedures for percent correct on Test 3 was significant ($F[2, 46] = 4.8$, $p < .05$). Follow-up dependent pairs t-tests revealed that subjects obtained significantly higher percent correct scores on Test 3 test following backward chaining ($M = 85.9\%$) than following forward chaining ($M = 66.2\%$; $t[23] = -3.5$, $p < .05$).

Completion time. Figure 19 shows the mean completion time for Test 3 for each of the training procedures. The overall comparison among training procedures for time to complete Test 3 approached significance ($F[2, 46] = 2.9$, $p = .06$). Follow-up dependent pairs t-tests revealed that subjects completed Test 3 more quickly following backward chaining ($M = 47.7$ s) than following forward chaining ($M = 73.3$ s; $t[23] = 2.6$, $p < .05$).

Location of errors. Descriptive statistics for location of Test 3 errors are presented in Table 10. Figure 20 shows the

mean number of errors made on Test 3 for the three training procedures. A 2 (Location of Errors) X 3 (Training Procedure) repeated measures ANOVA revealed a significant main effect for training procedure ($F[2, 22] = 7.2$, $p < .01$). Follow-up dependent pairs t-tests indicated that fewer errors were made on Test 3 following backward chaining ($M = 4.5$) than following forward chaining ($M = 17.9$; $t[23] = 3.3$, $p < .05$). No evidence was found that errors were made in different locations depending upon training procedure. Following forward chaining, the mean errors made on symbols from the beginning and end of the sequence were 6.0 and 11.8, respectively. Following backward chaining, the mean errors made on symbols from the beginning and end of the sequence were 2.4 and 2.0, respectively. Following whole task training, the mean errors made on symbols from the beginning and end of the sequence were 7.6 and 6.8, respectively.

Acquisition Variables

Descriptive statistics for the acquisition variables are presented in Table 11. The repeated measures MANOVA for acquisition variables was significant ($F[8, 16] = 13.2$, $p < .001$).

Figure 21 shows the significant effect of training procedure on total errors during acquisition ($F[2, 46] = 18.2$, $p < .001$). Forward chaining and backward chaining led to fewer total errors than did whole task training ($t[23] = -6.7$, $p < .05$ and $t[23] = -3.6$, $p < .05$, respectively).

Figure 22 shows the mean true errors for each training procedure. Training procedure had a significant effect on true errors during acquisition ($F[2, 46] = 16.4$, $p < .001$). Subjects made fewer true errors during forward and backward chaining than during whole task training ($t[23] = -6.4$, $p < .05$ and $t[23] = -3.3$, $p < .05$, respectively).

Figure 23 shows the mean total responses for each training procedure. The overall comparison among training procedures on total responses during acquisition reached univariate significance ($F[2, 46] = 9.3$, $p < .001$). Subjects made fewer responses during forward chaining than during backward chaining or whole task training ($t[23] = -2.4$, $p < .05$ and $t[23] = -4.7$, $p < .05$, respectively).

Figure 24 illustrates the effect of training procedure on time to learn a sequence ($F[2, 46] = 6.9$, $p < .01$). Subjects learned sequences more quickly by forward chaining than by backward chaining or whole task training ($t[23] = -3.1$, $p < .05$ and $t[23] = -4.3$, $p < .05$, respectively).

Discussion

The results of Experiment 3 further supported the overall hypothesis that the three procedures should result in different post-acquisition performance. Specifically, when asked to perform the component behaviors of a sequence in some other order (with the aid of the screen stimuli), backward chaining led to fewer errors than did forward chaining and whole task training.

Following backward chaining, then, the screen symbols had a stronger controlling role for responses than they did following forward chaining or whole task training. The poorer performance of the forward chaining group on Test 3 fits with Test 1 results from Experiment 2 which suggested that, by the end of forward chaining, the screen stimuli may no longer play an important role in controlling responding.

It is possible that backward chaining leads to fewer errors on Test 3 because of the order in which behaviors are added to subsequences. Because the sequence is built by adding behaviors at the beginning of each subsequence, subjects may be forced to look at the symbols in order to prevent making errors at the beginning of each new subsequence. During backward chaining, then, when responses are added to the subsequence, the only reliable discriminative stimulus for the response is the screen stimulus. By contrast, because responses are added to the end of the subsequence during forward chaining, there are three reliable discriminative stimuli or cues for the response: the previous response, the response produced stimulation of the previous response, and the screen stimulus. During forward chaining, stimulus control may shift to the response produced stimulation leading to poor performance on Test 3 and good performance on Test 1. During whole task training, subjects may rely on some combination of the screen symbols, the responses, and response

produced stimulation as discriminative stimuli for subsequent responses.

The results of this experiment are interesting for at least two reasons. One, backward chaining, for the first time on any measure, appears superior to the other two methods. Second, the performance of the component behaviors in some order other than that in which they were learned has implications for training subjects to perform sequences for which, subsequently, some degree of flexibility is desired.

Experiment 4

Experiment 4 continued the investigation of post-acquisition performance using Test 3. Three changes were made to the procedure followed for Experiment 3 in order to facilitate this investigation. The sequence length was increased from 10 responses to 12 responses, the number of subjects was increased from 24 to 30, and whole task training was replaced with random chaining.

The rationale for increasing the sequence length was to increase the difficulty of the task and therefore increase the number of errors and exaggerate patterns of errors. The aim of increasing the number of errors was especially important for the purpose of examining the location of errors on Test 3. The number of subjects was increased in order to improve the power to detect differences among chaining procedures.

As the focus of the research narrowed, it seemed appropriate to replace whole task training with a random chaining procedure. Random chaining has the advantage of being similar to backward and forward chaining as well as being different from them in very specific, identifiable ways. This allows for the testing of more specific hypotheses and the development of a more sound understanding of what is learned when a subject is taught a sequence of behaviors.

Random chaining involves teaching one randomly chosen response, then teaching that response plus a second randomly

chosen response and so on. During training, the randomly chosen responses are put into the sequence in the order in which they appear in the final sequence. For example, the subject might be taught response 4; then response 1 followed by response 4; then response 1 followed by response 4 followed by response 9 and so on.

The superior performance on Test 3 following backward chaining in Experiment 3 was tentatively explained as a function of role the screen symbols play as discriminative stimuli in the various chaining procedures. When responses are added to the subsequence during forward chaining, subjects have the advantage of three reliable discriminative stimuli or cues for the response: the previous response, the response produced stimulation of the previous response, and the screen stimulus. During backward chaining, when a response is added to the subsequence, the only reliable discriminative stimulus for it is the screen stimulus. Later in backward chaining, the previous response and its response produced stimulation become reliable discriminative stimuli. When responses are added to the subsequence during random chaining, the only reliable discriminative stimulus for the response is the screen stimulus. Previous responses and their response produced stimulation do not become reliable discriminative stimuli for subsequent responses in random chaining until late in training.

Based on this analysis of the chaining procedures, it was hypothesized that random chaining would lead to the best performance on the Test 3. Forward chaining was expected to lead to the poorest performance on Test 3 and backward chaining was expected to lead to intermediate performance.

Method

Subjects

Subjects were 30 undergraduate student volunteers (28 females, 2 males). The mean age of subjects was 27.5 (*SD* = 8.6, range 19 to 45).

Apparatus

The apparatus was the same as for Experiments 1-3.

Procedure

The procedure was the similar to that followed for Experiment 3 with the following exceptions: (a) sequences were 12 responses in length rather than 10, and, (b) random chaining replaced whole task training. The two additional keys that were covered with tape to allow for a 12 response sequence were O and L.

Variables and Design

The variables and design were the same as for Experiment 3. Because the sequence length for Experiment 4 was 12 responses, the location of errors on Test 3 was calculated so as to allow for examination of errors made on symbols from three sequence locations: the beginning (symbols 1 to 4), the middle (symbols 5

to 8), and the end (symbols 9 to 12) of the sequence. For example, the number of errors made prior to correctly pressing the keys that had corresponded to each symbol were counted. Then the number of errors made prior to correctly pressing the keys that had corresponded to symbols 1 through 4 were summed to provide a measure of the number of errors made prior correct responses for the symbols from the beginning of the sequence.

Data Analyses

The same data analytic strategy was used for Experiment 4 as was used for Experiment 3.

Results

Exact probabilities for all inferential statistics are reported in Appendix G. Only significant results are reported. Any comparisons not reported did not reach significance.

Test 3: Performance of Responses Following the Symbols

Descriptive statistics for Test 3 variables are presented in Table 12. A repeated measures MANOVA for Test 3 variables approached significance ($F[4, 26] = 2.3$, $p = .09$). Univariate comparisons among chaining procedures on both Test 3 variables were significant but follow-up tests should be interpreted with caution given the significance level of the multivariate test.

Percent correct. Figure 25 shows the mean percent correct on Test 3 for the three chaining procedures. The effect of chaining procedure on percent correct on Test 3 was significant at the univariate level ($F[2, 58] = 4.7$, $p < .05$). Follow-up

dependent pairs t-tests revealed that forward chaining (M = 76.1%) led to lower scores on Test 3 than did random chaining (M = 89.5%; t[29] = -2.7, p < .05).

Completion time. Figure 26 shows the mean completion time for Test 3 for the three chaining procedures. The univariate test revealed that chaining procedure had a significant effect on time to complete Test 3 (F[2, 58] = 5.7, p < .01). Follow-up dependent pairs t-tests revealed that subjects took longer to complete Test 3 following forward chaining (M = 75.8 s) than following random chaining (M = 46.2 s; t[29] = 2.9, p < .05).

Location of errors. Descriptive statistic for location of errors on Test 3 are presented in Table 13. A 3 (Location: beginning, middle, end) X 3 (Chaining Procedure) repeated measures ANOVA revealed significant main effects for chaining (F[2, 28] = 3.8, p < .05) and for location (F[2, 28] = 11.0, p < .001). Figure 27 shows the mean number of errors made on Test 3 as a function of chaining procedure. Figure 28 shows the mean errors made on symbols from each sequence location.

Dependent pairs t-tests revealed that more errors were made on Test 3 following forward chaining (M = 16.9) than following random chaining (M = 3.2; t[29] = 2.4, p < .05). More errors were made on Test 3 following backward chaining (M = 7.7) than following random chaining (t[29] = 2.4, p < .05).

Dependent pairs t-tests revealed that more errors were made on symbols from the beginning of the sequence (M = 13.0) than on

symbols from the end of the sequence (M = 5.6; t[29] = 4.4, p < .05). The comparison between mean errors on symbols from the middle (M = 9.2) and end of the sequence was also significant (t[29] = 2.6, p < .05), with more errors made on symbols from the middle of the sequence than on symbols from the end of the sequence.

Acquisition Variables

Descriptive statistics for the acquisition variables are presented in Table 14. A repeated measures MANOVA for the acquisition variables was significant (F[8, 22] = 4.8, p < .01).

Univariate comparisons were significant for only 2 of the 4 acquisition variables: true errors and learning time. The univariate comparisons for total errors and total responses during acquisition did not reach significance.

Figure 29 shows the effect of chaining procedure on true errors (F[2, 58] = 3.8, p < .05). Fewer true errors were made following forward chaining than following random chaining (t[29] = -3.3, p < .05).

Figure 30 shows the effect of chaining procedure on learning time (F[2, 58] = 3.2, p < .05). Subjects learned sequences more quickly by forward chaining than by random chaining (t[29] = -2.7, p < .05).

Discussion

The results of Experiment 4 supported the hypothesis that random chaining would lead to superior post-acquisition

performance on Test 3. Following random chaining, subjects obtained higher percent correct scores on Test 3 and completed the test faster than following forward or backward chaining.

The superior performance on Test 3 following random chaining supports the notion that, during random chaining, the external antecedent stimuli become more reliable discriminative stimuli than is the case during forward chaining. The hypothesis that backward chaining would lead to intermediate performance on Test 3 (as a result of the intermediate strength of the screen symbols as discriminative stimuli) was not supported by statistically significant differences. However, the rank order of the means and the medians for the three chaining procedures on completion time and of the means on percent correct are consistent with this hypothesis.

The results for location of errors on Test 3 may reflect the high number of errors made following forward chaining. During forward chaining, the stimulus control for responses likely shifts from the screen symbols to response produced stimulation for responses from the beginning of the sequence first, then for responses from the middle of the sequence, and finally for responses from the end of the sequence. The responses from the beginning of the sequence are learned first during forward chaining and therefore there is more opportunity for stimulus control to shift to response produced stimulation from the screen symbols. The high number of errors made on symbols from the

beginning of the sequence, the intermediate number of errors made on symbols from the middle of the sequence and the low number of errors made on symbols from the end of the sequence may reflect the fact that it was the forward chaining group that made the bulk of the errors made on Test 3.

Experiment 5

Experiment 5 used Test 3 to further investigate the performance of correct responses following the presentation of antecedent external stimuli. The purpose of Experiment 5 was to: (a) replicate the results of Experiment 4; and, (b) test hypotheses regarding the impact of proficiency criteria on Test 3 performance.

Subjects were exposed to two proficiency criteria (1 and 4) in Experiment 5. It was expected that increasing the proficiency criterion from 1 to 4 would result in subjects looking at the external stimuli less and would therefore affect performance on Test 3. Specifically, it was hypothesized that the repetitions of each subsequence and the complete sequence required at a proficiency criterion of 4 would result in poorer performance on Test 3 than at a proficiency criterion of 1. The more repetition required, the more likely it is that stimuli other than the screen symbols may become reliable discriminative stimuli. For example, response produced stimulation (e.g., feeling and seeing the performance of the response) may become reliable discriminative stimuli. Stimulus control for the responses may shift to stimuli other than the screen symbols if the other stimuli become reliable discriminative stimuli.

Method

Subjects

Subjects were 60 undergraduate student volunteers (45 females, 15 males). The mean age of subjects was 22.5 (*SD* = 7.1, range 18 to 60).

Apparatus

The apparatus was the same as for Experiments 1 to 4.

Procedure

Each subject was randomly assigned to one of three groups: forward chaining, backward chaining, or random chaining. Subjects learned two 12 response sequences to criterion. One sequence was performed to a proficiency criterion of 1 (PC 1), the other to a proficiency criterion of 4 (PC 4). The order of proficiency criteria was counterbalanced. Following the performance of each sequence to criterion, subjects completed Test 3.

Variables and Design

The variables were the same as for Experiment 3. The design was 3 between (Chaining Procedure) X 2 within (Proficiency Criterion) X 2 between (Order of Proficiency Criteria). For ease of presentation, where a proficiency criterion of 1 followed a proficiency criterion of 4, the order will be referred to as *order 1-4*. Where a proficiency criterion of 4 followed a proficiency criterion of 1, the order will be referred to as *order 4-1*.

Results

Complete inferential statistics for Experiment 5, including exact *p* values, are presented in Appendix H. In contrast to Experiments 1 to 4, nonsignificant comparisons are reported to facilitate the description of significant interactions.

Test 3: Performance of Responses Following the Symbols

Descriptive statistics for Test 3 variables are presented in Tables 15 and 16. A 3 (Chaining Procedure) X 2 (Proficiency Criterion) X 2 (Order) MANOVA performed on percent correct and completion time revealed a significant Chaining X Proficiency Criterion X Order interaction ($F[4, 108] = 2.5$, $p < .05$).

Percent correct. Figures 31 and 32 show the three way interaction for percent correct on Test 3. The univariate test for percent correct on Test 3 reached univariate significance for the three way interaction ($F[2, 54] = 3.6$, $p < .05$). Follow-up dependent pairs *t*-tests between mean percent correct scores on Test 3, following PC 1 and PC 4, for each order, revealed a similar pattern for all three chaining groups. The interaction is described in three steps. First, an overall description of the interaction is provided. Second, the results of the follow-up *t*-tests between PC 1 and PC 4 for each of the chaining groups at each order are presented. Third, the results of the post hoc Student Newman Keuls comparisons of the chaining groups for each of the orders are presented.

Overall, the interaction revealed the following pattern for all three chaining groups. For order 1-4, each chaining group obtained significantly higher percent correct scores on Test 3 after learning a sequence to PC 4 than after learning a sequence to PC 1. However, for order 4-1, there were no significant differences between mean percent correct scores on Test 3 following learning a sequence to PC 1 and to PC 4. When subjects had already been exposed to Test 3 following PC 1, then, PC 4 led to higher percent correct scores on Test 3 for all three chaining groups. When subjects were exposed to Test 3 following PC 4, percent correct scores were not higher the second time Test 3 was completed (following PC 1).

The forward chaining group obtained significantly higher percent correct scores on Test 3 following a proficiency criterion of 4 ($\underline{M}$ = 97.6%) than following PC 1 ($\underline{M}$ = 62.9%), only when the proficiency criterion of 4 was presented second (order 1-4; $\underline{t}[9] = -4.7$, $\underline{p} < .05$). For order 4-1, the difference between percent correct scores on Test 3 following forward chaining to a proficiency criterion of 4 ($\underline{M}$ = 51.0%) and PC 1 ($\underline{M}$ = 65.9%) was not significant.

For the backward chaining group, order 1-4, neither paired comparison was significant but the pattern of means was similar. When the proficiency criterion of 4 was presented first (order 4-1), there was no significant difference between the mean percent correct score on Test 3 after learning a sequence to a

proficiency criterion of 4 (M = 82.0%) and the mean score on Test 3 after learning a sequence to PC 1 (M = 70.3%).

For the random chaining group, percent correct scores on Test 3 were higher following a proficiency criterion of 4 (M = 98.5%) than following PC 1 (M = 91.8%) for order 1-4 only ($t[9]$ = -3.4, $p < .05$). The difference between scores on Test 3 at the proficiency criteria of 1 and 4 was not significant when the proficiency criterion 4 was presented first. The means for the random chaining group for order 4-1 were 85.3% at PC 4 and 88.2% at PC 1.

Depending upon the order in which proficiency criteria were presented, the pattern of mean percent correct scores for the three chaining groups differed at each proficiency criteria. For order 1-4, there were no significant differences between chaining groups on mean percent correct scores on Test 3 following PC 4 (the second test). At PC 1, however, the overall comparison among groups was significant ($F[2, 27] = 6.8$, $p < .01$). Student Newman Keuls comparisons with alpha set at .05 revealed that, for order 1-4, PC 1, the random chaining group (M = 91.8%) obtained significantly higher percent correct scores than the forward (M = 62.9%) and backward chaining groups (M = 65.1%; $p < .05$).

By contrast, for order 4-1, there were no significant differences between chaining groups on mean percent correct scores on Test 3 following PC 1 (the second test). At PC 4, however, the overall comparison among chaining groups was

significant (F[2, 27] = 5.0, p < .05). For order 4-1, PC 4, Student Newman Keuls comparisons, with alpha set at .05, indicated that the forward chaining group obtained significantly lower percent correct scores on Test 3 (M = 51.0%) than did the backward (M = 82.0%) and random chaining groups (M = 85.3%). Taken together, the results of comparisons among chaining groups at each order and proficiency criterion appear to reveal a practice effect. On the second test, regardless of proficiency criterion, there were no significant differences among chaining groups in terms of mean percent correct.

Completion time. Figures 33 and 34 show the three way interaction for Test 3 completion time (univariate F[2, 54] = 4.8, p < 05). This interaction is described in three steps. First, the overall pattern of the interaction is described. Second, the follow-up t-tests between completion time at PC 1 and PC 4 for each chaining group, at each order, are presented. Third, the post hoc Student Newman Keuls comparisons of the chaining groups for each order are presented.

Overall, this interaction revealed a similar patterns for the forward and random chaining groups and a different pattern for the backward chaining groups. For both the forward and random chaining groups, order 1-4, completion times were longer following PC 1 than following PC 4. For the forward and random chaining groups, order 4-1, there was no significant difference between completion times following PC 1 and PC 4. By contrast,

there were no significant differences between completion times following PC 1 and PC 4, at either order, for the backward chaining group.

Follow-up dependent pairs *t*-tests revealed that, for forward and backward chaining, order 1-4, there were no significant differences between completion times for Test 3 at the two proficiency criteria.

For random chaining, order 1-4, a significantly longer mean completion time was evident following PC 1 (M = 45.2 s) than following PC 4 (M = 35.1 s; $t[9] = 2.9$, $p < .05$). For random chaining, order 4-1, this difference was not significant. For random chaining, order 4-1, the mean completion times following PC 1 and PC 4 were 44.9 s and 57.0 s, respectively.

For order 1-4, the overall comparison among chaining groups for completion times following each proficiency criteria was not significant. The overall comparison did approach significance ($F[2, 27] = 3.2$, $p = .06$) for PC 1 where the means for forward, backward, and random chaining were 140.5 s, 69.8 s, and 45.2 s, respectively. However, the paired comparisons among the three procedures did not reach significance. For order 4-1 the overall comparison among chaining groups for mean Test 3 completion times following PC 4 was significant ($F[2, 27] = 4.7$, $p < .05$). Student Newman Keuls comparisons with alpha set at .05 indicated that the forward chaining group took longer to complete Test 3 (M = 144.6 s) than the backward (M = 71.4 s) and random chaining

groups (M = 57.0 s; $p < .05$). The comparison among chaining groups for order 4-1, proficiency criterion 1, was not significant.

Location of errors. The descriptive statistics for location of errors on Test 3 are presented in Tables 17 and 18. The MANOVA for location of errors on Test 3 revealed two significant interactions, a Chaining X Order X Location interaction ($F[4, 108] = 2.8$, $p < .05$) and a Location X Proficiency Criterion Interaction ($F[2, 53] = 5.0$, $p < .05$).

Figures 35 and 36 show the three way interaction for location of errors on Test 3. Note that the means reported in the following description of the interaction are not tabled but are the mean of the errors made at each proficiency criteria. The means for errors at each location at each proficiency criteria are reported in Tables 17 and 18.

The comparisons among errors made on symbols from different sequence locations by the forward chaining group did not reach significance for either order. The paired comparisons among errors made on symbols from different sequence locations following backward chaining, order 1-4, were not significant. The backward chaining group, order 4-1, made significantly more errors on symbols from the beginning of the sequence (M = 8.0) than on symbols from the end of the sequence (M = 4.3; $t[9] = 3.0$, $p < .05$).

For the random chaining group, order 1-4, errors were evenly distributed. Mean errors made on symbols from the beginning, middle, and end of the sequence were 0.5, 0.6, and 0.3, respectively. The same was true for the random chaining group, order 4-1. The means for errors made on symbols from the beginning, middle, and end of the sequence were 4.1, 0.7, and 1.2, respectively.

At order 1-4, the overall univariate comparison of chaining groups for errors made on symbols from the beginning of the sequence was significant ($F[2, 27] = 3.9$, $p < .05$). Student Newman Keuls comparisons, with alpha set at .05, revealed that the backward chaining group made significantly more errors on symbols from the beginning of the sequence ($M = 6.9$) than the random chaining group ($M = 0.5$; $p < .05$). The mean number of errors made on symbols from the beginning of the sequence at order 1-4 by the forward chaining group was 3.9. At order 1-4, the overall univariate comparisons of the chaining groups for errors made on symbols from the middle and end of the sequence did not reach significance.

At order 4-1, the overall univariate comparison of chaining groups for errors made on symbols from the beginning of the sequence did not reach significance. The same comparison for errors made on symbols from the middle of the sequence was significant ($F[2, 27] = 4.5$, $p < .05$). Student Newman Keuls comparisons, with alpha set at .05, revealed that the forward

chaining group made more errors on symbols from the middle of the sequence (M = 12.7) than did the random chaining group (M = 0.7; $p < .05$). Paired comparisons involving the backward chaining group did not reach significance (M = 5.3). The comparison among chaining groups for errors made on symbols from the end of the sequence was significant ($F[2, 27] = 7.6$, $p < .01$). Student Newman Keuls comparisons, with alpha set at .05, revealed that the forward chaining group made significantly more errors (M = 14.8) on symbols from the end of the sequence than the backward (M = 4.3) or random (M = 1.2) chaining groups. The comparison between the backward and random chaining groups was not significant.

Figure 37 shows the Location X Proficiency Criterion interaction. Following PC 1, significantly more errors were made on symbols from the beginning of the sequence (M = 7.7) than on symbols from the middle (M = 4.3) or end of the sequence (M = 4.9; $t[59] = 2.9$, $p < .05$ and $t[59] = 2.0$, $p = .05$). Following PC 4, errors were evenly distributed on Test 3. The mean number of errors made on symbols from the beginning, middle, and end of the sequence following PC 4 were 3.8, 3.5, and 4.4, respectively.

Comparisons between the mean number of errors made at each proficiency criteria on symbols from the beginning, middle, and end of the sequence did not reach significance.

Acquisition Variables

Descriptive statistics for the acquisition variables are presented in Tables 19 and 20. A 3 (Chaining) X 2 (Proficiency Criterion) X 2 (Order) MANOVA performed on the acquisition variables (total errors, true errors, total responses, and learning time) revealed a significant main effect for Chaining ($F[8, 104] = 2.2$, $p < .05$) and a significant Proficiency Criterion X Order interaction ($F[4, 51] = 6.3$, $p < .001$). The Chaining X Proficiency Criterion X Order interaction also approached significance ($F[8, 104] = 1.8$, $p = .08$). The univariate tests for the acquisition variables did not reach significance for either the Chaining or the Chaining X PC X Order interaction. The univariate tests for each of the acquisition variables were significant for the Proficiency Criterion X Order interaction.

Figure 38 shows the Proficiency Criterion X Order (PC X O) interaction ($F[1, 54] = 5.2$, $p < .05$) for total errors. Follow-up dependent pairs t-tests indicated that, for order 1-4, significantly more total errors were made following PC 1 than following PC 4 ($t[29] = 3.0$, $p < .05$). This difference was not significant for order 4-1.

Figure 39 shows the PC X O interaction for true errors during acquisition revealed a significant Proficiency Criterion X Order interaction ($F[1, 54] = 5.2$, $p < .05$). Follow-up dependent pairs t-tests indicated that, for order 1-4, significantly more

total errors were made following PC 1 than following PC 4 (t[29] = 2.8, p < .05). This difference was not significant for order 4-1.

Figure 40 shows the PC X O interaction for total responses during acquisition (F[1, 54] = 5.6, p < .05). Follow-up dependent pairs t-tests revealed that, for orders 1 and 2, subjects made significantly more responses at PC 4 than at PC 1 (t[29] = -6.4, p < .05 and t[29] = -12.9, p < .05, respectively).

Figure 41 shows the PC X O interaction for learning time (F[1, 54] = 16.8, p < .001). For both order 1-4 and order 4-1, subjects took longer to learn a sequence at PC 4 than at PC 1 (t[29] = -6.8, p < .05 and t[29] = -13.8, p < .05, respectively).

Discussion

The results of Experiment 5 did not provide support for the hypothesis that a proficiency criteria of 4 would lead to poorer performance on Test 3 than would a proficiency criteria of 1. Higher percent correct scores were obtained by all three chaining groups at PC 4 than at PC 1, at order 1-4. When the chaining groups were exposed to order 4-1, there were no differences between percent correct scores at PC 1 and PC 4. The Test 3 completion time results are consistent with the percent correct results for the forward and random chaining groups. Following both forward and random chaining at order 1-4, subjects took longer to complete Test 3 following PC 1 than following PC 4. This pattern of results suggests that the repetition required at

PC 4 improved Test 3 performance, providing the subject had completed Test 3 once already.

For both orders of PC presentation, percent correct scores on the first test supported Experiment 4 findings regarding Test 3 performance following the three types of chaining. At PC 1, order 1-4, the random chaining group obtained significantly higher percent correct scores than the forward and backward chaining groups. At PC 4, order 4-1, the forward chaining group obtained significantly lower percent correct scores and took longer to complete the test than the backward and random chaining groups.

Comparisons of location of errors for the three chaining groups reflected the tendency of the backward chaining group to make errors on symbols from the beginning of the sequence (order 1-4) and of the forward chaining groups to make errors on symbols from the end of the sequence (order 4-1). For both chaining procedures, the symbols on which the most errors were the last symbols learned, and therefore might not control responses as well as symbols that are seen more often.

Although the results were consistent with those of Experiments 3 and 4 with regards to Test 3 performance following chaining, they did not support the general hypothesis that an increase in proficiency criterion would lead to poorer Test 3 performance. A better methodology for future exploration of this hypothesis would be to make level of proficiency criterion a

between rather than a within subjects factor. This would eliminate the complication of practice effects.

Experiment 6

Experiment 6 examined three types of post-acquisition performance: (1) Whether subjects could accurately perform the sequence in the absence of the antecedent external stimuli used during training (Test 1); (2) Whether subjects could accurately order the antecedent external stimuli (Test 2); and, (3) Whether subjects could press the button that corresponded to the antecedent external stimuli that had been used during training when these stimuli were presented out of order (Test 3). The purposes of the experiment were to:

(1) Examine the performance on Test 1 following random chaining. Test 1 had not been presented following random chaining in the previous experiments.

(2) Examine the performance on Test 2 following random chaining. Test 2 had not been presented following random chaining in the previous experiments.

(3) Re-examine performance on the three tests following forward and backward chaining. Although all three tests had been presented following forward and backward chaining in previous experiments, the three tests had not been presented within a single experiment. In addition, Test 2 was redesigned for Experiment 6.

(4) Re-examine performance on Test 3 following random chaining.

With regards to Test 1, it was hypothesized that the forward chaining group would obtain the highest percent correct scores (as well as the best scores on other Test 1 variables) and that the random chaining group obtain the lowest percent correct scores on Test 1. The backward chaining group was expected to obtain intermediate scores. This hypothesis was based on the results of Experiment 1 and 2. The rationale for this hypothesis was based on the notion that during forward chaining, response produced stimulation should form more reliable discriminative stimuli than it will during backward chaining. Response produced stimulation should form the least reliable discriminative stimuli during random chaining.

With respect to performance on Test 2, it was hypothesized that the forward chaining group would obtain the highest percent correct scores, followed by the backward chaining group, followed by the random chaining group. Further, it was hypothesized that the forward chaining group would identify the most correct pairs and that the backward chaining group would identify the most reverse pairs. This hypothesis was based on the results of Experiments 1 and 2. Forward chaining gives subjects the most repetitive experience of seeing the screen stimuli in order. Backward chaining gives subjects a similar amount of experience but, because component behaviors are added to the beginning of the subsequence may lead to a reverse ordering of symbols. For

example, during backward chaining, subjects see symbol 12 for the first time before they see symbol 11 for the first time.

With respect to performance on Test 3, it was hypothesized that the random chaining group would obtain the highest percent correct scores, followed by the backward chaining group, followed by the forward chaining group. This hypothesis was based on the results of Experiments 3, 4, and 5. The theoretical rationale for this hypothesis is based on the relative reliability of response produced stimulation as discriminative stimuli during the three chaining procedures.

Method

Subjects

Subjects were 72 undergraduate student volunteers (43 females, 29 males). The mean age of subjects was 20.3 years (*SD* = 3.7; range 18 to 47).

Apparatus

The apparatus was the same as for Experiments 4-5.

Procedure

Subjects were randomly assigned to three groups: forward chaining, backward chaining, or random chaining. Each subject learned one 12 response sequence to a proficiency criterion of one. One third of subjects learned a sequence by each of three chaining methods (forward, backward, and random). Following the performance of a sequence to criterion, subjects completed Tests 1, 2, and 3. The presentation order of the three tests was

counterbalanced, with four subjects in each group receiving each possible order.

Test 1 was the same as those used in previous experiments. Test 2 was modified in order to address two problems with the original hand-drawn form of the tests used in Experiments 1 and 2. The original version of Test 2 presented subjects with drawings of the symbols used during training rather than the actual symbols, making it a less than perfect test of relationships between the symbols actually used. The instructions on the original version also resulted in some confusion for subjects following backward chaining, making the test an imperfect assessment of performance following backward chaining. This confusion was avoided in the new version by using clearer instructions (Appendix I).

At the beginning of Test 2, a message appeared on the monitor screen instructing subjects to turn over the instruction sheet that the experimenter had given them. The symbols that had been used during acquisition were then presented on the monitor screen in three rows of four symbols each. The instruction sheet (Appendix I) included a blank 3 X 4 grid. The sheet asked subjects to find the first symbol in the sequence they had just learned and place a "1" in the appropriate space on the grid then to find the second symbol and place a "2" in the appropriate space and so on.

Variables and Design

The design consisted of one between subjects factor with 3 levels (Chaining Procedure). Four sets of dependent variables were examined: Test 1 variables, Test 2 variables, Test 3 variables, and acquisition variables.

The variables were the same as those in preceding Experiments. For Tests 1 and 3, the location of errors was calculated for three sections of the sequence: beginning, middle, and end.

Data Analyses

The data analyses were similar to those of Experiment 5. Where multiple dependent variables were of interest, MANOVAs were performed to test for overall differences among the chaining groups. If multivariate significance was reached, those variables reaching univariate significance were followed-up with the Student Newman Keuls procedure for multiple paired comparisons.

Where within subjects variables were of interest (i.e., location of errors on Tests 1 and 3), repeated measures ANOVAs were utilized, followed by dependent pairs *t*-tests for paired comparisons.

Results

Exact probabilities for all inferential statistics are reported in Appendix J. Only significant comparisons are reported. Comparisons not reported did not reach significance.

Test 1: Performance of the Sequence in the Absence of the Symbols

Descriptive statistics for Test 1 variables are presented in Table 21. A MANOVA performed on Test 1 variables was significant ($F[10, 132] = 5.6$, $p < .001$). Each dependent variable reached univariate significance.

Percent correct. Figure 42 shows the mean percent correct scores for each of the chaining groups. Chaining procedure had a significant effect on percent correct ($F[2, 69] = 15.6$, $p < .001$). Follow-up Student Newman Keuls comparisons indicated that the forward chaining group obtained significantly higher percent correct scores ($M = 68.5$) than the backward ($M = 44.6$) and random chaining groups ($M = 25.4$; $p < .05$). The backward chaining group obtained significantly higher percent correct scores than the random chaining group ($p < .05$).

Consecutive correct responses. The number of correct responses on Test 1 was examined using three variables: (a) the number of correct responses made prior to the first error, (b) the longest series of correct responses, and (c) the sequence position of the first response in the longest series of correct responses.

Figure 43 shows the effect of chaining procedure on the number of correct responses prior to the first error on Test 1. The univariate comparison of the three chaining groups on this variable was significant ($F[2, 69] = 22.6$, $p < .001$). Follow-up Student Newman Keuls comparisons with alpha set at .05 revealed

that the forward chaining group made significantly more correct responses prior to the first error (M = 6.7) than either the backward (M = 2.0) or the random chaining group (M = 0.5; p < .05).

Figure 44 shows the mean number of responses in the longest series of correct responses on Test 1 for each chaining group. The overall univariate comparison among chaining procedures on this variable was significant (F[2, 69] = 21.9, p < .001). Follow-up Student Newman Keuls comparisons with alpha set at .05 revealed that the forward chaining group made longer series of correct responses (M = 7.9) than backward chaining group (M = 5.7) and the random chaining group (M = 2.5, p < .05). The backward chaining group made longer series of correct responses than the random chaining group (M = 2.5, p < .05).

Figure 45 shows the effect of chaining procedure on the sequence position of the first response in the longest series of correct responses. The univariate comparison of the three chaining groups revealed that the sequence location of the first response in the longest series of correct responses differed depending on the chaining procedure (F[2, 69] = 9.6, p < .001). Follow-up Student Newman Keuls comparisons with alpha set at .05 showed that the forward chaining group began their longest series of correct responses earlier in Test 1 (M = 2.4) than the backward (M = 4.5) and random chaining groups (M = 5.3, p < .05).

The random chaining group was significantly different from the backward chaining group ($p < .05$).

Completion time. Figure 46 shows the mean completion time for Test 1 for each chaining group. Chaining procedure had a significant effect on time to complete Test 1 ($F[2, 69] = 8.3$, $p < .001$). Follow-up Student Newman Keuls comparisons with alpha set at .05 revealed that the random chaining group took longer to complete Test 1 ($M = 173.5$ s) than either the forward ($M = 47.4$) or backward chaining groups ($M = 97.5$, $p < .05$).

Location of errors. Descriptive statistics for location of Test 1 errors are presented in Table 22. A repeated measures ANOVA, with number of errors on the first (responses 1-4), second (responses 5-8), and third (responses 9-12) sections of Test 1 as dependent variables revealed a significant Location X Chaining Procedure interaction ($F[4, 138] = 2.7$, $p < .05$). Figure 47 shows the mean errors made prior to correct responses at the beginning, middle, and end of Test 1 for each chaining procedure.

This interaction was decomposed into simple main effects. One-way ANOVAs and Student Newman Keuls comparisons revealed a number of differences between pairs of chaining procedures in terms of the location of errors on Test 1. The alpha level was set at .05 for the Student Newman Keuls comparisons.

On the first third of the test, the forward chaining group made the fewest errors ($M = 1.2$), the random chaining group made the most errors ($M = 25.5$), and the backward chaining group made

an intermediate number of errors (M = 15.6). Each of the paired comparisons reached significance ($p < .05$).

In the middle section of the test, both the forward (M = 3.0) and backward chaining groups (M = 8.8) made fewer errors than the random chaining group (M = 17.0, $p < .05$). A similar pattern emerged for the last third of the test. Both the forward (M = 5.5) and the backward chaining groups (M = 8.6) made fewer errors than the random chaining group (M = 16.2, $p < .05$).

No differences were evident among errors made on responses from the three sequence locations by the forward and backward chaining groups. The random chaining group made significantly more errors on the first four responses of Test 1 (M = 25.5) than on the middle (M = 17.0; $t[23] = 3.6$, $p < .05$) and final sections of the test (M = 16.2; $t[23] = 2.6$, $p < .05$).

Test 2: Ordering the Symbols

Descriptive statistics for Test 2 variables can be found in Table 23. The MANOVA performed on the three Test 2 variables, percent correct, correct pairs, and reverse pairs, was significant ($F[6, 136] = 18.7$, $p < .001$). Univariate comparisons of the three chaining groups reached significance for each of the three variables.

Percent correct. Figure 48 shows the mean percent correct on Test 2 for each chaining procedure. Chaining procedure had a significant effect on performance on Test 2, as assessed by percent correct scores ($F[2, 69] = 23.0$, $p < .001$). Answers on

Test 2 were considered correct if a subject correctly identified the position of a symbol in the original sequence. Follow-up Student Newman Keuls comparisons with alpha set at .05 revealed that the forward chaining group obtained significantly higher percent correct scores (M = 51.7) than the backward (M = 10.1) and random chaining groups (M = 14.2, p < .05).

Pairs. Figure 49 shows the effect of chaining procedure on correct pairs. The overall comparison among chaining groups on correct pairs was also significant (F[2, 69] = 8.4, p < .001). Follow-up Student Newman Keuls comparisons with alpha set at .05 indicated that the forward chaining group identified more correct pairs (M = 4.6) than the backward (M = 1.7) and random chaining groups (M = 1.8, p < .05).

Figure 50 shows the effect of chaining procedure on reverse pairs. The overall comparison among chaining groups on reverse pairs was significant (F[2, 69] = 32.5, p < .001). Student Newman Keuls comparisons with alpha set at .05 revealed that the backward chaining group identified more reverse pairs (M = 4.1) than the forward (M = 0.9) and random chaining groups (M = 0.5, p < .05).

Test 3: Performing Responses Following the Symbols

Descriptive statistics for Test 3 variables are presented in Table 24. The MANOVA performed on percent correct and completion time for Test 3 was significant (F[4, 138] = 10.5, p < .001).

Univariate comparisons of the three chaining groups were significant for both dependent variables.

Percent correct. Figure 51 shows the mean percent correct on Test 3 for the three chaining groups. The overall univariate comparison among chaining groups was significant ($F[2, 69] = 15.9$, $p < .001$). Follow-up Student Newman Keuls comparisons with alpha set at .05 showed that the random chaining group obtained significantly higher percent correct scores on Test 3 ($M = 85.6\%$) than the forward ($M = 52.9\%$) and backward chaining groups ($M = 58.4\%$, $p < .05$).

Completion time. Figure 52 shows the completion time on Test 3 for each chaining procedure. Chaining procedure had a significant effect on time to complete Test 3 ($F[2, 69] = 12.5$, $p < .001$). Student Newman Keuls comparisons with alpha set at .05 indicated that the forward chaining group took longer to finish Test 3 ($M = 132.5$ s) than either the backward ($M = 81.1$ s) or random chaining group ($M = 52.0$ s, $p < .05$).

Location of errors. Descriptive statistics for location of Test 3 errors are presented in Table 25. A repeated measures ANOVA, with number of errors on the symbols corresponding to the first 4 responses of the sequence, the second 4 responses, and the final 4 responses of the sequence as dependent variables revealed a significant Chaining Procedure X Location interaction ($F[4, 138] = 4.9$, $p < .01$). Figure 53 shows the mean errors made

on symbols from each sequence location (beginning, middle, and end) for the three chaining procedures.

This interaction was decomposed using the simple main effects technique. One-way ANOVAs and Student Newman Keuls comparisons revealed a number of differences between pairs of chaining procedures in terms of the location of errors on Test 3. Alpha was set at .05 for the Student Newman Keuls comparisons.

On the symbols that corresponded to responses 1 to 4 of the sequence, Student Newman Keuls comparisons indicated that the backward chaining group (M = 14.2) and the forward chaining group (M = 10.7) made significantly more errors than the random chaining group (M = 3.0, $p < .05$).

On the symbols that corresponded to responses 5 to 8 of the sequence, Student Newman Keuls comparisons revealed that the forward chaining group (M = 12.0) made significantly more errors than the random chaining group (M = 1.0, $p < .05$). The backward chaining group made a mean of 6.6 errors on symbols that corresponded to responses 5 to 8 of the sequence.

On the symbols that corresponded to the last four responses of the sequence, Student Newman Keuls comparisons showed that the forward chaining group (M = 13.8) made significantly more errors than the backward (M = 3.8) and random chaining groups (M = 1.0, $p < .05$).

No significant differences were evident among mean errors on each section of Test 3 for the forward chaining group. The

backward chaining group made significantly fewer errors on symbols from the end (M = 3.8) and middle sections (M = 6.6) than on symbols from the beginning of the sequence (M = 14.2); t[23] = 4.2, $p < .05$ and t[23] = 2.7, $p < .05$ respectively). The random chaining group made significantly more errors on symbols from the beginning of the sequence (M = 3.0) than on symbols from the middle of the sequence (M = 1.0; t[23] = 2.6, $p < .05$).

Acquisition Variables

Means, medians, and standard deviations for the acquisition variables are presented in Table 26. The MANOVA performed on the four acquisition variables (total errors, true errors, total responses, and learning time) was significant (F[8, 134] = 3.9, $p < .001$).

Figure 54 shows the effect of chaining procedure on total errors during acquisition. The overall univariate comparison among chaining groups on total errors was significant (F[2, 69] = 6.7, $p < .01$). Student Newman Keuls comparisons with alpha set at .05 showed that the backward chaining group made more total errors than either the forward or the random chaining group, $p < .05$.

Figure 55 shows mean true errors for each chaining group. The overall univariate comparison among chaining groups on true errors was significant (F[2, 69] = 8.8, $p < .001$). Student Newman Keuls comparisons with alpha set at .05 revealed that the

backward chaining group made more true errors than the forward and random chaining groups, $p < .05$.

Figure 56 shows the means for total responses during acquisition for the three chaining groups. The overall univariate comparison of the chaining groups on total responses during acquisition was significant ($F[2, 69] = 8.3$, $p < .001$). Follow-up Student Newman Keuls comparisons indicated that the backward chaining group made more responses during learning than either the forward or the random chaining group, $p < .05$.

Figure 57 shows the means for learning time for each chaining groups. The overall univariate comparison of the three chaining groups on learning time was significant ($F[2, 69] = 10.5$, $p < .001$). Follow-up Student Newman Keuls comparisons indicated that the forward chaining group learned the sequences faster than either the backward or the random chaining group, $p < .05$.

Discussion

In summary, dependent variables related to Test 1 provide support for the hypothesis that forward chaining results in the best performance on this test, followed by backward chaining, followed by random chaining. The forward chaining group obtained higher percent correct scores and made a greater number of consecutive correct responses than the other two groups. The backward chaining group obtained higher percent correct scores and made a greater number of consecutive correct responses on

Test 1 than did the random chaining group. Random chaining resulted in the poorest performance on Test 1, including significantly longer completion time.

The ranking of the chaining groups on Test 1 variables again provided support for the idea that, during forward chaining, response produced stimulation (the feeling of performing a response, the sight of performing a response, etc.) forms more reliable discriminative stimuli than it does during backward or random chaining.

Test 2 results supported the hypotheses that forward chaining would lead to superior performance on Test 2, random chaining would lead to the poorest performance on Test 2, and backward chaining would result in intermediate performance. Following backward chaining, subjects identified the sequence position of fewer antecedent external stimuli in the sequence than they did following forward chaining or random chaining. Forward chaining led to the identification of more correct pairs than did backward or random chaining. Backward chaining resulted in the identification of more reverse pairs of antecedent external stimuli.

The results of Test 3 supported the hypothesis that subjects would more accurately perform the responses corresponding to the antecedent external stimuli following random chaining than following forward or backward chaining. The results also supported the hypothesis that subjects would exhibit the least

accurate performance on Test 3 following forward chaining. These results provide support for the idea that, following random chaining, the screen symbols retain a controlling role over responding because they are the most reliable discriminative stimuli available. By contrast, following forward chaining, other stimuli have become reliable discriminative stimuli (e.g., response produced stimulation) and control responding.

General Discussion

Theoretical Implications

The theoretical questions behind the investigation of what subjects can do following learning a sequence by a chaining are questions of stimulus control: What is the stimulus control for responding following chaining? Does the stimulus control depend upon the chaining procedure?

Test 1, which required subjects to perform the sequence in the absence of the antecedent external stimuli, was best mastered following forward chaining. Accurate performance on this task implies that some stimulus other than the antecedent external stimulus is exerting stimulus control over responding. The most likely source of stimulus control appears to be response produced stimulation.

It is possible that a shift in stimulus control occurs during forward chaining. Early in training, the screen symbols may exert control over responding whereas later in training, control may shift to response produced stimulation. This notion is supported by the findings regarding the location of errors on Test 1 following forward chaining (Experiment 6). Following forward chaining, more errors were made on Test 1 near the end of the sequence than near the beginning. Errors on Test 1 may be made because the screen symbols and not response produced stimulation are exerting stimulus control over responding. The end of the sequence is the last portion of the sequence learned

in forward chaining and the responses from the end of the sequence may still be controlled by screen symbols rather than by response produced stimulation. Further support for the notion that stimulus control may shift from the screen symbols to response produced stimulation would be evident if the forward chaining group made more errors on symbols from the beginning of the sequence on Test 3. Errors on Test 3 may be made because the screen symbols are not exerting effective stimulus control over responding. If there is a shift in stimulus control from the screen symbols to response produced stimulation over the course of forward chaining, it would most likely occur for those responses that are first learned. However, no evidence was found that the forward chaining group made more errors from the beginning of the sequence on Test 3.

In contrast to Test 1, Test 3 required subjects to respond correctly to the antecedent external stimuli when they were presented out of order. This test was best mastered by random chaining. Good performance on this test supports the notion that stimulus control is exerted by the antecedent external stimuli following random chaining. These stimuli do appear to be the most reliable discriminative stimuli during random chaining. The response produced stimulation (e.g., the feeling of performing a particular button press) is a less reliable source of discriminative stimuli because responses are added to the sequence randomly. That is, during training, the feeling of

performing response 2 is not always followed by the performance of response 3. Consistent with the above discussion of shifts in stimulus control, however, it might be hypothesized that sufficient repetitions of the complete sequence following random chaining would lead to a reduction in the discriminative strength of the external stimuli. No evidence was found to support this hypothesis in Experiment 5.

The stimulus control for responding following chaining is likely resides in some combination of the external stimuli (the screen symbols) and response produced stimulation (feeling and seeing the response). The most accurate performance of a sequence following acquisition will likely be under conditions where the external stimuli are present and the responses are to be performed in order. However, the present findings support the idea that the nature of stimulus control varies as a function of the chaining procedure and may shift over the course of forward chaining. The external stimuli exert more stimulus control following forward chaining than following the other chaining procedures whereas response produced stimulation exerts more stimulus control following random chaining than following the other chaining procedures. If the idea that stimulus control shifts from external stimuli to response produced stimuli over the course of forward chaining is accepted, another question arises: Why would such a shift occur?

The shift in stimulus control from external stimuli to internal response produced stimulation may be a hard-wired aspect of human learning (Watters, personal communication). There may be a tendency for humans to come under control of the most immediate reliable stimuli available. That is, where possible, humans may come under the stimulus control of response produced stimulation rather than relying on external stimuli. This shift of control seems most likely to occur in situations where there is a range of stimuli present, from obvious external stimuli (e.g., the screen symbols) to more subtle internal stimuli (e.g., feeling and seeing the performance of a button press), and where these two types of stimuli are equally reliable discriminative stimuli. Forward chaining is an example of this type of situation.

A tendency for shifts in stimulus control from external stimuli to more immediate internal stimuli may have developed due to its survival value. If immediate response produced stimulation is available as a reliable discriminative stimulus, the organism is not so much at the mercy of a changing external environment where external stimuli may disappear or change. The organism is also likely to perform task more quickly if relying on response produced stimulation for stimulus control over a behavioral sequence and speed is an adaptive (and often reinforced) characteristic of behavior in many situations.

The transfer of control for sequence behaviors from publicly observable external stimuli to subtle and private internal stimuli may also explain the smooth and apparently effortless (to the observer) performance of much practiced behavioral sequences such as dance routines. As a dancer repeatedly practices a dance routine, the stimulus control for the routine shifts from publicly observable discriminative stimuli (e.g., instructions from others, musical sounds) to subtle internal stimuli (physical sensations). Eventually, other than the initial discriminative stimulus to start the sequence (the start of the music), the entire sequence is controlled by subtle private discriminative stimuli.

When differences in post-acquisition performance are explained by referring to stimulus control and shifts in stimulus control, questions arise regarding distinctions between types of discriminative stimuli. When stimulus control shifts, does it shift from one stimulus to another or from one type of stimulus to another? When performing a sequence, a subject is exposed to many stimuli and these stimuli change each time a behavior is performed. All of the stimulus changes that may potentially become discriminative stimuli for subsequent responses may be conceptualized as being response produced.

For example, when performing a sequence like those taught in the present experiments, a subject is exposed to the following response produced stimulation: (a) the subject feels movement in

the arm, hand, and finger, (b) the subject sees the movement of the arm, hand, and finger, (c) the subject feels the button, feels the movement of the button, and if the program is set to act on the particular button press, the subject sees the screen symbol change. When the subject completes a response by removing their finger from the button there is further response produced stimulation (feeling and seeing the movement in the arm, hand, and finger; seeing the key move up and stop moving; etc.). All of these stimulus changes may be conceptualized as response produced in the sense that, if the response of pressing the button did not occur, none of them would be experienced.

Some distinctions between types of response produced stimulation may be useful but are not necessary. For example, a distinction could be made between publicly observable response produced stimulation (e.g., seeing the movement, seeing the screen symbol change) and private response produced stimulation (e.g., feeling the movement). This distinction would become fuzzy, however, if another subject shared the key and could feel the key stop moving. Another distinction that might be useful is between natural stimulus change (e.g., seeing and feeling the movement) and artificial stimulus change (e.g., seeing the screen stimulus change). Again, this distinction is fuzzy when one considers the stimulation of feeling the key move - is the key movement natural or artificial?

If one accepts the idea that all stimulation is, in essence, response produced stimulation, the theoretical positions derived from Skinner's Law of Chaining by Thvedt, Zane, and Walls (1984) are not as distinct as they stated. The two positions were that: (a) the responses were the discriminative stimuli for subsequent responses, or (b) the responses produced the discriminative stimuli for subsequent responses. Thvedt, Zane, and Walls (1984) further assumed that the first position included two possibilities: (a) that the response itself was the discriminative stimulus, and (b) that the internal and natural stimulation produced by the response were the discriminative stimuli. They were therefore, making a distinction between natural response produced stimulation and artificial response produced stimulation. If all stimuli are conceptualized as being response produced, however, research should not ask which position best describes the learning mechanism underlying chaining. Instead, research should ask when, why, and under what conditions does stimulus control shift from one type of response produced stimulation to another.

Another plausible explanation for the present results can be found in the cognitive literature on skill acquisition and controlled versus automatic processing. From this perspective, there are three stages of skill acquisition: cognitive, associative, and autonomous (Anderson, 1985). At the cognitive stage, subjects rehearse a set of facts related to the

performance of a skill and encode these facts in declarative memory. After more practice, subjects reach the associative stage in which errors in performance are identified and corrected. Connections among the elements of the skill are strengthened at this stage and declarative information is encoded in procedural form. The final stage of skill acquisition, the autonomous stage, is characterized by increasingly accurate and rapid performance. During the autonomous stage, which lasts indefinitely, performance improves and verbal mediation of the behavior is reduced.

Progress through the three stages of skill acquisition is dependent on practice. With enough practice, a qualitative shift in the fluency of performance is observed. This shift in performance is hypothesized to be the result of a shift from controlled processing to automatic processing. Shiffrin and Schneider and colleagues (Schneider & Shiffrin, 1977; Shiffrin & Dumais, 1981; Shiffrin & Schneider, 1977) have articulated a theory of controlled and automatic processing based on an extensive series of experiments on cognitive processing.

According to this theory, controlled processing is slow, effortful, requires conscious attention, and is used when dealing with novel and unfamiliar tasks. By contrast, automatic processing is rapid, seemingly effortless, does not require conscious control, and is not limited by short term memory capacity. Automatic processing develops as a result of practice

in which subjects repeatedly process information in a consistent manner.

Schneider and Fisk (1983) applied the theory of controlled and automatic processing to skill acquisition. They hypothesized that the qualitative difference between novice and expert performance of a skill is a result of these qualitatively different forms of information processing. From this perspective, the elements of practice and similarity of training and test procedure are important explanatory factors in the development of speed and accuracy. The more similar the training procedure is to the test procedure, the more practice subjects will have before testing. The more practice subjects have before testing the more automatic (rapid and accurate) test performance will be.

From this perspective then, superior performance on Test 1 following forward chaining would be explained as a function of the development of more automatic processing through practice. In forward chaining, the subjects got considerable practice at performing the sequence behaviors in order, which is what they were required to do on Test 1. From a cognitive perspective, forward chaining would be hypothesized to lead to a shift in processing from controlled to automatic. From a behaviorist perspective, forward chaining would be hypothesized to lead to a shift of stimulus control from external to internal.

The present results can be adequately explained by both the cognitive and the behaviorist perspectives. However, the research was undertaken from a behaviorist perspective and the experiments were designed to test hypotheses generated by the behaviorist framework. Adopting a cognitive explanatory framework would have implications for further research on post-acquisition behavior. For example, subjects might be overtrained to insure a higher level of automatic processing and then asked to perform a secondary task that required controlled processing. Other dependent variables, such as response latencies, would also be of interest from this perspective. The present research has established some clear post-acquisition patterns. Future research could make a valuable contribution by testing more specific hypotheses generated from either a behaviorist or a cognitive perspective or could attempt to establish the relative adequacy of each explanatory framework.

Practical Implications

The practical implications of the results are limited by the nature of the task and the population used. The extent to which the present results are generalizable to applied settings where more powerful reinforcers are used depends, in part, on the extent to which the percent correct scores can be considered effective reinforcers. As noted in the introduction, Martin et al. (1981) suggested that the nature of reinforcers and their delivery may account for differences between findings in human

and animal chaining research. Such differences in reinforcers and their delivery may also lead to differences between the present findings and those in applied settings. However, provided the results are replicable using applied tasks with appropriate populations, a number of tentative recommendations can be made regarding the use of the chaining procedures.

Forward chaining. Taken together, the results of the six experiments indicated that forward chaining leads to superior performance of the sequence in the absence of the external antecedent stimuli and superior ordering of the external antecedent stimuli. Forward chaining leads to the poorest performance of responses following the presentation of the external antecedent stimuli in an order other than that used during training.

In view of these results, forward chaining may be the training procedure of choice for teaching tasks when it is desirable that antecedent stimuli (e.g., instructions, flash cards, a script) used during training be faded following training. It should also be investigated as the procedure of choice for tasks that require learners to order the antecedent external stimuli following training, for example, in situations where the learner will later be required to give directions to other individuals. Forward chaining does not appear to lead to ease in emitting sequence responses, out of sequence order following antecedent external stimuli used during training and

may be best avoided for tasks where such performance is desired following training.

If the nature of the task, setting, or subjects dictates that the number of errors made during training be kept at a minimum, the acquisition results indicate that forward chaining would be an appropriate choice of procedures. For example, when teaching tasks such as surgical or navigational procedures where errors are expensive and dangerous, forward chaining would limit the expense and danger. In settings where one teacher is responsible for teaching a number of individuals a task simultaneously (e.g., a classroom), minimizing errors by using forward chaining would save time and provide for relatively smooth progress through the task by a group of individuals. Forward chaining may be the procedure of choice for teaching children as children perform best when the density of reinforcement is high. A high density of reinforcement will be easiest to maintain if relatively few errors are made during training.

Backward chaining. Backward chaining did not emerge as clearly superior on any of the three tests. Due to practical considerations, random chaining may sometimes be difficult and time consuming to program. Arranging for a random order of response acquisition requires more planning time and record keeping than simply arranging for a backward chaining order. In such cases, backward chaining may be the procedure of choice

where it is desirable that subjects are able to produce sequence responses following the presentation of external antecedent stimuli in an order other than that used during training.

In general, however, the findings failed to identify any significant advantages associated with backward chaining. This failure is of note due to the fact that basic learning sources have long asserted that behavior sequences should be taught by backward chaining (e.g., Ferster & Perrott, 1968; Gilbert, 1962a, 1962b; Kazdin, 1989; Rachlin, 1976). The present post-acquisition findings suggest that backward chaining is best characterized as a mediocre procedure.

Whole task training. Whole task training produced relatively good performance on Test 2 and may be an appropriate choice when ordering of the antecedent external stimuli is desirable following training. However, if errors during acquisition need to be minimized at the same time, forward chaining would be the superior choice of procedures. In contrast to earlier research findings (Martin et al., 1981; Spooner, 1984; Watters & Scott, in press), no evidence was found for the superiority of whole task training in terms of acquisition time. It is possible that the length and homogeneity of the sequence used here made it difficult to learn quickly by whole task training.

Random chaining. Random chaining produced the best performance on Test 3 and the worst performance on Tests 1 and 2.

Random chaining would be the procedure of choice when it is desirable that the external antecedent stimuli retain control over responses following training. Scott and Watters (in press) found that random chaining and backward chaining were indistinguishable in terms of acquisition variables. In Experiment 6 the significant differences between random chaining and backward chaining on acquisition variables favored the random chaining group who made fewer total errors, fewer true errors, and fewer responses during acquisition. Taken together with the acquisition results, the post-acquisition results further support the use of random chaining over backward chaining, particularly for tasks where control by external antecedent stimuli is desirable following training.

Future Research and Limitations

A methodological refinement that would enhance the current findings would be to introduce a timed version of Test 2. Both Test 1 and Test 3 were timed to provide an additional measure of performance. In terms of conceptual completeness, it may be interesting to examine subjects performance on a number of other post-acquisition tests. The three tests used here represent three conditions under which to assess subjects' post-acquisition performance: (a) absence of screen symbols and performance of responses (Test 1), (b) presence of screen symbols and no performance of responses (Test 2), and (c) presence of screen symbols and performance of responses (Test 3).

Test 1 removed the screen symbols and asked subjects to perform the sequence in their absence. A variation of this test would be to instruct subjects to perform a randomly chosen response from the sequence and then asking them to perform the next response from the sequence.

Test 2 removed the responses from the screen symbol-response contingency and asked subjects to order the entire sequence of screen symbols. Another possibility would be to present a single randomly chosen screen symbol and asking subjects to identify the following screen symbol.

Test 3 presented the screen symbols and required subjects to perform the appropriate responses. Subjects were asked to make the appropriate response following each screen symbol when the symbols were presented out of order. Another variation of this condition is the used for the criterion behavior at the end of the acquisition phase where the screen symbols are presented in order and the subject is required to perform the sequence in order. Presenting the screen symbols in order and requesting the subject to perform the correct responses has been as a measure of retention by Watters (in press). Another variation on this condition would be to instruct subjects to perform a randomly chosen response and then ask them to identify the following screen stimulus. That is, subjects would be asked to make each of the responses from the sequence and to identify the stimulus that appeared during training following each response.

The experiments reported here have the advantage of being controlled laboratory experiments. As such they allow for clear conclusions regarding the impact of chaining procedures on particular types of post-acquisition performance following the performance to criterion of a particular type of behavioral sequence. The results show patterns for each type of post-acquisition behavior that are consistent enough to be described with confidence. The experiments also allow for preliminary suggestions regarding the practical applications of various chaining techniques. The suggestions, though preliminary, are refinements to those in extant literature because they begin to answer the question, Which technique for which types of behaviors? rather than the more general question, Which technique?

Laboratory experiments also have significant limitations. In particular, the nature of the task and the population restricts generalization to more heterogeneous tasks, more difficult tasks, and other populations. This presents a problem in chaining research because chaining procedures are frequently applied to heterogeneous behavioral chains, like assembly tasks, and to handicapped populations or to children. Further research should test the hypotheses suggested by the present results using varying behavioral sequences and populations. Further research should also examine what subjects can do after longer periods of time have elapsed since acquisition of a sequence by chaining.

The advantage of the present experiments is that the very control that limits the generalizability of the findings, allows for the identification of clear patterns. Once such patterns are identified, research of a more applied nature can test the specific hypotheses generated in more controlled settings.

References

Anderson, J.R. (1985). Cognitive Psychology and Its Implications. NY: W.H. Freeman.

Azrin, N.H, Schaeffer, R.M., & Wesolowski, M.D. (1976). A rapid method of teaching profoundly retarded persons to dress by a reinforcement-guidance method. Mental Retardation, 14, 29-33.

Bellamy, G.T., Horner, R.H., & Inman, D.P. (1979). Vocational Habilitation of Severely Retarded Adults: A direct service technology. Baltimore: University Park Press.

Chartrand, D. (1990). The effects of timeout on the learning of sequenced behavior by humans. Unpublished manuscript.

Cornell, R.K. (1975). The relative effectiveness of forward chaining, reverse chaining, and current teaching methods in the acquisition, retention, and transfer in spelling (Doctoral dissertation, University of Oregon, 1974). Dissertation Abstracts International, 48, 4976-4977.

Ferster, C.B. & Perrott, M.C. (1968). Behavior Principles. New York: Appleton-Century-Crofts.

Gilbert, T.F. (1962a). Mathetics: The technology of education. The Journal of Mathetics, 1, 7-73.

Gilbert, T.F. (1962b). Mathetics II: The design of teaching exercises. The Journal of Mathetics, 2, 7-56.

Giles, G.M. & Morgan, J.H. (1989). Training functional skills following herpes simplex encephalitis: A single case study. Journal of Clinical and Experimental Neuropsychology, 11, 311-318.

Kazdin, A.E. (1989). Behavior Modification in Applied Settings. California: Brooks/Cole.

Martin, G., Koop, S., Turner, G., & Hanel, F. (1981). Backward chaining versus total task presentation to teach assembly tasks to severely retarded persons. Behavior Research of Severe Developmental Disabilities, 2, 117-136.

Martin, G. & Pear, J. (1983). Behavior Modification: What it is and how to do it. (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Martin, G. & Pear, J. (1988). Behavior Modification: What it is and how to do it. (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Nannay, R.W. (1971). The effectiveness of teaching a psychomotor task via forward and backward chaining. Dissertation Abstracts International, 31, 3922.

Nettelbeck, T. & Kirby, N.H. (1976). A comparison of part and whole training methods with mildly mentally retarded workers. Journal of Occupational Psychology, 49, 115-120.

Rachlin, H. (1976). Behavior and Learning. San Francisco: W.H. Freeman.

Ross, A.O. (1985). To form a more perfect union: It is time to stop standing still. Behavior Therapy, 16, 195-204.

Schneider, W. & Fisk, A.D. (1983). Attention theory and mechanisms for skilled performance. In R.A. Magill (Ed.), Memory and Control of Action, pp. 119-143. NY: North-Holland.

Schneider, W. & Shiffrin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Schwartz, B. (1984). Psychology of Learning and Behavior. (2nd ed.). New York: Norton.

Scott, J.M. (1990). Teaching sequences of behaviors to humans by forward chaining, backward chaining, and whole task training. Unpublished doctoral dissertation, University of Ottawa, Canada.

Scott, J.M. & Watters, R.G. (1992). Human sequenced behavior: A comparison of forward chaining, backward chaining, random chaining, and whole task training. Manuscript submitted for publication.

Shiffrin, R.M. & Dumais, S.T. (1981). The development of automatism. In J.R. Anderson (Ed.), Cognitive Skills and Their Acquisition. Hillsdale, NJ: Lawrence Erlbaum.

Shiffrin, R.M. & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. Psychological Review, 84, 127-190.

Simek, T.C. & O'Brien, R.M. (1989). A chaining-mastery discrimination training program to teach little leaguer to hit a baseball. Human Performance, 7, 73-84.

Skinner, B.F. (1938). The Behavior of Organisms. New York: Appleton-Century-Croft.

Spooner, F. (1984). Comparisons of backward chaining and total task presentation in training severely handicapped persons. Education and Training of the Mentally Retarded, 19, 15-22.

Thvedt, J.E., Zane, T., & Walls, R.T. (1984). Stimulus functions in response chaining. American Journal of Mental Deficiency, 88, 661-667.

Watters, R.G. (1989). Human sequenced behavior: A comparison of forward chaining, backward chaining, and whole task training procedures. Journal of Human Movement Studies, 16, 183-190.

Watters, R.G. (1990). Human sequenced behavior: A comparison of forward chaining, backward chaining, and whole task training at different sequence lengths. Journal of Human Movement Studies, 18, 25-35.

Watters, R.G. (in press). Retention of human sequenced behavior following forward chaining, backward chaining, and whole task training procedures. Journal of Human Movement Studies.

Watters, R.G. & Scott, J.M. (in press). Teaching human sequenced behavior with and without prompts: A comparison of forward chaining, backward chaining, and whole task procedures. Journal of Human Movement Studies.

Weber, N.J. (1978). Chaining strategies for teaching sequenced motor tasks to mentally retarded adults. American Journal of Occupations Therapy, 32, 385-389.

Weiss, K.M. (1978). A comparison of forward and backward procedures for the acquisition of response chains in humans. Journal of the Experimental Analysis of Behavior, 29, 255-259.

Wightman, D.C. & Sistrunk, F. (1987). Part-task training strategies in simulated carrier landing final-approach training. Human Factors, 29, 245-254.

Wilcox, B. (1974). The teaching of serial tasks using chaining strategies. British Journal of Educational Psychology, 44, 175-183.

Table 1

Means, Standard Deviations and Medians for Test 1 Variables by Chaining Procedure: Experiment 1 (N = 18)

	Chaining Procedure		
Variable	Forward	Backward	Whole Task
Percent correct			
M	95.0	94.5	82.3
SD	9.6	10.6	24.1
Median	100.0	100.0	100.0
Correct responses prior to first error			
M	8.2	7.2	7.8
SD	3.1	3.9	2.7
Median	10.0	10.0	10.0
Longest series			
M	9.0	9.1	8.1
SD	1.6	1.6	2.2
Median	10.0	10.0	10.0
Beginning of longest series			
M	1.6	1.4	1.9
SD	1.2	1.1	1.8
Median	1.0	1.0	1.0
Completion time			
M	8.1	15.1	21.9
SD	4.0	11.7	23.9
Median	6.5	15.0	12.1

Table 2

Means, Standard Deviations and Medians for Location of Test 1 Errors by Chaining Procedure: Experiment 1 (N = 18)

	Chaining Procedure		
Location	Forward	Backward	Whole Task
Beginning of sequence (responses 1 to 5)			
M	0.4	0.3	2.1
SD	1.4	0.5	6.3
Median	0.0	0.0	0.0
End of sequence (responses 6 to 10)			
M	0.2	0.5	1.8
SD	0.5	1.9	3.6
Median	0.0	0.0	0.0

Table 3

Means, Standard Deviations and Medians for Test 2 Variables by Chaining Procedure: Experiment 1 (N = 18)

Variable	Chaining Procedure		
	Forward (n = 6)	Backward (n = 6)	Whole Task (n = 6)
Percent correct			
M	58.3	56.7	38.3
SD	34.3	42.7	31.3
Median	60.0	60.0	25.0
Correct pairs			
M	4.3	4.5	4.0
SD	2.7	3.8	3.3
Median	4.0	3.5	3.5
Reverse pairs			
M	1.0	0.8	0.5
SD	0.9	1.0	1.2
Median	1.0	0.5	0.0

Table 4

Means, Standard Deviations and Medians for Acquisition Variables: Experiment 1 (N = 18)

	Chaining Procedure		
Location	Forward	Backward	Whole Task
Total errors			
M	41.3	65.7	211.9
SD	28.7	68.6	132.6
Median	31.5	44.0	183.5
True errors			
M	14.6	34.6	169.1
SD	26.5	65.5	132.3
Median	3.0	7.5	134.0
Total responses			
M	286.6	361.3	399.1
SD	71.6	152.5	216.9
Median	257.5	295.0	363.0
Learning time			
M	502.7	688.2	598.7
SD	150.1	385.3	266.5
Median	449.4	543.7	570.8

Table 5

Means, Standard Deviations and Medians for Test 1 Variables by Chaining Procedure: Experiment 2 (N = 24)

	Chaining Procedure		
Variable	Forward	Backward	Whole Task
Percent correct			
M	91.1	77.6	76.7
SD	14.3	29.1	23.2
Median	100.0	91.0	87.1
Correct responses prior to first error			
M	8.2	5.2	5.5
SD	2.5	4.7	3.2
Median	10.0	5.0	5.0
Longest series			
M	8.5	7.8	6.4
SD	2.0	2.8	2.4
Median	10.0	10.0	5.5
Beginning of longest series			
M	1.6	1.8	2.6
SD	1.3	1.7	2.0
Median	1.0	1.0	1.0
Completion time			
M	15.4	37.3	27.2
SD	9.2	40.5	15.7
Median	12.1	17.6	22.1

Table 6

Means, Standard Deviations and Medians for Location of Test 1 Errors by Chaining Procedure: Experiment 2 (N = 24)

	Chaining Procedure		
Location	Forward	Backward	Whole Task
Beginning (responses 1 to 5)			
M	0.2	3.9	0.8
SD	0.5	9.1	1.5
Median	0.0	0.5	0.0
End (responses 6 to 10)			
M	1.2	4.1	4.1
SD	2.8	10.2	5.5
Median	0.0	0.0	1.5

Table 7

Means, Standard Deviations and Medians for Test 2 Variables by Chaining Procedure: Experiment 2 (N = 24)

	Chaining Procedure		
Variable	Forward	Backward	Whole Task
Percent correct			
M	64.6	33.8	52.1
SD	30.2	38.0	26.7
Median	65.0	15.0	45.0
Correct pairs			
M	5.1	3.4	4.1
SD	3.1	3.0	2.3
Median	4.5	2.0	4.0
Reverse pairs			
M	0.8	1.8	0.9
SD	1.0	2.4	0.8
Median	0.5	1.0	1.0

Table 8

Means, Standard Deviations and Medians for Acquisition Variables: Experiment 2 (N = 24)

	Chaining Procedure		
Location	Forward	Backward	Whole Task
Total errors			
M	34.0	56.4	151.4
SD	10.9	31.8	100.4
Median	33.0	56.0	119.5
True errors			
M	6.5	28.1	110.5
SD	8.5	27.9	90.9
Median	3.5	18.5	85.0
Total responses			
M	147.6	192.0	255.3
SD	31.7	72.1	141.0
Median	140.5	186.0	216.5
Learning time			
M	326.8	450.8	495.6
SD	78.1	171.0	220.1
Median	308.5	390.5	412.0

Table 9

Means, Standard Deviations and Medians for Test 3 Variables by Chaining Procedure: Experiment 3 (N = 24)

	Chaining Procedure		
Variable	Forward	Backward	Whole Task
Percent correct			
M	66.2	85.8	72.2
SD	27.1	16.3	25.8
Median	72.8	91.1	78.5
Completion time			
M	73.3	47.7	71.1
SD	44.9	21.5	50.7
Median	60.1	42.8	51.1

Table 10

Means, Standard Deviations and Medians for Location of Test 3 Errors by Chaining Procedure: Experiment 3 (N = 24)

	Chaining Procedure		
Location	Forward	Backward	Whole Task
Symbols 1 to 5			
M	6.0	2.4	7.6
SD	8.6	3.4	14.3
Median	1.5	1.0	1.5
Symbols 6 to 10			
M	11.8	2.0	6.8
SD	15.3	3.8	9.1
Median	4.0	0.0	3.5

Table 11

Means, Standard Deviations and Medians for Acquisition Variables: Experiment 3 (N = 24)

Location	Chaining Procedure		
	Forward	Backward	Whole Task
Total errors			
M	41.8	69.0	162.4
SD	22.6	89.4	95.5
Median	33.5	44.0	141.5
True errors			
M	11.0	39.1	127.1
SD	19.9	85.7	93.3
Median	4.0	11.0	104.5
Total responses			
M	153.8	206.8	280.5
SD	42.3	124.7	145.3
Median	138.5	165.0	257.0
Learning time			
M	328.6	471.7	541.0
SD	88.7	249.3	239.0
Median	302.9	396.0	492.1

Table 12

Means, Standard Deviations and Medians for Test 3 Variables by Chaining Procedure: Experiment 4 (N = 30)

	Chaining Procedure		
Variable	Forward	Backward	Random
Percent correct			
M	76.1	80.7	89.5
SD	26.3	18.3	9.9
Median	85.8	85.7	92.3
Completion time			
M	75.8	60.5	46.2
SD	54.2	27.3	15.3
Median	52.1	47.0	41.8

Table 13

Means, Standard Deviations and Medians for Location of Test 3 Errors by Chaining Procedure: Experiment 4 (N = 30)

Location	Chaining Procedure		
	Forward	Backward	Random
Symbols 1 to 4			
M	6.5	4.7	1.8
SD	10.8	5.6	2.4
Median	1.0	3.0	1.0
Symbols 5 to 8			
M	6.2	1.7	1.2
SD	13.6	3.4	2.1
Median	1.0	0.0	0.5
Symbols 9 to 12			
M	4.2	1.3	0.1
SD	9.2	3.8	0.3
Median	0.5	0.0	0.0

Table 14

Means, Standard Deviations and Medians for Acquisition Variables: Experiment 4 (N = 30)

	Chaining Procedure		
Location	Forward	Backward	Random
Total errors			
M	52.2	68.7	65.5
SD	18.0	47.5	21.1
Median	49.0	57.0	59.0
True errors			
M	10.3	29.8	21.2
SD	13.4	43.1	18.8
Median	5.0	17.5	15.5
Total responses			
M	227.1	257.5	257.1
SD	53.0	79.0	62.7
Median	224.0	235.5	243.0
Learning time			
M	503.6	587.0	601.1
SD	192.4	192.2	153.5
Median	454.5	533.5	603.2

Table 15

Means, Standard Deviations, and Medians for Test 3 Variables: Experiment 5, Order 1-4 (n = 30)

Variable	Chaining Procedure		
	Forward (n = 10)	Backward (n = 10)	Random (n = 10)
Percent correct			
M	62.9 (97.6)	65.1 (89.3)	91.8 (98.5)
SD	22.4 (2.7)	24.9 (20.8)	5.1 (3.2)
Median	61.6 (98.0)	69.9 (98.0)	88.9 (100.0)
Completion time			
M	140.5 (45.2)	69.8 (52.1)	45.2 (35.1)
SD	147.7 (9.5)	28.8 (31.8)	13.4 (4.8)
Median	84.8 (45.0)	60.9 (41.5)	40.6 (33.8)

Note. Values for variables at PC 4 are in parentheses, other values are for PC 1.

Table 16

Means, Standard Deviations, and Medians for Test 3 Variables: Experiment 5, Order 4-1 (n = 30)

Variable	Chaining Procedure		
	Forward (n = 10)	Backward (n = 10)	Random (n = 10)
Percent correct			
M	65.9 (51.0)	70.3 (82.0)	88.2 (85.3)
SD	31.8 (34.4)	29.8 (23.7)	20.7 (19.8)
Median	71.8 (34.1)	80.1 (89.0)	98.0 (100.0)
Completion time			
M	96.4 (144.6)	89.5 (71.4)	44.9 (57.0)
SD	54.0 (88.8)	88.2 (64.1)	17.6 (46.0)
Median	85.2 (153.7)	65.4 (44.5)	40.4 (37.9)

Note. Values for variables at PC 4 are in parentheses, other values are for PC 1.

Table 17

Means, Standard Deviations, and Medians for Location of Test 3 Errors: Experiment 5, Order 1-4 (n = 30)

Location	Chaining Procedure		
	Forward (n = 10)	Backward (n = 10)	Random (n = 10)
Symbols 1 to 4			
M	7.5 (0.2)	10.0 (3.7)	0.7 (0.2)
SD	9.7 (0.6)	11.7 (9.2)	1.3 (0.4)
Median	1.5 (0.0)	4.5 (0.0)	0.0 (0.0)
Symbols 5 to 8			
M	4.3 (0.2)	2.8 (1.2)	1.1 (0.0)
SD	7.1 (0.4)	6.9 (3.5)	1.1 (0.0)
Median	2.5 (0.0)	0.0 (0.0)	1.0 (0.0)
Symbols 9 to 12			
M	7.6 (0.2)	6.0 (0.6)	0.4 (0.2)
SD	10.0 (0.4)	8.6 (1.3)	0.7 (0.4)
Median	4.0 (0.0)	2.0 (0.0)	0.0 (0.0)

Note. Values for variables at PC 4 are in parentheses, other values are for PC 1.

Table 18

Means, Standard Deviations, and Medians for Location of Test 3 Errors: Experiment 5, Order 4-1 (n = 30)

Location	Chaining Procedure		
	Forward (n = 10)	Backward (n = 10)	Random (n = 10)
Symbols 1 to 4			
M	12.0 (10.2)	12.5 (3.4)	3.3 (4.9)
SD	20.6 (11.8)	19.6 (6.0)	8.4 (7.0)
Median	0.0 (7.0)	5.0 (0.5)	0.0 (0.0)
Symbols 5 to 8			
M	8.7 (16.6)	7.6 (3.0)	1.2 (0.1)
SD	12.4 (18.3)	15.9 (5.5)	3.2 (0.3)
Median	2.5 (13.5)	0.5 (1.0)	0.0 (0.0)
Symbols 9 to 12			
M	8.1 (21.5)	5.7 (2.8)	1.4 (0.9)
SD	10.9 (20.8)	10.1 (7.5)	4.1 (2.0)
Median	3.5 (22.5)	0.5 (0.0)	0.0 (0.0)

Note. Values for variables at PC 4 are in parentheses, other values are for PC 1.

Table 19

Means, Standard Deviations, and Medians for Acquisition Variables: Experiment 5, Order 1-4 (n = 30)

Location	Chaining Procedure		
	Forward (n = 10)	Backward (n = 10)	Random (n = 10)
Total errors			
M	93.6 (49.7)	111.0 (47.3)	62.9 (43.3)
SD	121.3 (18.3)	77.8 (13.1)	16.4 (8.9)
Median	53.0 (46.5)	93.5 (42.5)	57.0 (43.5)
True errors			
M	49.9 (8.3)	61.4 (6.7)	21.3 (3.9)
SD	113.9 (10.4)	70.4 (6.0)	17.7 (2.6)
Median	13.0 (5.0)	42.5 (5.0)	19.5 (3.0)
Total responses			
M	285.1 (524.8)	358.7 (500.7)	267.5 (454.8)
SD	186.9 (145.1)	158.8 (99.7)	60.6 (40.2)
Median	223.0 (516.0)	320.0 (469.0)	262.0 (440.5)
Learning time			
M	583.4 (945.8)	713.4 (977.2)	647.0 (913.9)
SD	338.1 (163.1)	360.6 (191.6)	165.5 (78.4)
Median	493.3 (914.5)	593.3 (917.1)	596.8 (922.0)

Note. Values for variables at PC 4 are in parentheses, other values are for PC 1.

Table 20

Means, Standard Deviations, and Medians for Acquisition Variables: Experiment 5, Order 4-1

	Chaining Procedure		
Location	Forward	Backward	Random
Total errors			
M	69.0 (57.1)	58.5 (66.7)	73.7 (59.5)
SD	44.5 (21.8)	24.4 (42.5)	57.3 (35.1)
Median	50.5 (52.5)	51.0 (56.0)	62.5 (52.0)
True errors			
M	21.3 (15.4)	19.0 (25.5)	29.8 (17.9)
SD	35.9 (18.7)	17.7 (35.5)	44.9 (31.1)
Median	8.0 (10.5)	13.5 (16.0)	19.5 (8.5)
Total responses			
M	256.9 (521.1)	239.9 (571.2)	276.3 (507.2)
SD	107.0 (83.2)	55.8 (169.1)	133.8 (99.5)
Median	214.5 (518.0)	230.0 (496.5)	241.5 (492.0)
Learning time			
M	518.2 (1006.4)	451.1 (1057.6)	602.3 (1137.3)
SD	220.5 (290.3)	100.0 (283.4)	282.1 (289.4)
Median	437.9 (887.4)	436.0 (1051.3)	525.2 (1108.4)

Note. Values for variables at proficiency criterion of 4 are in parentheses, other values are for PC 1.

Table 21

Means, Standard Deviations and Medians for Test 1 Variables by Chaining Procedure: Experiment 6 (N = 72)

Variable	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Percent correct			
M	68.5	44.6	25.4
SD	26.6	31.4	21.3
Median	75.0	32.9	18.5
Correct responses prior to first error			
M	6.7	2.0	0.5
SD	4.1	3.7	1.5
Median	7.0	0.0	0.0
Longest series			
M	7.9	5.7	2.5
SD	3.0	3.2	2.2
Median	8.0	5.0	2.0
Beginning of longest series			
M	2.4	4.5	5.3
SD	1.8	2.7	2.6
Median	1.0	5.3	5.8
Completion time			
M	47.4	97.5	173.5
SD	35.0	75.8	166.8
Median	33.7	74.2	114.2

Table 22

Means, Standard Deviations and Medians for Location of Test 1 Errors by Chaining Procedure: Experiment 6 (N = 72)

Location	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Symbols 1 to 4			
M	1.2	15.6	25.5
SD	3.1	15.9	19.0
Median	0.0	10.5	18.5
Symbols 5 to 8			
M	3.0	8.8	17.0
SD	9.4	9.8	11.7
Median	0.0	5.0	14.0
Symbols 9 to 12			
M	5.5	8.6	16.2
SD	8.4	14.1	12.6
Median	2.0	0.0	16.5

Table 23

Means, Standard Deviations and Medians for Test 2 Variables by Chaining Procedure: Experiment 6 (N = 72)

Variable	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Percent correct			
M	51.7	10.1	14.2
SD	25.9	21.6	22.7
Median	50.0	0.0	8.3
Correct pairs			
M	4.6	1.7	1.8
SD	3.2	2.2	2.8
Median	4.0	1.0	1.0
Reverse pairs			
M	0.9	4.1	0.5
SD	1.2	2.7	0.6
Median	0.0	4.0	0.0

Table 24

Means, Standard Deviations and Medians for Test 3 Variables by Chaining Procedure: Experiment 6 (N = 72)

Variable	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Percent correct			
M	52.9	58.4	85.6
SD	26.5	22.2	13.9
Median	48.5	58.5	92.3
Completion time			
M	132.5	81.1	52.0
SD	90.6	32.7	16.3
Median	104.0	77.6	51.2

Table 25

Means, Standard Deviations and Medians for Location of Test 3 Errors by Chaining Procedure: Experiment 6 (N = 72)

Location	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Symbols 1 to 4			
M	10.7	14.2	3.0
SD	13.3	12.3	4.3
Median	6.0	10.5	1.0
Symbols 5 to 8			
M	12.0	6.6	1.0
SD	15.3	8.8	1.5
Median	2.0	2.0	0.0
Symbols 9 to 12			
M	13.8	3.8	1.0
SD	11.1	7.4	1.9
Median	13.0	0.0	0.0

Table 26

Means, Standard Deviations and Medians for Acquisition Variables: Experiment 6 (N = 72)

Location	Chaining Procedure		
	Forward (n = 24)	Backward (n = 24)	Random (n = 24)
Total errors			
M	59.3	115.0	81.1
SD	22.3	72.4	52.4
Median	52.5	99.5	60.0
True errors			
M	15.8	72.2	37.9
SD	17.0	61.4	50.6
Median	8.5	59.0	21.0
Total responses			
M	239.5	361.4	286.6
SD	59.7	136.8	102.0
Median	216.5	337.5	258.0
Learning time			
M	511.4	784.9	667.5
SD	173.7	230.6	212.9
Median	480.4	781.3	621.9

Figures

Note. The error bars on all figures represent the standard error of the means.

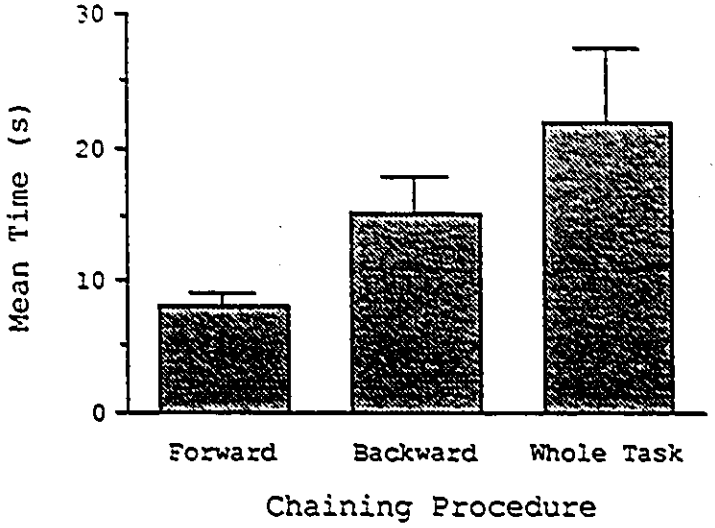


Figure 1. Mean Test 1 completion time as a function of chaining procedure, Experiment 1

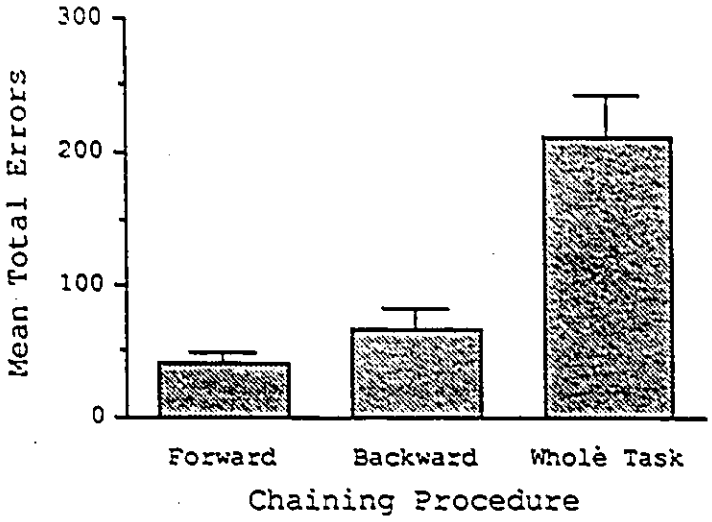


Figure 2. Mean total errors as a function of chaining procedure, Experiment 1

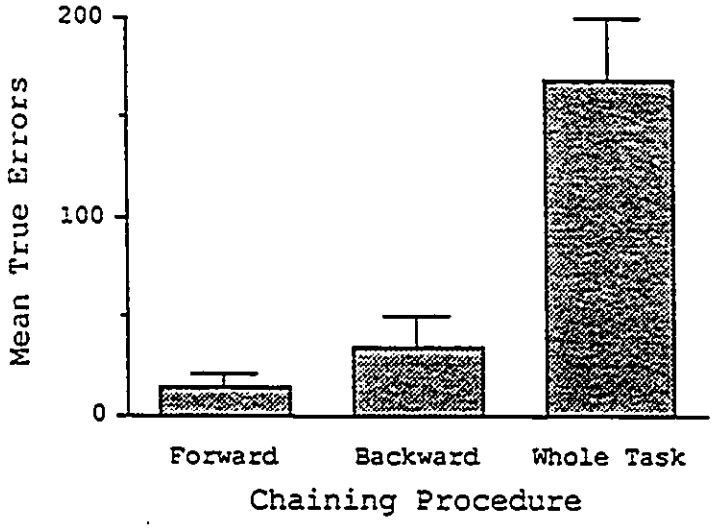


Figure 3. Mean true errors as a function of chaining procedure, Experiment 1

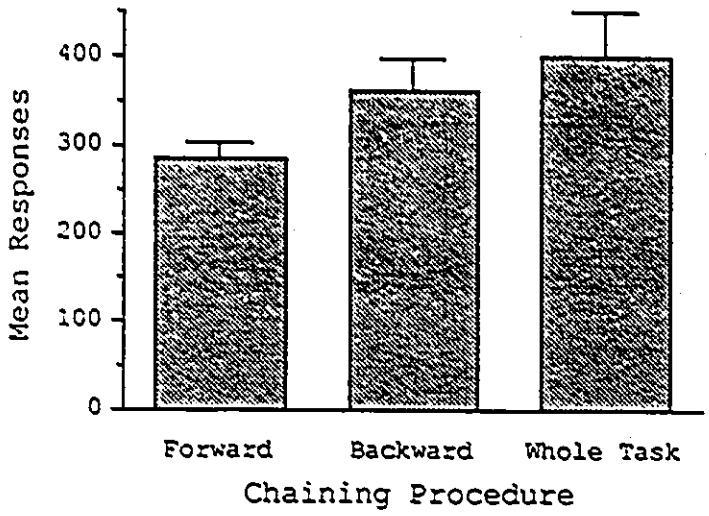


Figure 4. Mean total responses as a function of chaining procedure, Experiment 1

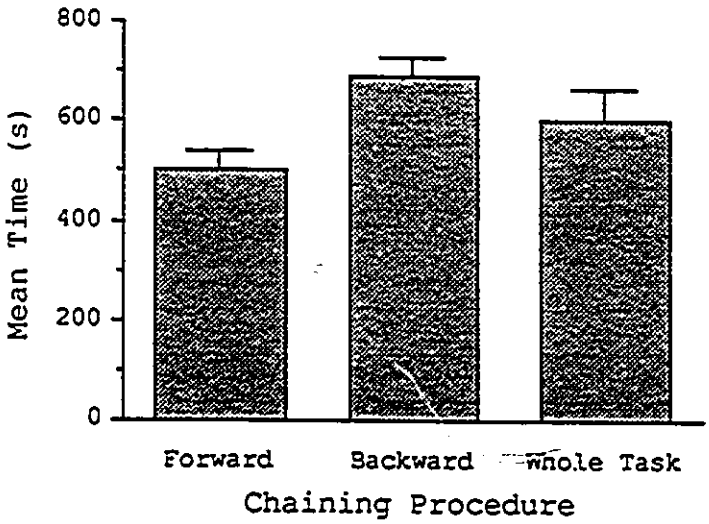


Figure 5. Mean learning time as a function of chaining procedure, Experiment 1

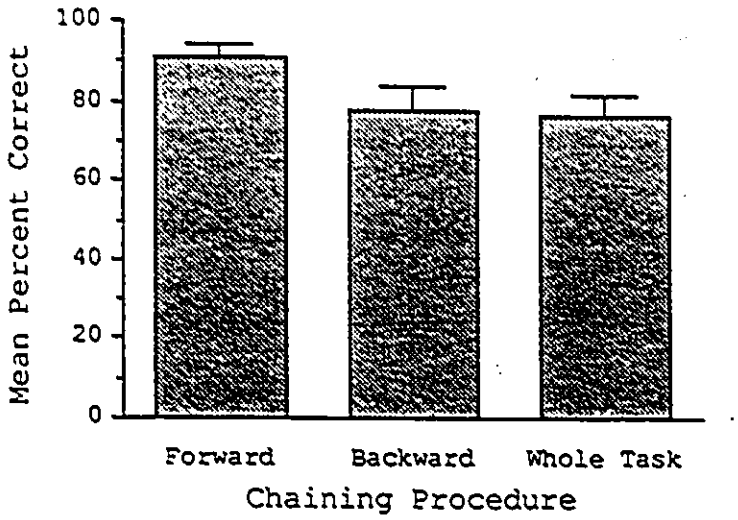


Figure 6. Mean percent correct on Test 1 as a function of chaining procedure, Experiment 2

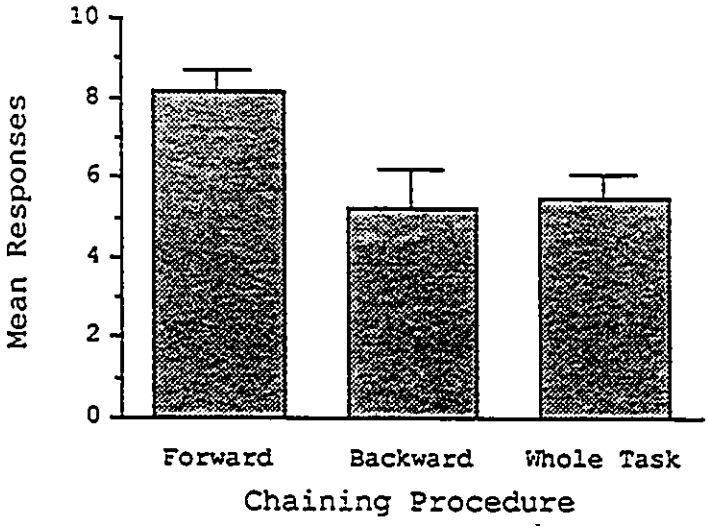


Figure 7. Mean correct responses prior to first error on Test 1 as a function of chaining procedure, Experiment 2

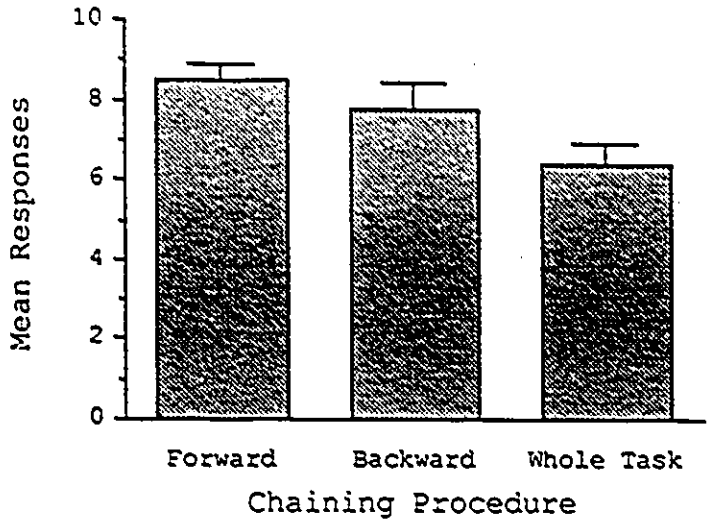


Figure 8. Mean longest series of correct responses on Test 1 as a function of chaining procedure, Experiment 2

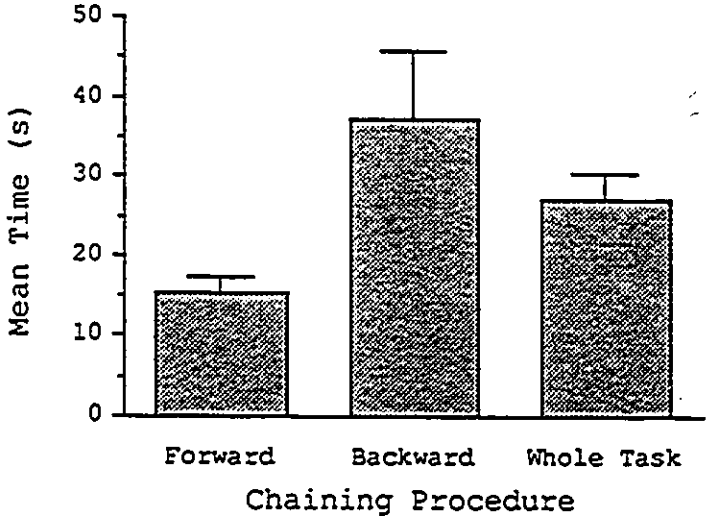


Figure 9.. Mean Test 1 completion time as a function of chaining procedure, Experiment 2

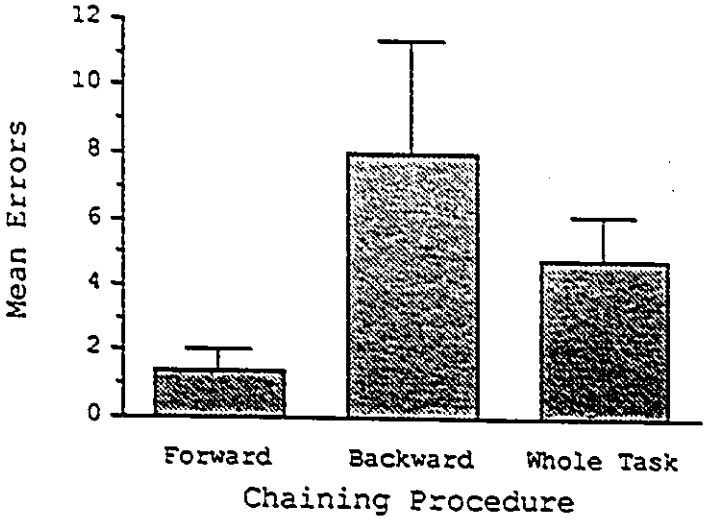


Figure 10. Mean errors on Test 1 as a function of chaining procedure, Experiment 2

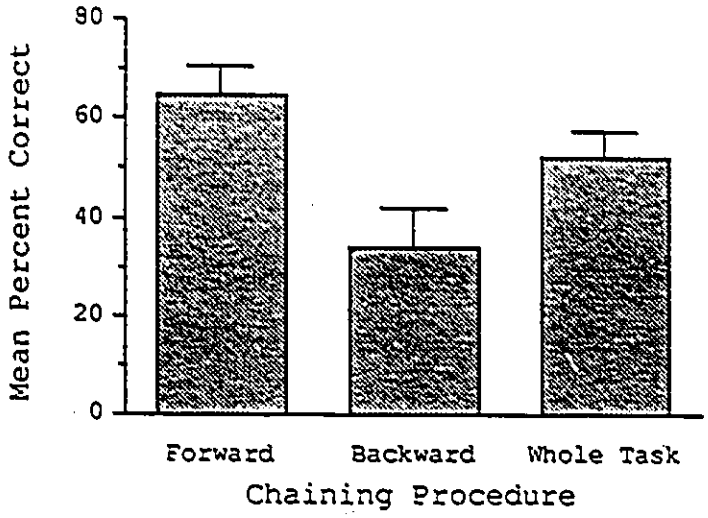


Figure 11. Mean percent correct on Test 2 as a function of chaining procedure, Experiment 2

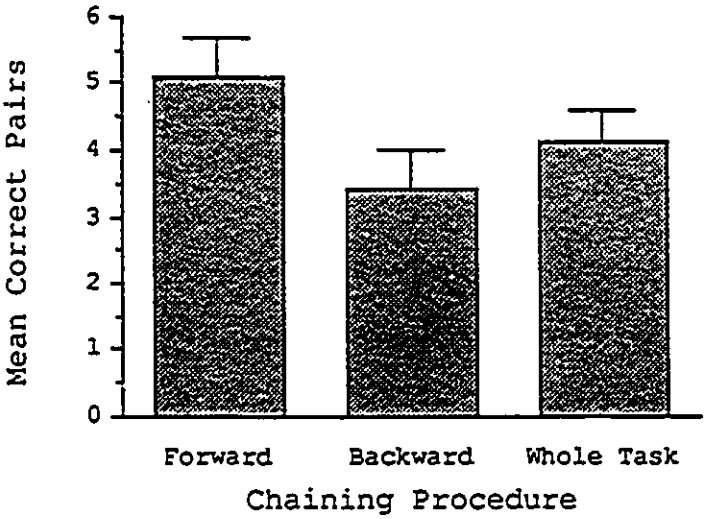


Figure 12. Mean correct pairs on Test 2 as a function of chaining procedure, Experiment 2

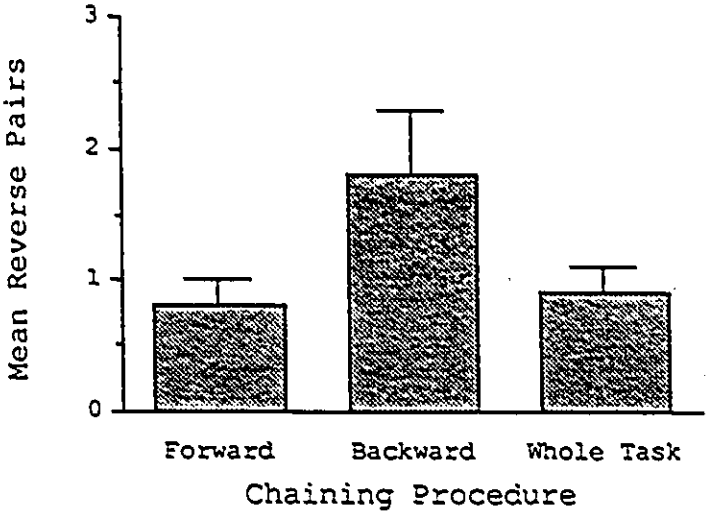


Figure 13. Mean reverse pairs on Test 2 as a function of chaining procedure, Experiment 2

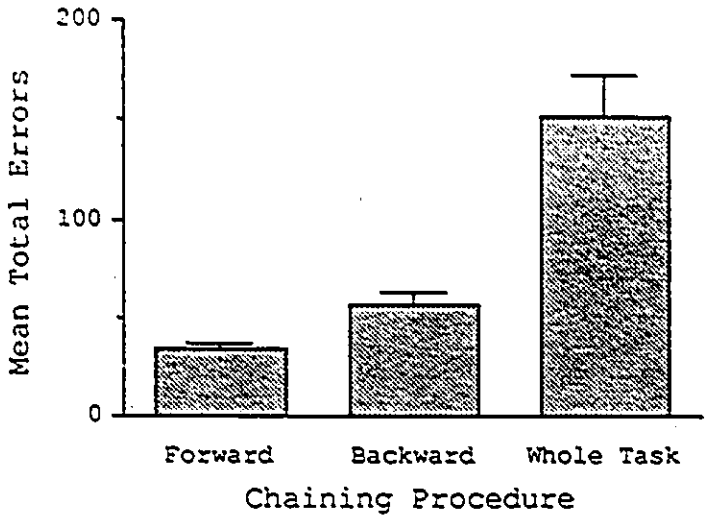


Figure 14. Mean total errors as a function of chaining procedure, Experiment 2

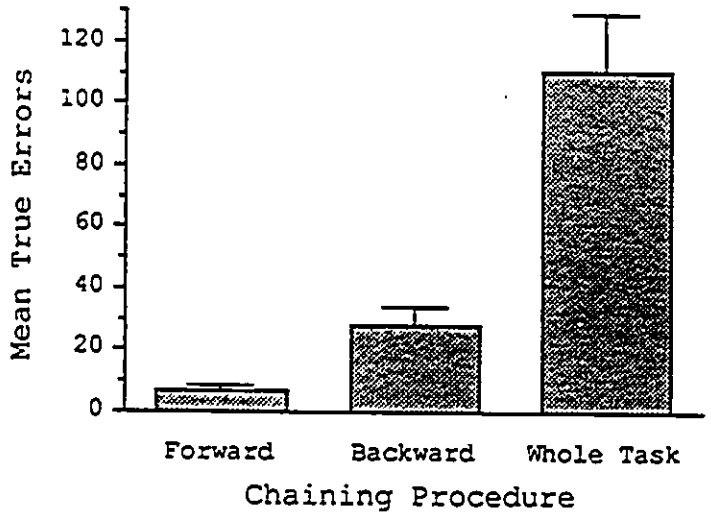


Figure 15. Mean true errors as a function of chaining procedure, Experiment 2

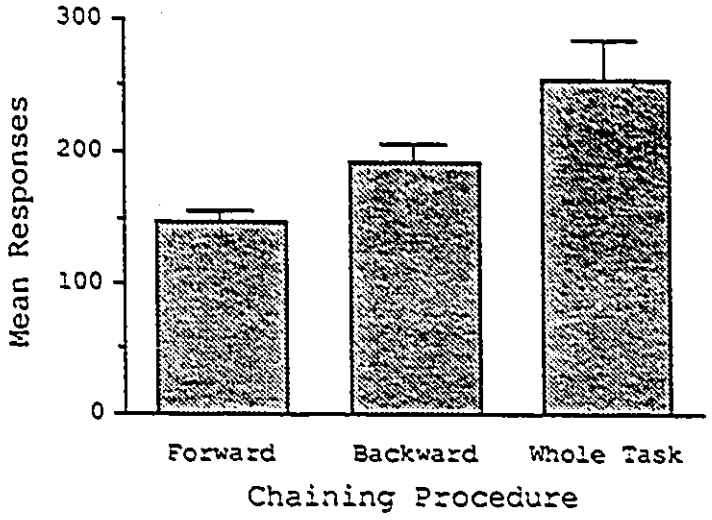


Figure 16. Mean total responses as a function of chaining procedure, Experiment 2

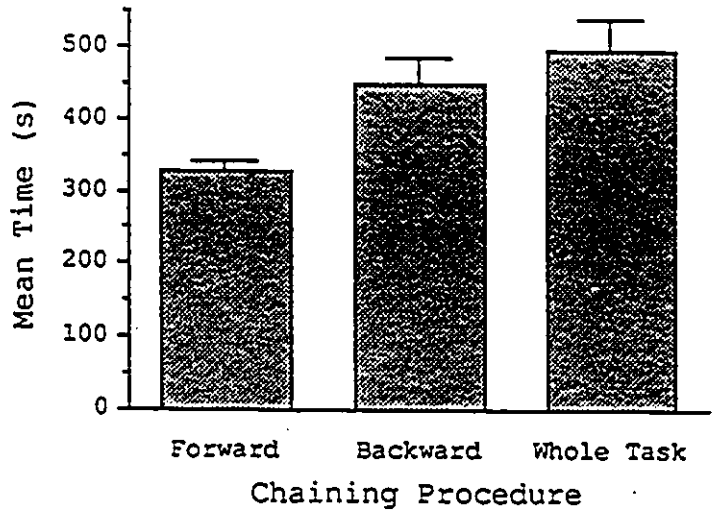


Figure 17. Mean learning time as a function of chaining procedure, Experiment 2

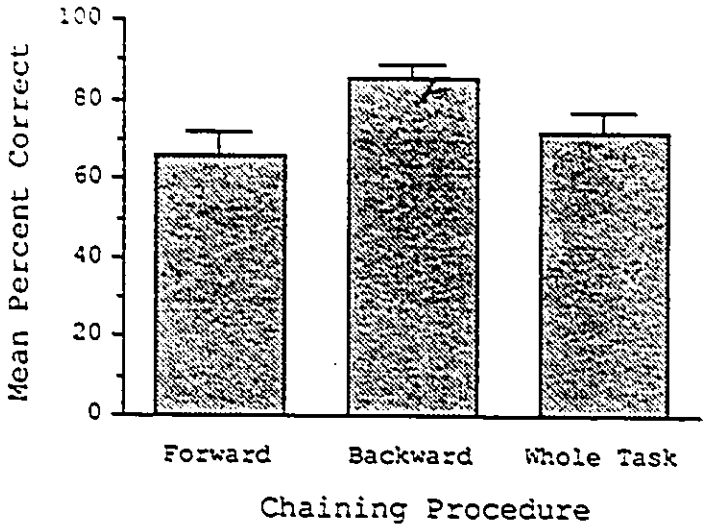


Figure 18. Mean percent correct on Test 3 as a function of chaining procedure, Experiment 3

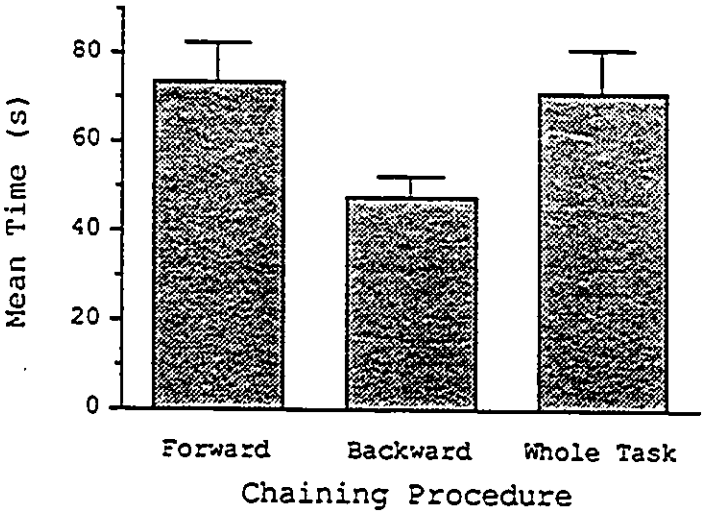


Figure 19. Mean Test 3 completion time as a function of chaining procedure, Experiment 3

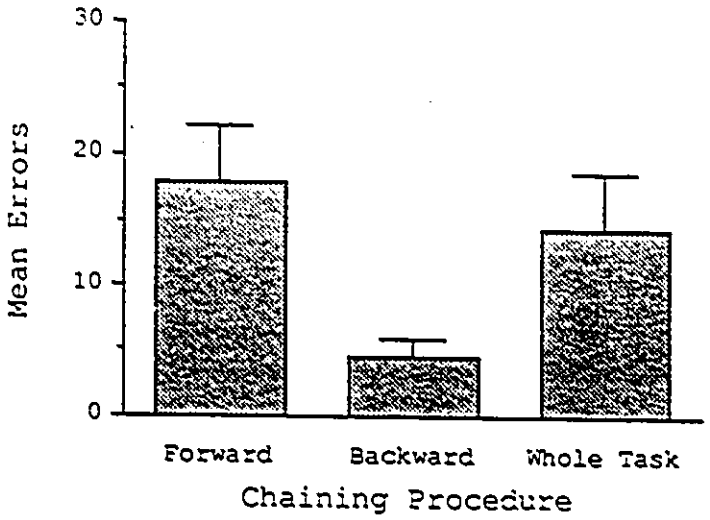


Figure 20. Mean errors on Test 3 as a function of chaining procedure, Experiment 3

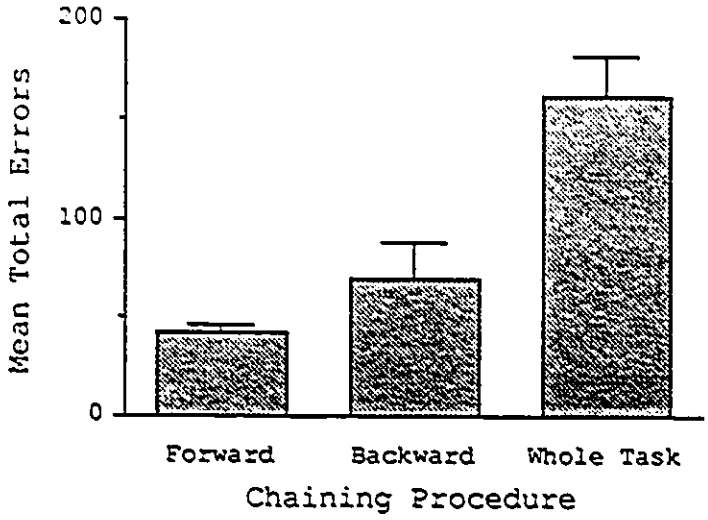


Figure 21. Mean total errors as a function of chaining procedure, Experiment 3

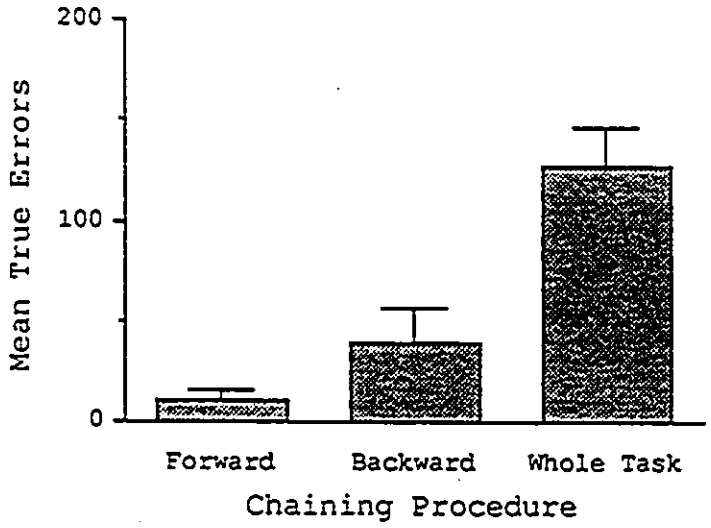


Figure 22. Mean true errors as a function of chaining procedure, Experiment 3

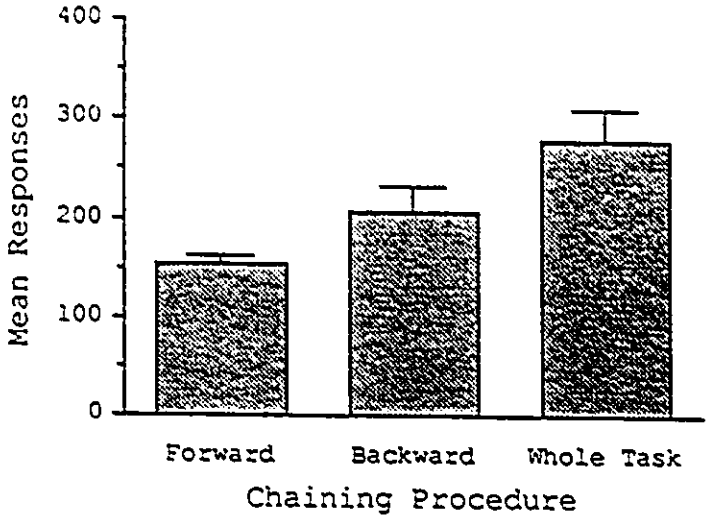


Figure 23. Mean total responses as a function of chaining procedure, Experiment 3

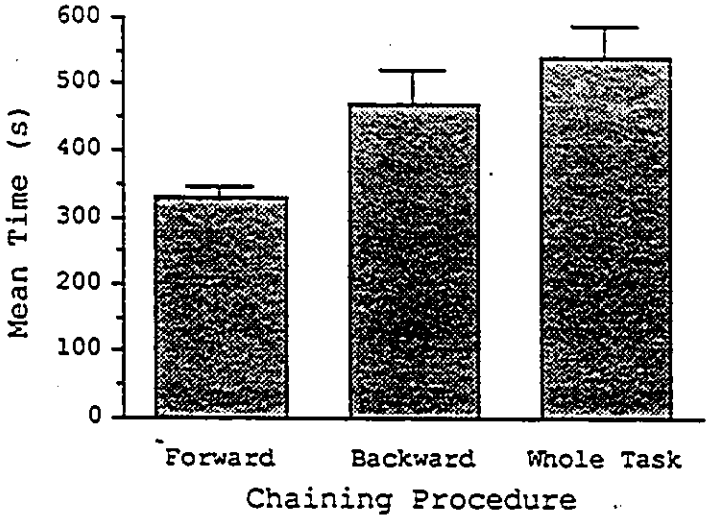


Figure 24. Mean learning time as a function of chaining procedure, Experiment 3

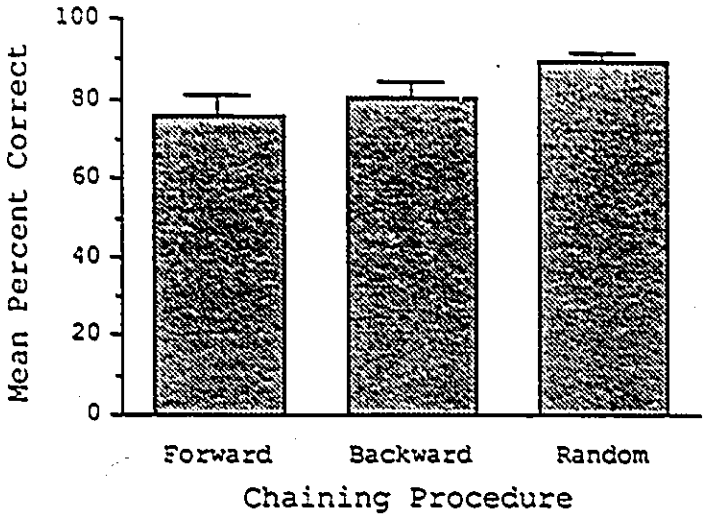


Figure 25. Mean percent correct on Test 3 as a function of chaining procedure, Experiment 4

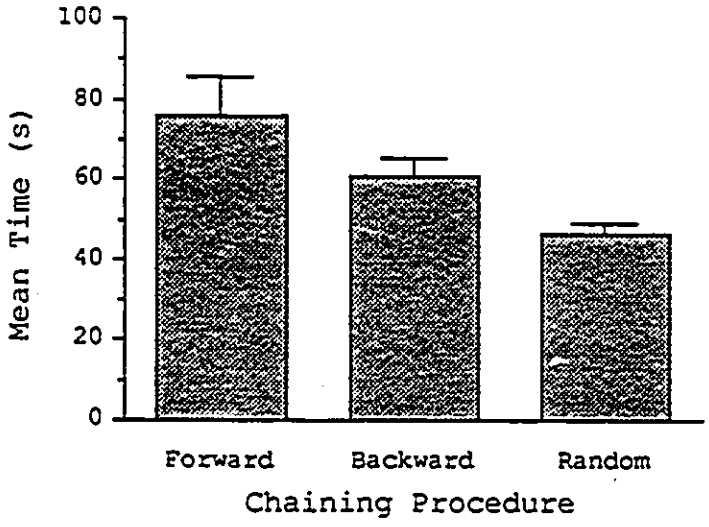


Figure 26. Mean Test 3 completion time as a function of chaining procedure, Experiment 4

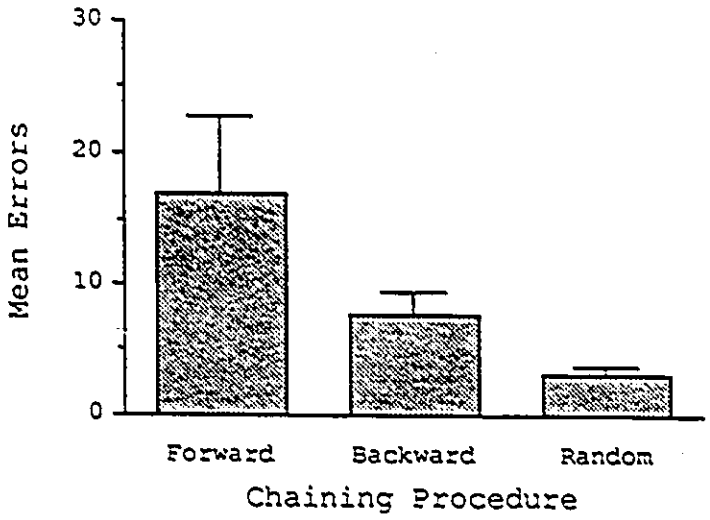


Figure 27. Mean errors on Test 3 as a function of chaining procedure, Experiment 4

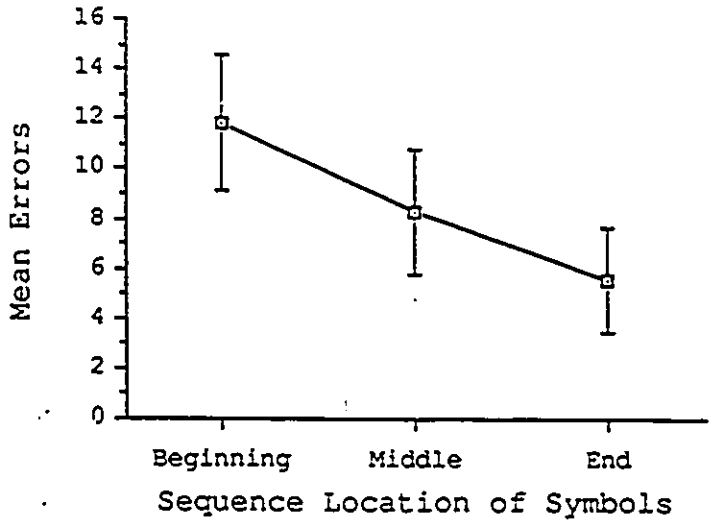


Figure 28. Mean errors on Test 3 as a function of sequence location of symbols, Experiment 4

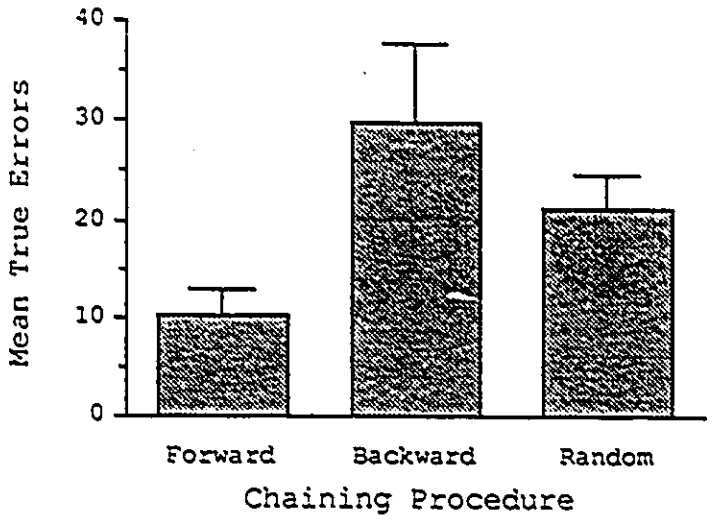


Figure 29. Mean true errors as a function of chaining procedure, Experiment 4

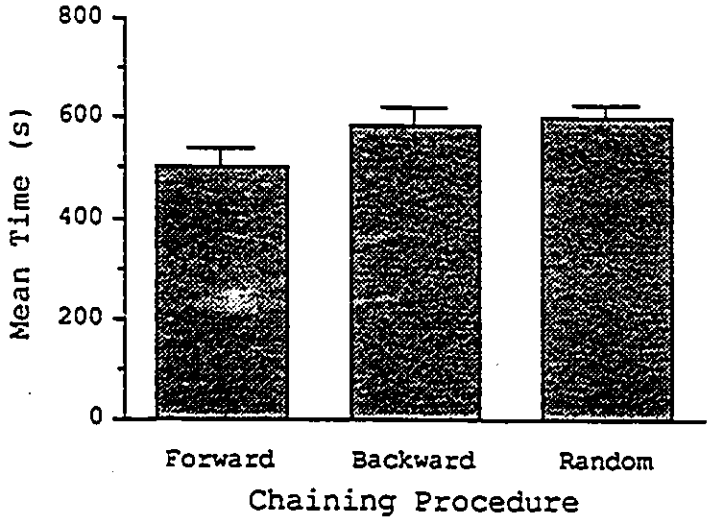


Figure 30. Mean learning time as a function of chaining procedure, Experiment 4

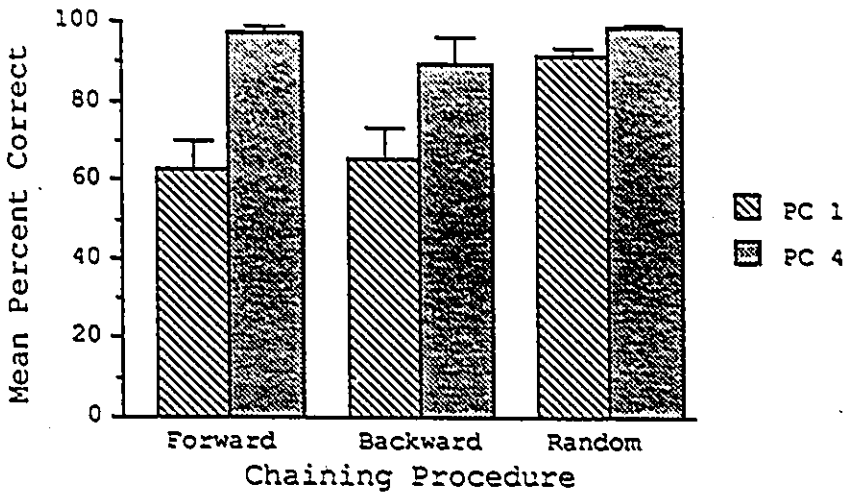


Figure 31. Mean percent correct on Test 3 as a function of chaining procedure and proficiency criterion, Order 1-4, Experiment 5

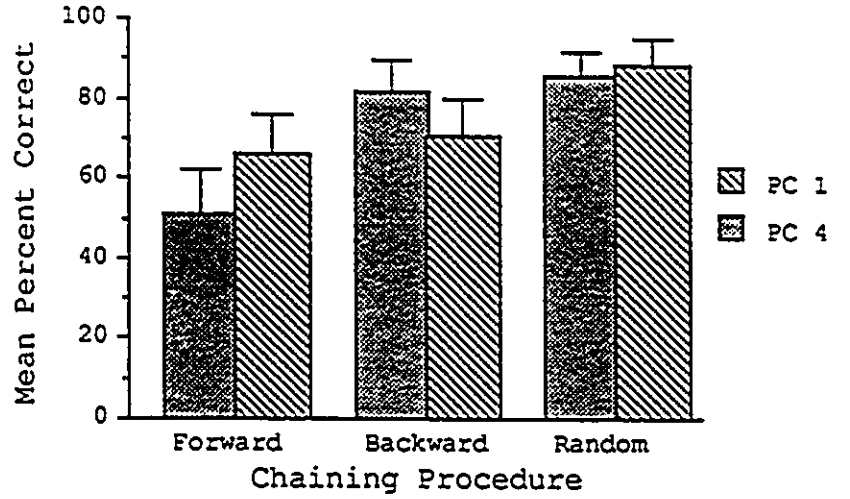


Figure 32. Mean percent correct on Test 3 as a function of chaining procedure and proficiency criterion, Order 4-1, Experiment 5

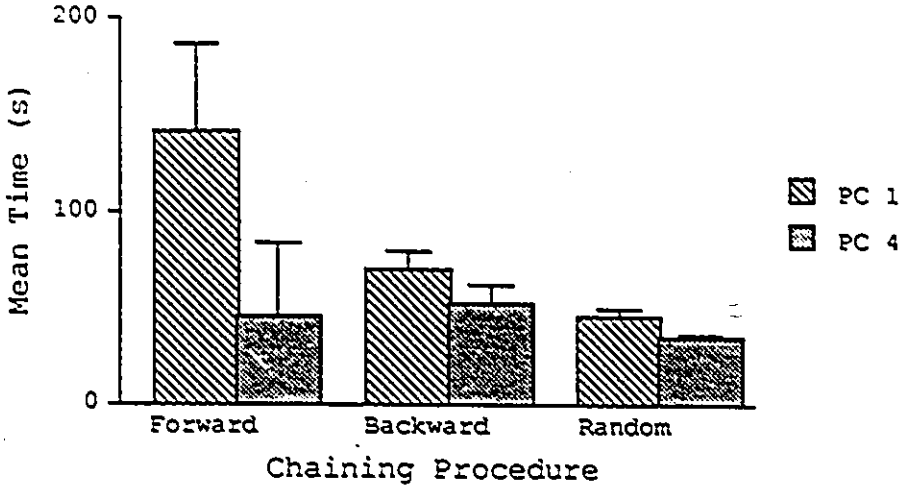


Figure 33. Mean Test 3 completion time as a function of chaining procedure and proficiency criterion, Order 1-4, Experiment 5

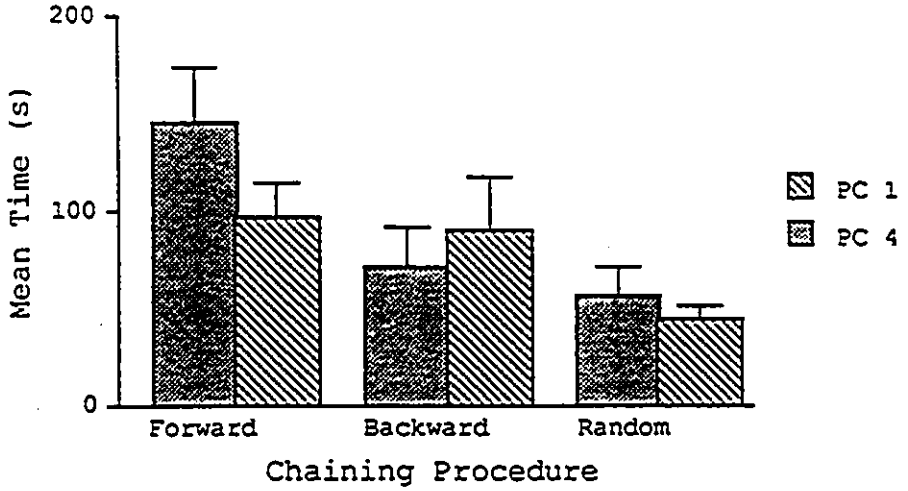


Figure 34. Mean Test 3 completion time as a function of chaining procedure and proficiency criterion, Order 4-1, Experiment 5

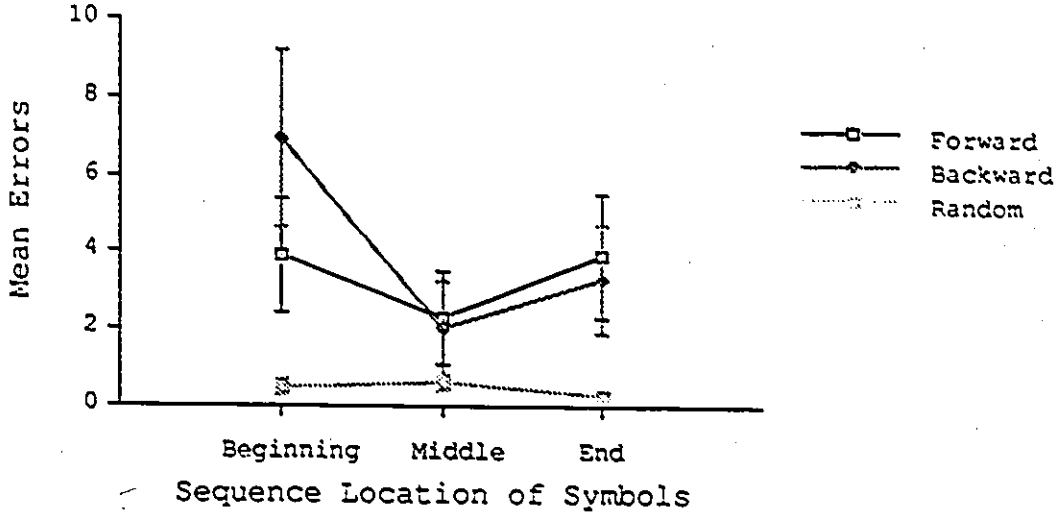


Figure 35. Mean errors on Test 3 as a function of sequence location of symbols and chaining procedure, Order 1-4, Experiment 5

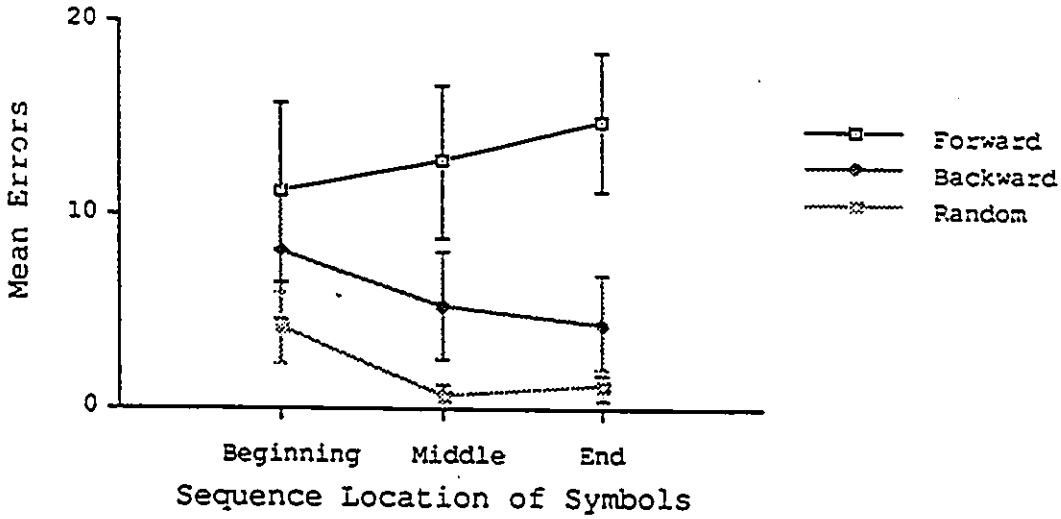


Figure 36. Mean errors on Test 3 as a function of sequence location of symbols and chaining procedure, Order 4-1, Experiment 5

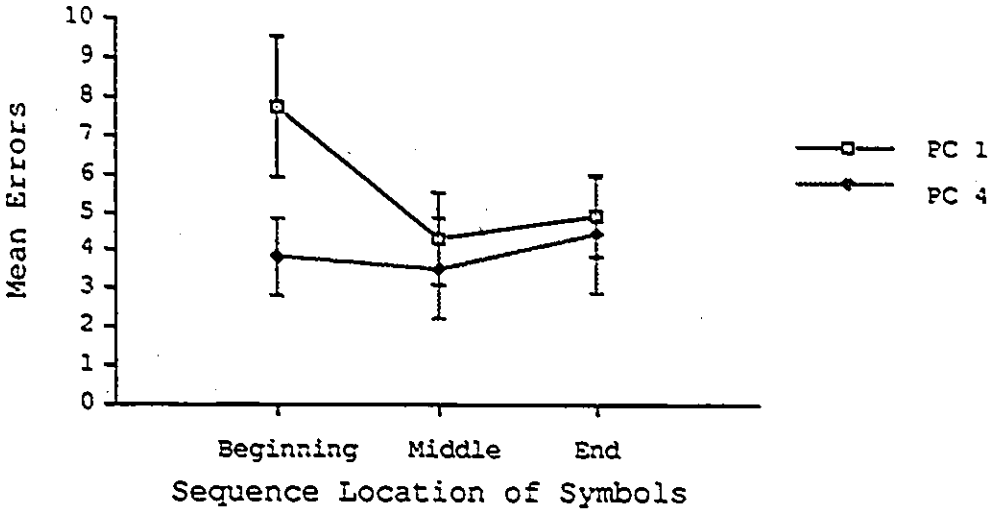


Figure 37. Mean errors on Test 3 as a function of sequence location of symbols and proficiency criterion, Experiment 5

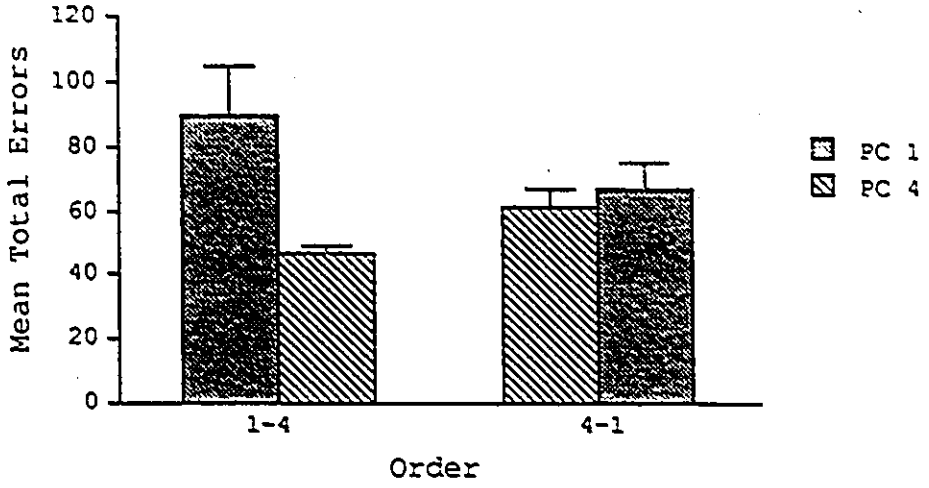


Figure 38. Mean total errors as a function of order and proficiency criterion, Experiment 5

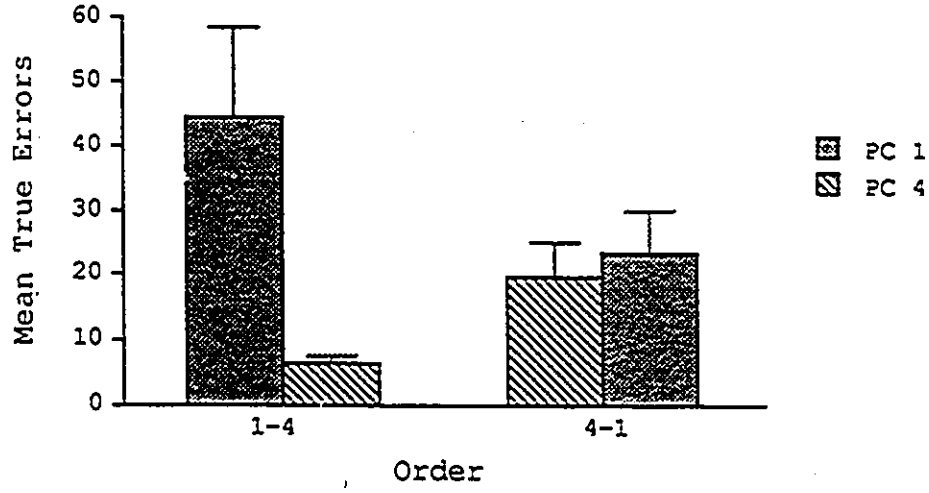


Figure 39. Mean true errors as a function of order and proficiency criterion, Experiment 5

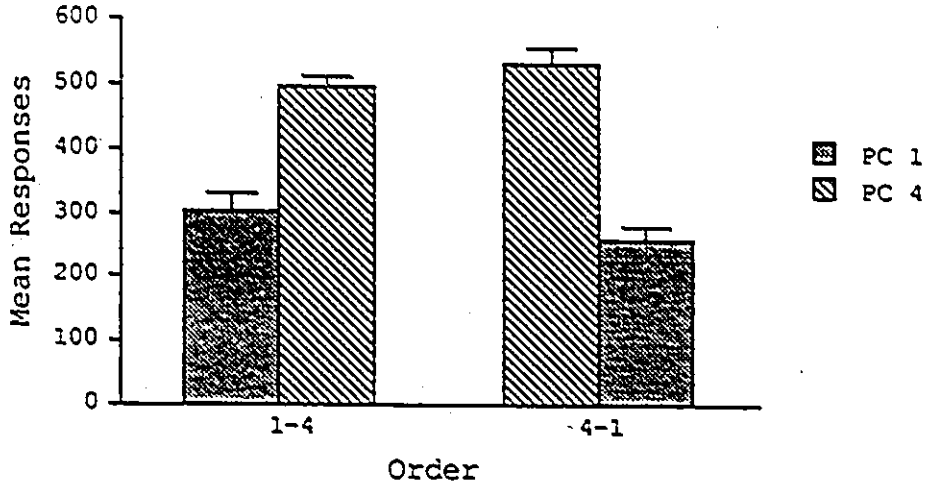


Figure 40. Mean total responses as a function of order and proficiency criterion, Experiment 5

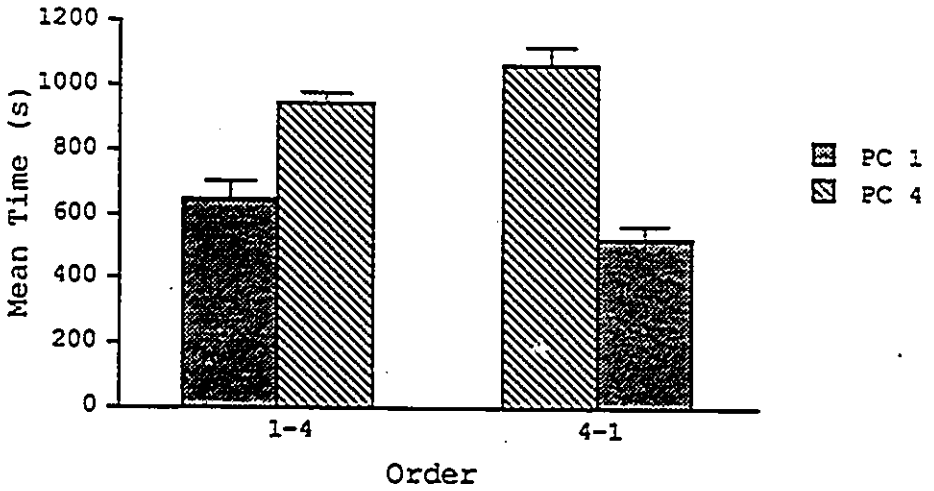


Figure 41. Mean learning time as a function of order and proficiency criterion, Experiment 5

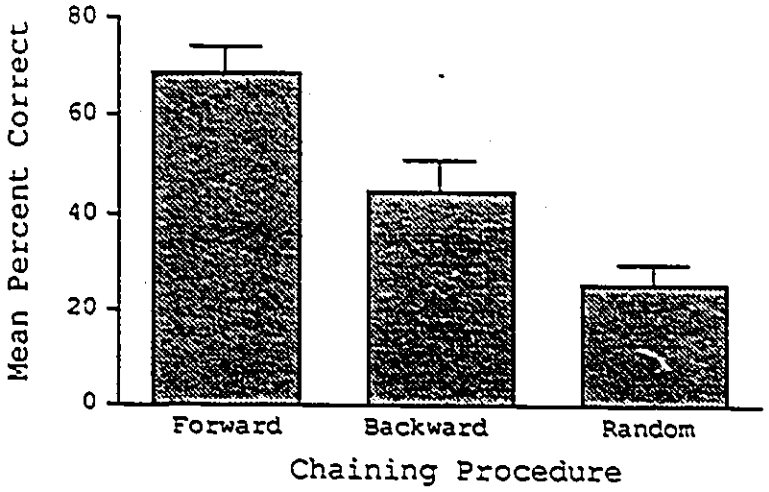


Figure 42. Mean percent correct on Test 1 as a function of chaining procedure, Experiment 6

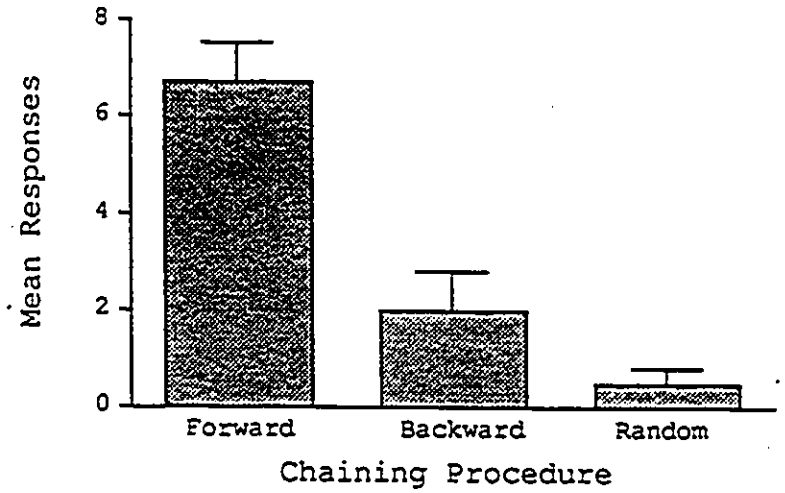


Figure 43. Mean correct responses prior to first error on Test 1 as a function of chaining procedure, Experiment 6

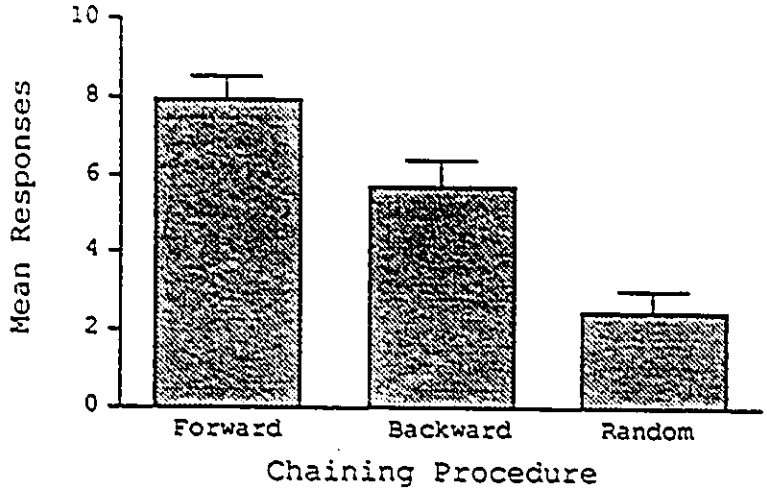


Figure 44. Mean longest series of correct responses on Test 1 as a function of chaining procedure, Experiment 6

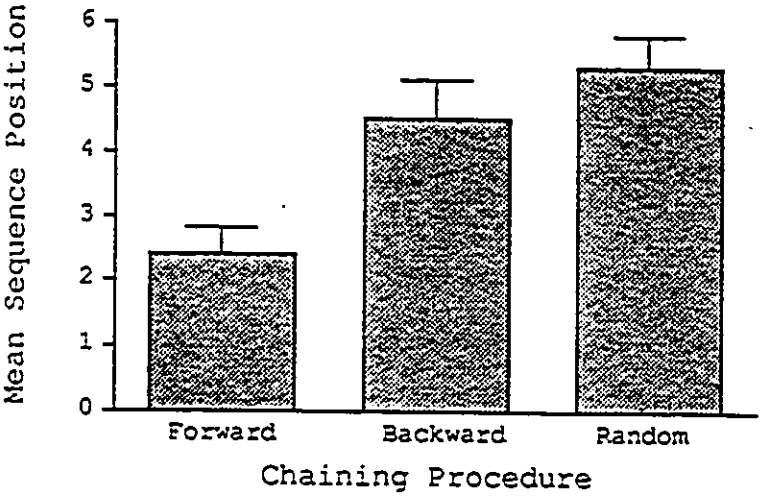


Figure 45. Mean sequence position of first response in longest series of correct responses on Test 1 as a function of chaining procedure, Experiment 6

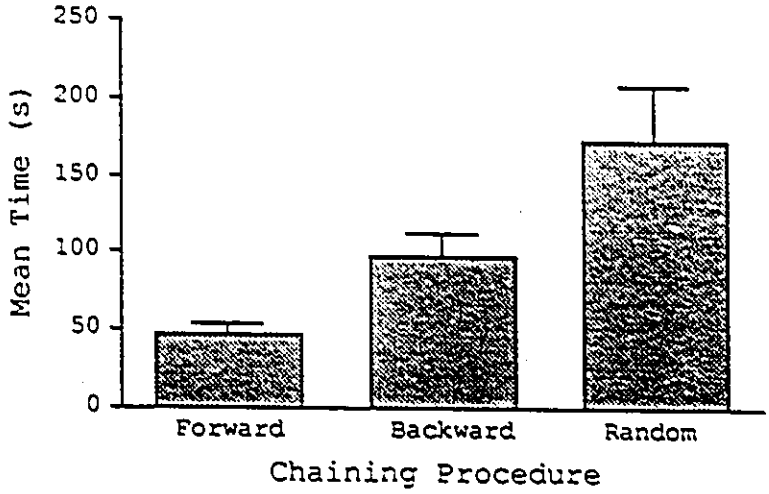


Figure 46. Mean Test 1 completion time as a function of chaining procedure, Experiment 6

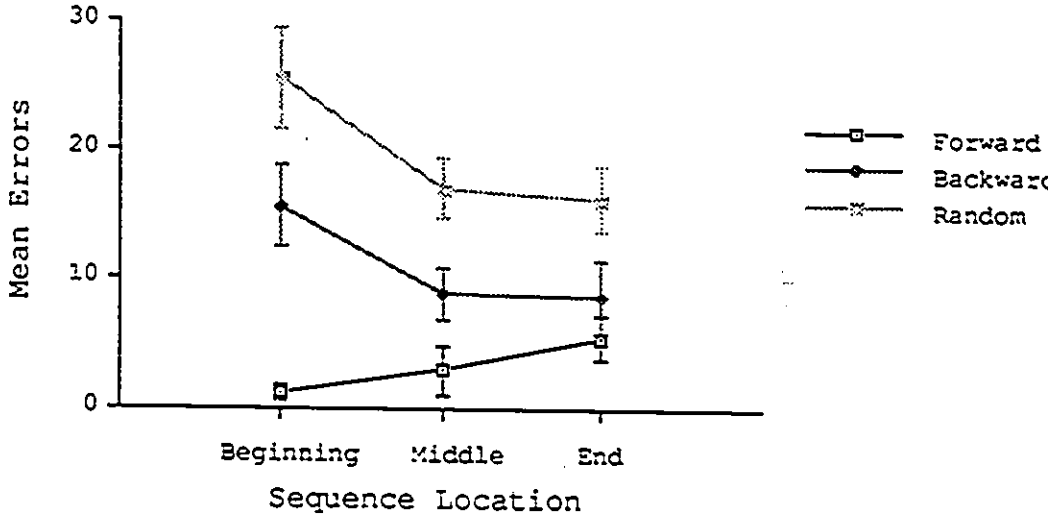


Figure 47. Mean errors on Test 1 as a function of sequence location and chaining procedure, Experiment 6

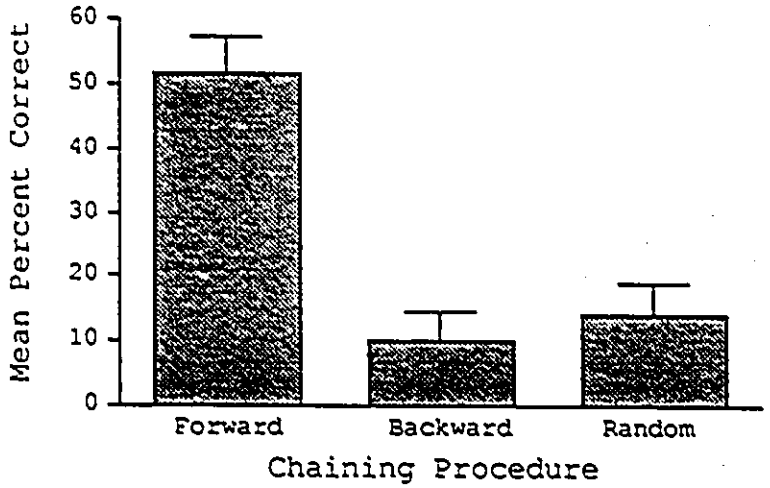


Figure 48. Mean percent correct on Test 2 as a function of chaining procedure, Experiment 6

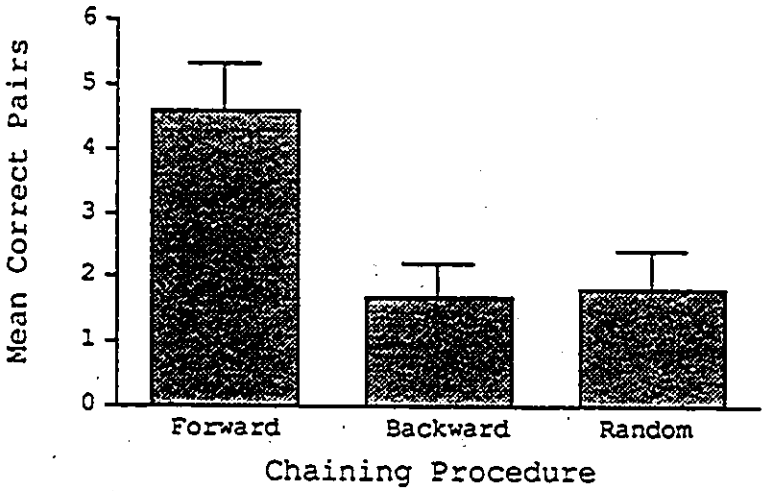


Figure 49. Mean correct pairs on Test 2 as a function of chaining procedure, Experiment 6

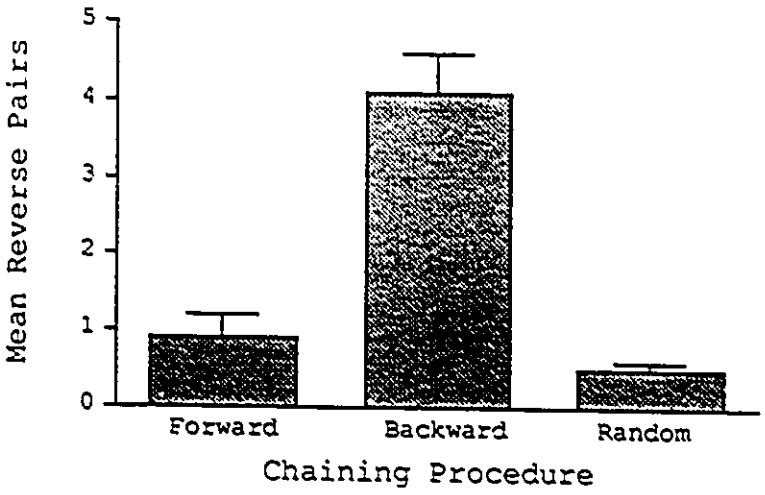


Figure 50. Mean reverse pairs on Test 2 as a function of chaining procedure, Experiment 6

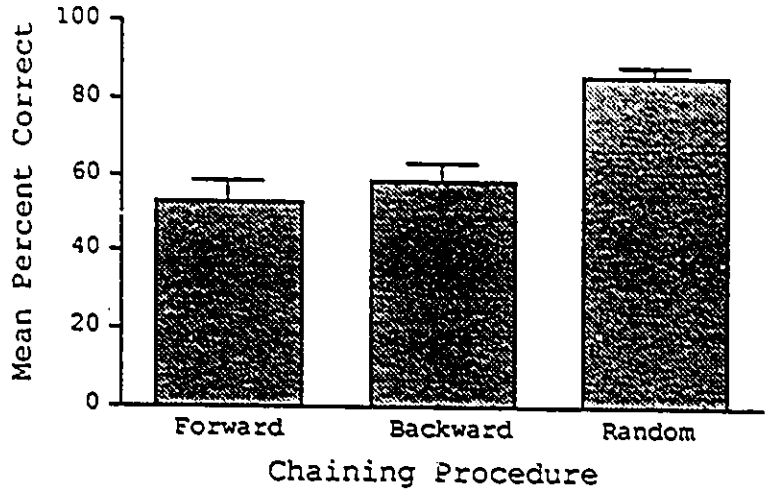


Figure 51. Mean percent correct on Test 3 as a function of chaining procedure, Experiment 6

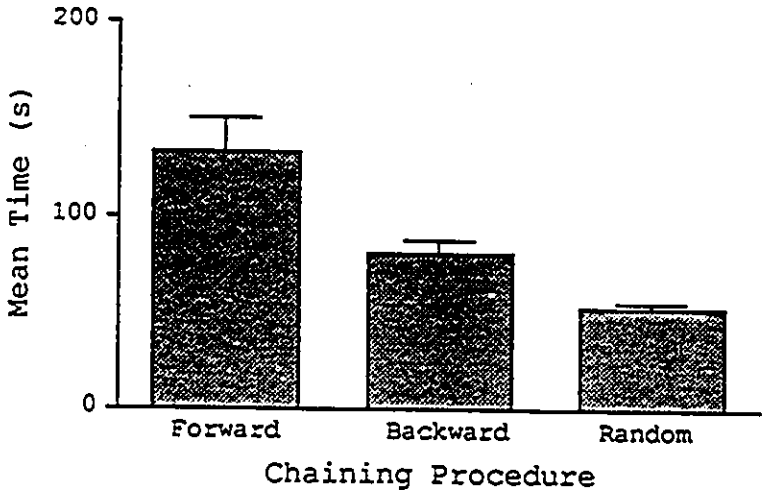


Figure 52. Mean Test 3 completion time as a function of chaining procedure, Experiment 6

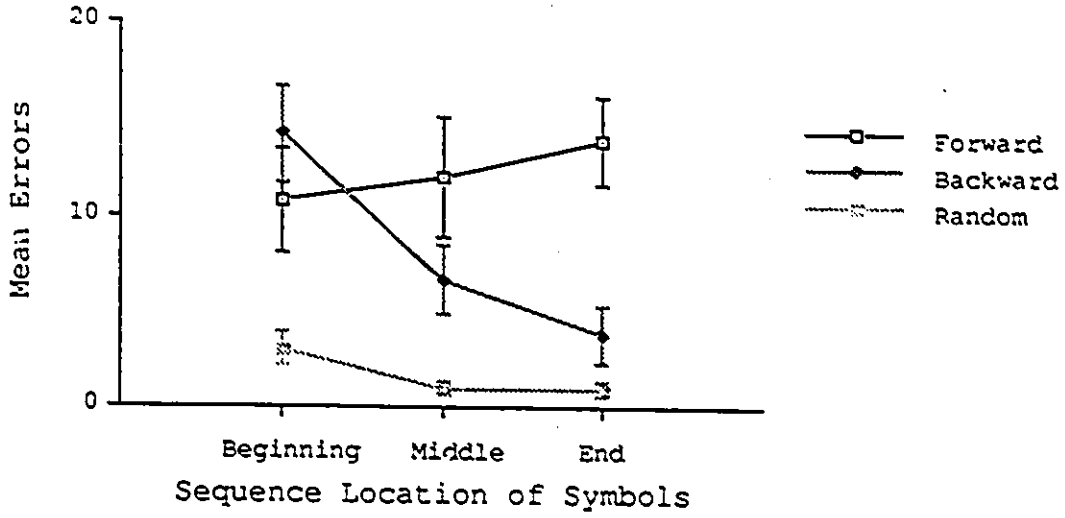


Figure 53. Mean errors on Test 3 as a function of sequence location of symbols and chaining procedure, Experiment 6

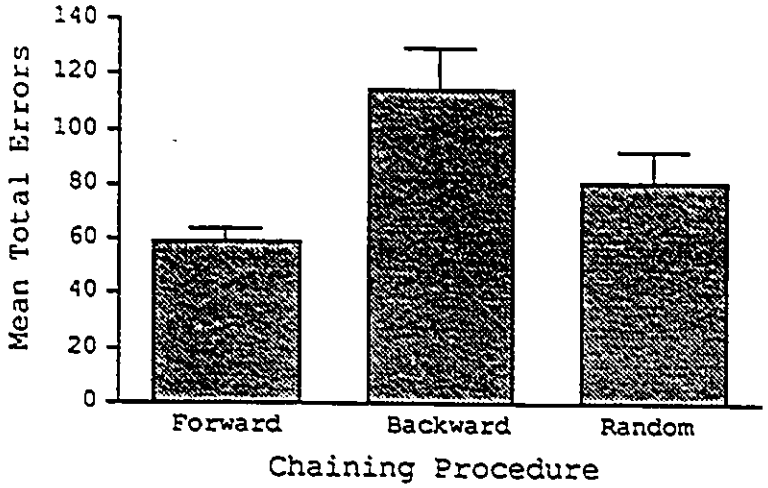


Figure 54. Mean total errors as a function of chaining procedure, Experiment 6

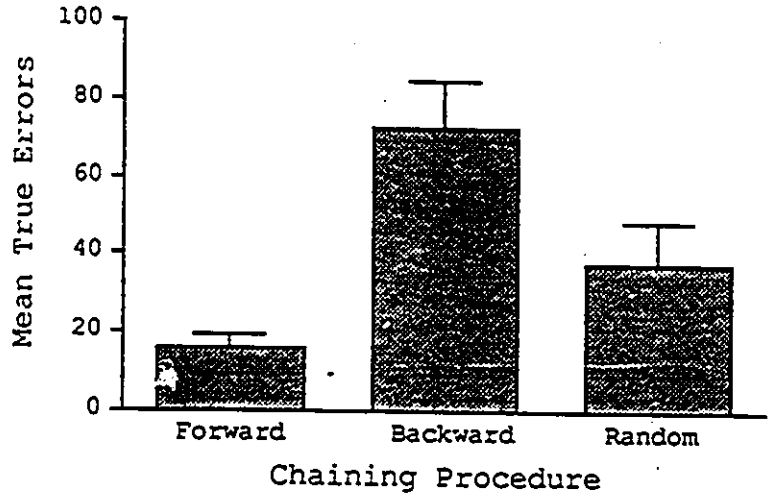


Figure 55. Mean true errors as a function of chaining procedure, Experiment 6

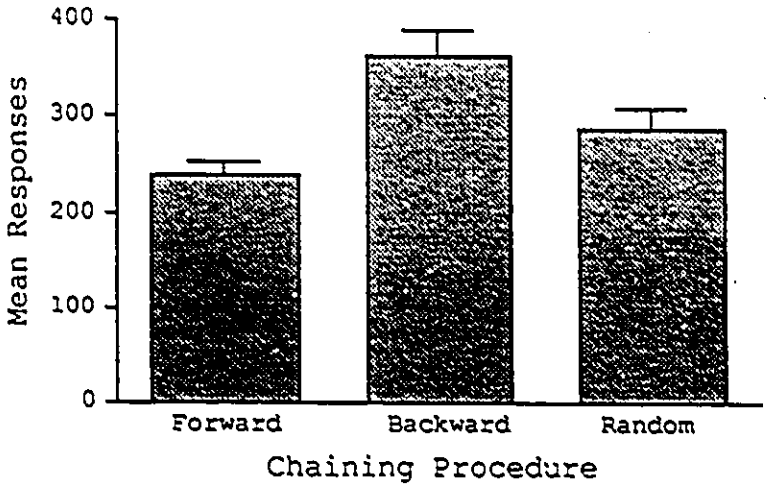


Figure 56. Mean total responses as a function of chaining procedure, Experiment 6

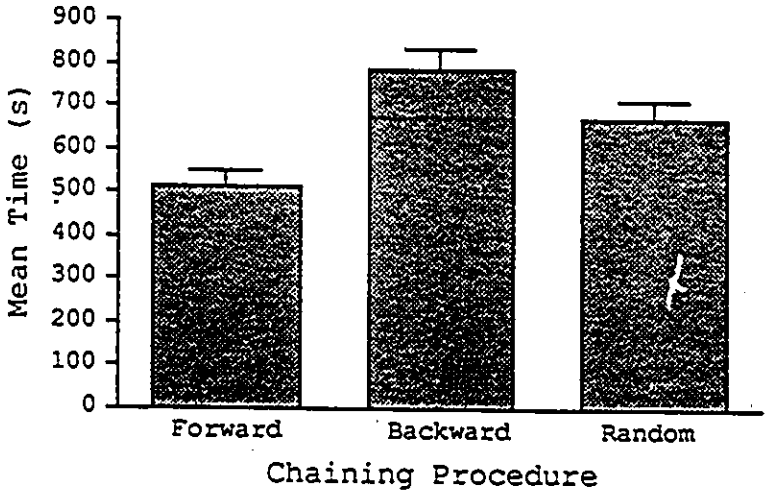


Figure 57. Mean learning time as a function of chaining procedure, Experiment 6

Appendix A

Instructions to Subjects

There is no deception involved in this experiment. It is a straightforward experiment to study the learning of sequences of responses. You will learn three separate sequences, each consisting of 10 button pushes. You will use these 10 buttons (experimenter: point to buttons). There will be a short break after you learn each sequence.

You will learn the sequences and perform them under different conditions. Usually, before you make each button push, a symbol will appear on the screen. Each symbol goes with a particular button. For example, a "?" may be the bottom left button. remembering which button goes with which symbol may help you to remember the sequence of correct button pushes. Each sequence involves a new set of symbols.

each time you are to get ready to begin a sequence, you will hear the computer make a "boing" noise. When you press the wrong button, there will be a "zit" sound and the computer will wait for you to get it right.

Periodically, your percent correct score will appear on the screen. You will learn some sequences all at once. You will learn other sequences a bit at a time, building up slowly, getting one button push correct before going on to the next, then getting two button pushes correct before going on, then getting three in a row correct, etc.

Decide which hand to use and use it throughout the experiment.

Do you have any questions?

Remember, at any time you wish, you can stop and leave (experimenter: have subject read and sign consent form).

I'll be going into this room (experimenter: point to room) while you're learning each sequence so that I won't be hanging around making you nervous. I'll come back after you learn each sequence. I'll turn off the overhead lights so you won't be bothered by reflections on the screen (experimenter: leave small lamp in corner on).

Remember, use only one hand. Also, remember that the "boing" will be your signal to get ready to perform the sequence.

Note. The wording of these subjects was changed to reflect changes in the procedures of experiments 4 to 6. For example, the sentence "You will learn some sequences all at once" was dropped for these experiments because it referred to whole task training. Similarly, the wording was changed to reflect the appropriate length of sequences and number of button pushes.

Appendix B

CONSENT FORM

I have listened to a description of the experiment and I wish to participate. From the description of the experiment I understand that:

(a) The experiment is concerned with the learning of sequences of behavior.

(b) I will be required to learn a sequence of key presses on a computer keyboard.

(c) There is no known discomfort or risk associated with the procedure.

(d) My name will not be used in connection with the analysis or reporting of the data.

(e) The session will last 20 to 40 minutes.

(f) I may, without any necessity of giving an explanation, discontinue my participation at any time and such action will not affect my class marks.

(g) A summary of the results will be available when the study is completed.

DATE: _____

PARTICIPANT'S NAME (PRINTED): _____

PARTICIPANT (SIGNATURE): _____

EXPERIMENTER (SIGNATURE): _____

Appendix C

Test 2

FOR THE LAST SEQUENCE YOU LEARNED, THESE WERE THE SCREEN SYMBOLS USED:

PLEASE NUMBER THE SYMBOLS IN THE ORDER THAT THEY APPEARED IN THE COMPLETE 10 BUTTON PRESS SEQUENCE BY WRITING A "1" UNDER THE FIRST SYMBOL IN THE SEQUENCE, A "2" UNDER THE SECOND SYMBOL IN THE SEQUENCE ETC.

Appendix D

Inferential Statistics for Experiment 1

Test 1 Variables

Multivariate $F(8, 10) = 0.6$, $p = .49$

	Univariate $F(2, 34)$	$t(17)$
Completion time	3.8, $p = .03$	
F/B		-2.4, $p = .03$
F/W		-2.6, $p = .02$
B/W		-1.1, $p = .28$

Appendix D continued

Inferential Statistics for Experiment 1

Location of Test 1 Errors

Effect	F
Chain	$F(2, 16) = 1.7$, $p = .22$
Location	$F(1, 17) = 0.2$, $p = .88$
Chain X Location	$F(2, 16) = 0.3$, $p = .74$

Appendix D continued

Inferential Statistics for Experiment 1

Test 2 Variables

Multivariate $F(6, 28) = 0.4$, $p = .89$.

Appendix D continued

Inferential Statistics for Experiment 1

Acquisition Variables

Multivariate $F(8, 10) = 47.8$, $p = .00$*

	Univariate $F(2, 34)$	$t(17)$
Total errors	26.5, $p = .00$	
F/B		-2.4, $p = .03$
F/W		-5.7, $p = .00$
B/W		-4.8, $p = .00$
True errors	22.3, $p = .00$	
F/B		-2.1, $p = .05$
F/W		-5.2, $p = .00$
B/W		-4.4, $p = .00$
Total responses	4.7, $p = .02$	
F/B		-3.4, $p = .00$
F/W		-2.6, $p = .02$
B/W		-0.9, $p = .39$
Learning time	3.4, $p = .04$	
F/B		-3.0, $p = .01$
F/W		-1.7, $p = .11$
B/W		1.0, $p = .32$

Here and throughout the appendices, $p = .00$ and $p = .000$ refer to the value provided by SPSS as $p = .000$.

Appendix E

Inferential Statistics for Experiment 2

Test 1 Variables

Multivariate $F(10, 14) = 4.3$, $p = .007$

	Univariace $F(2, 46)$	$t(23)$
Percent correct	3.2, $p = .052$	
F/B		2.1, $p = .050$
F/W		3.1, $p = .005$
B/W		0.1, $p = .910$
Correct responses prior to first error	4.4, $p = .018$	
F/B		2.3, $p = .030$
F/W		3.4, $p = .003$
B/W		-0.2, $p = .833$
Longest series of correct responses	4.5, $p = .017$	
F/B		0.8, $p = .408$
F/W		3.9, $p = .001$
B/W		1.8, $p = .080$
Beginning of longest series	2.2, $p = .124$	
Completion time	4.3, $p = .020$	
F/B		-2.7, $p = .012$
F/W		-3.8, $p = .001$
B/W		1.0, $p = .309$

Appendix E continued

Inferential Statistics for Experiment 2

Location of Test 1 Errors

Effect	F	t(23)
Chain	$F(2, 22) = 6.3$, $p = .007$	
F/B		-1.9, $p = .069$
F/W		-2.9, $p = .009$
B/W		0.8, $p = .414$
Location	$F(1, 23) = 4.0$, $p = .059$	
Chain X Location	$F(2, 22) = 2.3$, $p = .124$	

Appendix E continued

Inferential Statistics for Experiment 2

Test 2 Variables

Multivariate $F(6, 18) = 3.1$, $p = .030$

	Univariate $F(2, 46)$	$t(23)$
Percent correct	8.4, $p = .001$	
F/B		4.1, $p = .000$
F/W		1.7, $p = .103$
B/W		-2.3, $p = .028$
Correct pairs	3.2, $p = .048$	
F/B		2.6, $p = .018$
F/W		1.4, $p = .172$
B/W		-1.1, $p = .277$
Reverse pairs	3.4, $p = .040$	
F/B		-2.2, $p = .038$
F/W		-0.6, $p = .575$
B/W		1.7, $p = .099$

Appendix E continued

Inferential Statistics for Experiment 2

Acquisition Variables

Multivariate $F(8, 16) = 19.4$, $p = .000$

	Univariate $F(2, 46)$	$t(23)$
Total errors	30.4, $p = .000$	
F/B		-3.3, $p = .003$
F/W		-5.9, $p = .000$
B/W		-5.3, $p = .000$
True errors	28.0, $p = .000$	
F/B		-3.7, $p = .001$
F/W		-5.7, $p = .000$
B/W		-5.0, $p = .000$
Total responses	9.9, $p = .000$	
F/B		-2.8, $p = .010$
F/W		-4.0, $p = .001$
B/W		-2.2, $p = .038$
Learning time	8.3, $p = .001$	
F/B		-3.5, $p = .002$
F/W		-3.9, $p = .001$
B/W		-0.9, $p = .372$

Appendix F

Inferential Statistics for Experiment 3

Test 3 Variables

Multivariate $F(4, 20) = 4.4$, $p = .011$

	Univariate $F(2, 46)$	$t(23)$
Percent correct	4.8, $p = .013$	
F/B		-3.5, $p = .002$
F/W		-0.8, $p = .440$
B/W		2.2, $p = .037$
Completion time	2.9, $p = .063$	
F/B		2.6, $p = .015$
F/W		0.2, $p = .877$
B/W		-2.3, $p = .035$

Location of Test 3 Errors

Effect	F	$t(23)$
Chain	$F(2, 22) = 7.2$, $p = .004$	
F/B		3.3, $p = .003$
F/W		0.6, $p = .557$
B/W		-2.3, $p = .035$
Location	$F(1, 23) = 1.6$, $p = .214$	
Chain X Location	$F(2, 22) = 2.3$, $p = .123$	

Appendix F continued

Inferential Statistics for Experiment 3

Acquisition Variables

Multivariate $F(8, 16) = 13.2$, $p = .000$

	Univariate $F(2, 46)$	$t(23)$
Total errors	18.2, $p = .000$	
F/B		-1.6, $p = .134$
F/W		-6.7, $p = .000$
B/W		-3.6, $p = .002$
True errors	16.4, $p = .000$	
F/B		-1.6, $p = .115$
F/W		-6.4, $p = .000$
B/W		-3.3, $p = .003$
Total responses	9.3, $p = .000$	
F/B		-2.4, $p = .024$
F/W		-4.7, $p = .000$
B/W		-2.0, $p = .062$
Learning time	6.9, $p = .002$	
F/B		-3.1, $p = .005$
F/W		-4.3, $p = .000$
B/W		-0.9, $p = .362$

Appendix G

Inferential Statistics for Experiment 4

Test 3 Variables

Multivariate $F(4, 26) = 2.3$, $p = .088$

	Univariate $F(2, 58)$	$t(29)$
Percent correct	4.7, $p = .013$	
F/B		-1.1, $p = .289$
F/R		-2.7, $p = .011$
B/R		-2.2, $p = .039$
Completion time	5.7, $p = .006$	
F/B		1.6, $p = .118$
F/R		2.9, $p = .007$
B/R		2.3, $p = .031$

Location of Test 3 Errors

Effect	F	$t(29)$
Chain	$F(2, 28) = 3.8$, $p = .036$	
F/B		1.8, $p = .083$
F/R		2.4, $p = .023$
B/R		2.4, $p = .024$
Location	$F(2, 28) = 11.0$, $p = .000$	
1 to 4/5 to 8		2.2, $p = .033$
1 to 4/9 to 12		4.6, $p = .000$
5 to 8/9 to 12		2.6, $p = .015$
Chain X Location	$F(4, 26) = 1.6$, $p = .217$	

Appendix G continued

Inferential Statistics for Experiment 4

Acquisition Variables

Multivariate $F(8, 22) = 4.8$, $p = .002$

	Univariate $F(2, 58)$	$t(29)$
Total errors	2.6, $p = .079$	
F/B		-1.9, $p = .073$
F/R		-3.6, $p = .001$
B/R		0.4, $p = .731$
True errors	3.8, $p = .028$	
F/B		-2.3, $p = .027$
F/R		-3.3, $p = .002$
B/R		1.0, $p = .316$
Total responses	2.7, $p = .073$	
F/B		-1.8, $p = .081$
F/R		-2.4, $p = .021$
B/R		-0.0, $p = .978$
Learning time	3.2, $p = .047$	
F/B		-1.8, $p = .078$
F/R		-2.7, $p = .011$
B/R		-0.3, $p = .743$

Appendix H

Inferential Statistics for Experiment 5

Test 3 Variables

Effect	Multivariate F(df)	Exact p
Chaining X PC X Order	$F(4, 108) = 2.5$	.048
Chaining X PC	$F(4, 108) = 1.3$	.286
PC X Order	$F(2, 53) = 6.2$	.004
Chaining X Order	$F(4, 108) = 1.3$	.256
Chaining	$F(4, 108) = 4.3$	.003
Proficiency Criterion	$F(2, 53) = 4.6$	.015
Order	$F(2, 53) = 2.4$	.099

Effect	Univariate $F(2, 54)$	Exact p
Chaining X PC X Order		
Percent correct	3.6	.033
Completion time	4.8	.012

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t-tests for Chaining X PC X Order Interaction: Order 1-4

Chaining Group	*t*(9)	Exact *p*
Forward		
Percent correct PC 1/PC 4	-4.7	.001
Completion time PC 1/PC 4	2.1	.064
Backward		
Percent correct PC 1/PC 4	-2.3	.049
Completion time PC 1/PC 4	1.3	.238
Random		
Percent correct PC 1/PC 4	-3.4	.008
Completion time PC 1/PC 4	2.9	.017

Follow-up One-way ANOVAs for Chaining X PC X Order Interaction: Order 1-4

Variable	*F*(2, 27)	Exact *p*
Percent correct		
PC 1	6.8	.004
PC 4	1.7	.202
Completion time		
PC 1	3.2	.056
PC 4	1.9	.164

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t-tests for Chaining X PC X Order Interaction: Order 4-1

Chaining Group	$t(9)$	Exact p
Forward		
Percent correct PC 1/PC 4	1.4	.204
Completion time PC 1/PC 4	-1.7	.134
Backward		
Percent correct PC 1/PC 4	-1.4	.192
Completion time PC 1/PC 4	0.9	.420
Random		
Percent correct PC 1/PC 4	0.4	.705
Completion time PC 1/PC 4	-0.9	.391

Follow-up One-way ANOVAs for Chaining X PC X Order Interaction: Order 4-1

Variable	$F(2, 27)$	Exact p
Percent correct		
PC 1	1.8	.186
PC 4	5.0	.014
Completion time		
PC 1	2.1	.138
PC 4	4.7	.018

Appendix H continued

Inferential Statistics for Experiment 5

Location of Test 3 Errors

Effect	Multivariate F(df)	Exact p
Chaining X PC X Order X Location	$F(4, 108) = 0.6$	.699
Chaining X PC X Order	$F(2, 54) = 1.2$	.297
Chaining X PC X Location	$F(4, 108) = 0.9$	.449
Chaining X Order X Location	$F(4, 108) = 2.8$	.029
PC X Order X Location	$F(2, 53) = 0.8$	.448
Chaining X PC	$F(2, 54) = 2.6$	.085
Chaining X Order	$F(2, 54) = 2.4$	.097
Chaining X Location	$F(4, 108) = 2.0$	.105
PC X Order	$F(1. 54) = 3.0$	.088
PC X Location	$F(2, 53) = 5.0$	.011
Order X Location	$F(2, 53) = 2.4$	.105
Chaining	$F(2, 54) = 5.6$	.006
PC	$F(1, 54) = 4.1$	.049
Order	$F(1, 54) = 6.5$	.014
Location	$F(2, 53) = 1.2$	.297

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up One-way ANOVAs for Chaining X Order X Location Interaction

Sequence Location	Univariate $F(2, 27)$	Exact p
Beginning (symbols 1 to 4)		
Order 1-4	3.9	.032
Order 4-1	1.0	.377
Middle (symbols 5 to 8)		
Order 1-4	0.8	.440
Order 4-1	3.1	.062
End (symbols 9 to 12)		
Order 1-4	0.9	.412
Order 4-1	4.5	.021

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t-tests for Chaining X Order X Location Interaction: Order 1-4

Chaining Group	t(9)	Exact p
Forward		
1 to 4/5 to 8	0.8	.462
1 to 4/9 to 12	-0.0	.984
5 to 8/9 to 12	-2.4	.040
Backward		
1 to 4/5 to 8	2.4	.042
1 to 4/9 to 12	1.3	.223
5 to 8/9 to 12	-0.6	.535
Random		
1 to 4/5 to 8	-0.3	.758
1 to 4/9 to 12	0.6	.591
5 to 8/9 to 12	1.1	.299

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t-tests for Chaining X Order X Location Interaction: Order 4-1

Chaining Group	t(9)	Exact p
Forward		
1 to 4/5 to 8	-0.6	.553
1 to 4/9 to 12	-1.0	.342
5 to 8/9 to 12	-1.0	.330
Backward		
1 to 4/5 to 8	2.4	.040
1 to 4/9 to 12	3.0	.014
5 to 8/9 to 12	1.1	.303
Random		
1 to 4/5 to 8	1.9	.084
1 to 4/9 to 12	2.6	.027
5 to 8/9 to 12	-0.5	.621

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t-Tests for PC X Location Interaction

Sequence Location	$t(59)$	Exact p
Beginning (symbols 1 to 4)		
PC 1/PC 4	2.3	.027
Middle (symbols 5 to 8)		
PC 1/PC 4	0.5	.597
End (symbols 9 to 12)		
PC 1/PC 4	0.8	.412

Appendix H continued

Inferential Statistics for Experiment 5

Acquisition Variables

Effect	Multivariate F(df)	Exact p
Chaining X PC X Order	$F(8, 104) = 1.8$	.078
Chaining X PC	$F(8, 104) = 1.0$	.446
PC X Order	$F(4, 51) = 6.3$	.000
Chaining X Order	$F(8, 104) = 0.9$	.561
Chaining	$F(8, 104) = 2.2$	.030
Proficiency Criterion	$F(4, 51) = 200.2$	.000
Order	$F(4, 51) = 0.2$	.962

Effect	Univariate $F(1, 54)$	Exact p
PC X Order		
Total errors	5.2	.026
True errors	5.2	.026
Total responses	16.8	.000
Learning time	5.6	.021
Chaining		
Total errors	0.4	.679
True errors	0.4	.668
Total responses	0.9	.421
Learning time	0.4	.667

Appendix H continued

Inferential Statistics for Experiment 5

Follow-up t- Tests Between PC 1 and PC 4 for PC X Order Interaction

Variable	t(29)	Exact p
Total errors		
Order 1-4	3.0	.006
Order 4-1	0.9	.388
True errors		
Order 1-4	2.8	.010
Order 4-1	0.7	.501
Total responses		
Order 1-4	-6.4	.000
Order 4-1	-12.9	.000
Learning time		
Order 1-4	-6.8	.000
Order 4-1	-13.8	.000

Appendix I

Test 2 Experiment 6 (following page)

Subject # _____

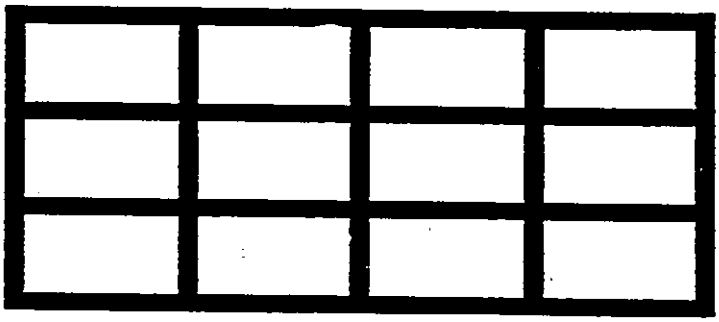
INSTRUCTION AND ANSWER SHEET

There are 3 rows of 4 symbols displayed on the monitor screen. The symbols are the 12 that were used in the sequence that you learned.

Below is a diagram with 3 rows of 4 spaces each. Look at the symbols displayed on the screen. Find the symbol that was FIRST in the sequence that you learned. Print a "1" in its space on the diagram below. Then find the symbol that was SECOND in the sequence that you learned. Print a "2" in its space on the diagram. Do this for each of the 12 symbols.

For example, if the first symbol was a "Q", and the "Q" is displayed in the top right corner of the screen, print "1" in the top right corner space of the diagram.

IF YOU DO NOT UNDERSTAND THE INSTRUCTIONS, PLEASE TELL THE EXPERIMENTER NOW.



WHEN YOU FINISH FILLING IN THE DIAGRAM, PRESS THE **BOTTOM RIGHT BUTTON** IN ORDER TO CONTINUE.

Appendix J

Inferential Statistics for Experiment 6

Test 1 Variables

Multivariate $F(10, 132) = 5.6$, $p = .000$

	Univariate $F(2, 69)$	Exact p
Percent correct	15.6	.000
Correct responses prior to first error	22.6	.000
Longest series of correct responses	21.9	.000
Beginning of longest series	9.6	.000
Completion time	8.3	.001

Appendix J continued

Inferential Statistics for Experiment 6

Location of Test 1 Errors

Effect	Multivariate *F*(df)	Exact *p*
Chaining X Location	*F*(4, 138) = 2.7	.031
Chaining	*F*(2, 69) = 16.9	.000
Location	*F*(2, 68) = 4.7	.013

Follow-up Dependent Pairs t-tests for Chaining X Location Interaction

Chaining Group	*t*(23)	Exact *p*
Forward		
1 to 4/5 to 8	-0.9	.383
1 to 4/9 to 12	-2.3	.033
5 to 8/9 to 12	-0.9	.375
Backward		
1 to 4/5 to 8	2.2	.038
1 to 4/9 to 12	1.9	.069
5 to 8/9 to 12	0.1	.942
Random		
1 to 4/5 to 8	3.6	.001
1 to 4/9 to 12	2.6	.017
5 to 8/9 to 12	0.5	.652

Appendix J continued

Inferential Statistics for Experiment 6

Test 2 Variables

Multivariate $F(6, 136) = 18.7$, $p = .000$

	Univariate $F(2, 69)$	Exact p
Percent correct	23.0	.000
Correct pairs	8.4	.001
Reverse pairs	32.5	.000

Appendix J continued

Inferential Statistics for Experiment 6

Test 3 Variables

Multivariate $F(4, 138) = 10.5$, $p = .000$

	Univariate $F(2, 69)$	Exact p
Percent correct	15.9	.000
Completion time	12.5	.000
Reverse pairs	32.5	.000

Appendix J continued

Inferential Statistics for Experiment 6

Location of Test 3 Errors

Effect	Multivariate *F*(df)	Exact *p*
Chaining X Location	*F*(4, 138) = 4.9	.001
Chaining	*F*(2, 69) = 11.2	.000
Location	*F*(2, 68) = 3.9	.025

Follow-up dependent Pairs t-tests for Chaining X Location Interaction

Chaining Group	*t*(23)	Exact *p*
Forward		
1 to 4/5 to 8	-0.6	.590
1 to 4/9 to 12	-1.4	.189
5 to 8/9 to 12	-0.6	.539
Backward		
1 to 4/5 to 8	2.7	.014
1 to 4/9 to 12	4.2	.000
5 to 8/9 to 12	1.7	.096
Random		
1 to 4/5 to 8	2.6	.015
1 to 4/9 to 12	2.2	.043
5 to 8/9 to 12	0.0	1.000

Appendix J continued

Inferential Statistics for Experiment 6

Acquisition Variables

Multivariate $F(8, 134) = 3.9$, $p = .000$

	Univariate $F(2, 69)$	Exact p
Total errors	6.7	.002
True errors	8.8	.000
Total responses	8.3	.001
Learning time	10.5	.000