

Attention and Structure in Sequence Learning

Asher Cohen, Richard I. Ivry, and Steven W. Keele
University of Oregon

In this study we investigated the role of attention, sequence structure, and effector specificity in learning a structured sequence of actions. Experiment 1 demonstrated that simple structured sequences can be learned in the presence of attentional distraction. The learning is unaffected by variation in distractor task difficulty, and subjects appear unaware of the structure. The structured sequence knowledge transfers from finger production to arm production (Experiment 2), suggesting that sequence specification resides in an effector-independent system. Experiments 3 and 4 demonstrated that only structures with at least some unique associations (e.g., any association in Structure 15243... or 4 to 3 in Structure 143132...) can be learned under attentional distraction. Structures with all items repeated in different orders in different parts of the structure (e.g., Sequence 132312...) require attention for learning. Such structures may require hierarchic representation, the construction of which takes attention.

One of the remarkable capabilities of humans is their ability to learn a variety of novel tasks involving complex motor sequences. They learn to play the violin, knit, serve tennis balls, and perform a variety of language tasks such as speaking, typing, writing, or producing sign. This study addresses three features that might be involved in such learning: attention, structure of the sequence, and effector specificity. These three features will be discussed in succession.

Attention and Sequence Learning

A large variety of evidence indicates that attention is important in verbal learning. For example, the classic study by Peterson and Peterson (1959) showed that a numeric distractor produced a dramatic loss of recall of short letter strings. Similarly, Fisk and Schneider (1984) found judgment of frequency of previously presented words to drop to chance level when the words were presented concurrently with a numeric distractor. The learning was prevented even though the secondary numeric task was very different from the frequency judgment task. On the basis of these findings, Fisk and Schneider argued that general attentional resources are necessary for modifications of long-term memory.

Does the learning of motor sequences also require attention? This question is especially relevant in light of the hypothesis that sequential learning can involve a different memory system, sometimes called *procedural memory*, than verbal learning or other declarative memory systems (cf. Mishkin & Appenzeller, 1987; Mishkin, Malamut, & Bachevalier, 1984; Schacter, 1987; Squire, 1986). A study by Nissen and Bullemer (1987), after which our own was modeled, investigated the role of attention in sequence learning. Subjects in their study performed a tapping task in which they pressed one of

four keys in response to an asterisk at one of four spatial positions. In one condition, the signals came on in a particular sequence of 10 events, with the same order repeating cyclically (hereafter, structured sequence). In another condition the locations were randomly selected (hereafter, random sequence). Reaction times became much faster with practice in the structured sequence condition, indicating that learning had occurred. However, when a secondary tone-counting task was performed concurrently with the reaction time task, learning of the structured sequence failed to occur. Reaction times to the structured sequence stimuli were no faster than to random sequence stimuli. Learning was prevented even though the secondary tone task is very different from the tapping task and is processed by a different modality. Nissen and Bullemer (1987) concluded that attention, in the sense of freedom from distraction, is necessary for the structured sequence to be learned.

In another experiment Nissen and Bullemer (1987) showed that Korsakoff patients were able to learn the structured sequence in the tapping task. Korsakoff patients have impaired declarative memory, and consequently they were unaware of the presence of the structured sequence. Moreover, when Nissen, Knopman, and Schacter (1987) administered scopolamine to healthy subjects, performance was reduced on a verbal memory task but not on the sequence learning task. Despite their preserved sequence learning, the subjects receiving scopolamine performed poorly on a task designed to assess awareness of the structured sequence. Nissen and her colleagues drew two important conclusions from these studies: First, awareness and attention need not be equated: Only the latter is required for sequential learning to occur. Second, the results indicate a dissociation between neural systems responsible for structured sequence learning from those responsible for declarative memory.

Sequence Structure

Although attention is necessary to learn the Nissen and Bullemer structured sequences, it is possible that other structured sequences can be acquired without attention. Consider

This research was supported by a contract (N00014-87-K-0279) from the Office of Naval Research to Steven Keele and Richard Ivry. Richard Ivry is now at the University of California, Santa Barbara.

Correspondence concerning this article should be addressed to Steven Keele, Department Psychology, University of Oregon, Eugene, Oregon 97403.

Nissen and Bullemer's sequence. If the four spatial asterisk positions from left to right are labeled 1 through 4, their structured sequence can be symbolized as 4231324321. Each location repeats at least once in a cycle. Moreover, different stimuli follow each repetition. For example, the asterisk is presented three times in Position 2 and is followed by Positions 3, 4, and 1 on successive presentations. The inclusion of different pairwise orders within the overall sequence poses problems for a mechanism of sequential learning based on serial associations (Keele & Summers, 1976). A representation not based solely on linear associations might be needed in order to learn structured sequences that contain pairwise ambiguities.

Hierarchical coding provides one alternative to the problem of learning sequences with repeated items. Substantial evidence argues that humans often represent sequences by hierarchic structures (see Keele, 1981, 1986, and Jordan & Rosenbaum, in press, for reviews). In a hierarchic representation, a sequence is coded as a series of groups, with each group being more finely divided at a lower level. Such coding is sometimes revealed by the pauses between subgroups of a motor sequence (e.g., Povel & Collard, 1982; Rosenbaum, Kenny, & Derr, 1983). Other phenomena also indicate that the sequence representation is hierarchically organized. These involve the reaction times to prepare movements for execution (e.g., Klapp, Anderson, & Berrian, 1973; Rosenbaum, Inhoff, & Gordon, 1984; Sternberg, Monsell, Knoll, & Wright, 1978) and patterns of errors and transfer of performance from one sequence to another (e.g., Gordon & Meyer, 1987; MacKay, 1982).

Structured sequences with repeated items can be learned with hierarchic coding. When a particular element is followed by one event in one case and another event in another case, a plan provided by a higher level in the hierarchy provides an additional cue to ensure the correct ordering. An example of how such an additional cue can help is provided by the connectionist model of sequencing of Jordan (1986a, 1986b). Because many human skills involve the same elementary movements assembled into different sequences, one might speculate that a powerful system of hierarchic coding underlies the remarkable ability of humans to learn sequences.

A central issue addressed in this article is the relation between attention and the nature of the sequence structure. Can structured sequences in which at least one element is *uniquely* associated with another element (e.g., Element a is *always* followed by Element b) be learned even when attention is distracted as argued by Posner (1984)? Conversely, is attention a prerequisite for learning structured sequences in which all of the elements are repeated with different pairwise associations?

Effector Independence

A second issue to which these studies are directed concerns whether the knowledge of a structured sequence is encoded by the effector that produced the sequence or whether the sequence knowledge is independent of the effector system.

It long has been known that animals and humans often code their movements with reference to external space rather

than to particular muscles. Rats in learning a maze learn the spatial layout of their environment (Tolman, 1948). Conditioned responses of the finger to shock change from finger extension to flexion when the hand is turned over (Wickens, 1943). Subjects code tactual stimulation to fingers that are placed at particular locations not in terms of finger stimulated but in terms of location at which the finger was placed (Attneave & Benson, 1969). The association between stimulus position and response in a reaction time setting is between stimulus and response key, not stimulus and responding finger, because compatibility effects remain unchanged when the hands are crossed (Wallace, 1971).

Although such results suggest that response coding is in terms of space and not effector, such may not be the case for rapid motor sequences. Nonetheless, a similar conclusion appears to emerge. One's writing looks similar whether produced by hand, wrist, forearm, elbow, and even by foot or mouth (Bernstein, 1947, as reported in Keele, Cohen, & Ivry, in press; Raibert, 1977), suggesting that sequences of action are encoded independently of the effector. Studies of neurologically impaired patients also suggest that the code that represents sequences of graphemes in handwriting is shared with oral spelling (Margolin, 1984; Hillis & Caramazza, 1988). Such analyses, however, do not rule out the possibility that some structured sequence representation is in fact tied to particular effectors (see especially Wright, in press). Moreover, the performance with unpracticed effectors and in neurologically impaired patients is often very slow. Possibly such slowness reflects degraded sequence knowledge specific to handwriting.

One study to address whether sequence knowledge is independent of effectors was conducted by Pew (1974). Unknown to subjects, one segment of an apparently random visual track repeated on each practice trial over many sessions. Subjects, though never expressing awareness of the repeated segment, improved on it more than on nonrepeated segments. Part, but not all, of that improvement transferred when the repeated segment of the visual track appeared in mirror image. Such mirroring required a different sequence of muscles—what had been agonist became antagonist, and vice versa. The result is ambiguous, however. Does lack of complete transfer indicate that some sequence knowledge is effector specific, or does it indicate that putting the track in a mirror image simply changes the spatial array?

The present study used a transfer paradigm from fingers to arm to assess whether sequence knowledge gained in the course of practice with one effector transfers to another. The procedure allowed a quantitative assessment of how much sequence knowledge is effector independent. In addition, the structured sequence was learned under dual-task conditions and without awareness. Such a procedure allowed a determination of whether the sequence knowledge that exists outside declarative knowledge is effector independent.

Experiment 1

This experiment examined whether a simple sequence of events can be learned in the presence of an attentional distractor and whether the amount of distraction has any influ-

ence on the amount of learning. The main task involved key pressing to *X* marks which appeared at one of three positions on a computer monitor. Three keys corresponding to the three positions were used for response. Subjects were instructed to respond as fast as they could to the *X* mark with the key corresponding to its spatial position. The *X* disappeared from the screen with the subject's response, and another *X* appeared 200 ms later. In one condition the visual signals occurred in a five-element structured sequence that repeated cyclically with no noticeable break between cycles. Symbolizing the positions from left to right as 1 through 3, one structure used was 32121 In the other condition, the signals occurred randomly at the three positions, with the constraint that the signal could not appear successively at the same position. Learning was indexed by the degree to which reaction times for subjects in the structured sequence condition became faster with practice than the reaction times of subjects in the random control group.

The distraction task was similar to that used by Nissen and Bullemer (1987). A tone was presented in each interval between a response to one stimulus and the onset of the next stimulus. The tone could be either high pitched or low pitched. After a block of trials, the subject had to report the number of high-pitched tones.

If learning occurred under distraction, the question might be raised whether learning was due to residual attention. To address this issue, the level of difficulty of the secondary tone task was varied. Pilot work had indicated that simply increasing the proportion of target high-pitched tones on the secondary task increased reaction time on the primary tapping task. This is a clear indication that the more difficult tone condition demands more attention. If it is the case that attention is necessary for learning the structured sequence, then there should be less learning of the structured sequence in the tapping task in the presence of the more difficult tone task.

Another issue concerns whether the tone task is sufficiently demanding to prevent awareness of the structured sequence being learned. This issue was assessed at the completion of the dual-task blocks by altering the primary task from one of responding to the current signal to predicting the next one (Nissen & Bullemer, 1987). For example, if the structured sequence was 32121, the first signal presented was the *X* in Position 3. A subject's task was to press the key corresponding to the signal he or she thought would follow 3, a correct response being Key 2. Following the press, and regardless of the subject's response, the *X* moved to the correct location. This provided both feedback (because it told the subject the correct response of the previous trial) and served as the next stimulus. The subject again had to predict the position of the next stimulus and so on. To assess the level of awareness, a comparison can be made between subjects in the structured sequence conditions and subjects in the random sequence conditions. The random groups had no opportunity to learn the structure during the initial tapping task. However, during the generation phase, the signals for all of the subjects occurred in a structured order. Thus, it could be determined whether either the initial level of generation performance or the rate of learning the sequence in the generation phase was better for the structured sequence groups, which had already practiced the structure, than for the random groups.

It should be emphasized that there is a fundamental difference between the tapping task used in the first part of the experiment and the generation task used in the second part. The learning of the structured sequence in the tapping task was *implicit*: It could be inferred only from the reaction time of the subjects. In the generation task the subjects were asked to *explicitly* produce the structured sequence. Amnesic patients typically show marked deficits in tasks requiring explicit knowledge acquisition and very little deficit in tasks requiring implicit knowledge acquisition (e.g., Squire, 1986).

The structured sequence involved two unique associations and one ambiguous association to complete one cycle of the structured sequence. Thus, for the sequence 32121, there was a unique association between Position 3 and 2 and a unique association between Position 2 and 1 (which occurred twice in the cycle). However, Position 1 could be followed by either Position 2 or 3, depending on the phase of the structured sequence. Thus, the association of Position 1 with the next position was ambiguous. Nevertheless, the presence of unique associations in the structured sequence should be sufficient to allow a learning mechanism based largely on linear associations to operate, given the short length of the structure (Keele & Summers, 1976). The assumption guiding our sequence selection in Experiment 1 will be more directly explored in Experiments 3 and 4.

Method

Subjects. Forty-eight subjects, recruited from introductory psychology courses at the University of Oregon, were randomly assigned to four different groups of 12 subjects each.

Apparatus. The stimuli were presented on a Video 100 monitor controlled by an Apple IIe microcomputer.

Procedure. Each subject performed first in a dual-task condition of tapping and tone counting tasks and then in a generation task. Each trial started with the appearance of an *X* mark in one of three horizontal positions on the monitor. The index, middle, and ring fingers of a subject's right hand rested on three microswitch keys. The subject's task was to press as fast as possible the key in the position corresponding to that of the *X* mark. After each response, a tone occurred 40, 80, or 120 ms later. The tone could either be high pitched or low pitched. The onset of the next *X* mark occurred 200 ms after the onset of the response. Each subject received two blocks of practice trials with 30 trials in each block. There were 10 subsequent training blocks with 100 trials in each.

The subject was to count the number of high-pitched tones, ignoring the low-pitched ones, and report the number after the block of 100 trials. For half the subjects, the number of high-pitched tones ranged from 25 to 50 tones per 100 tone block. For the other half the number of high-pitched tones ranged from 50 to 75. The former condition will be called the *easy tone task* and the latter the *difficult tone task*. The tone task was run concurrently with the tapping task in both the practice and the training phases. Subjects were given feedback on the tone task, and the instructions emphasized accuracy.

During the practice blocks the *X* marks appeared at random in the three positions, with the restriction that the same location could not be used successively. For half of the subjects in the easy tone condition and for half of the subjects in the difficult tone condition, the signals continued to occur at random positions, again with the restriction of no successive repetitions. These groups constituted controls against which to compare performance of the other two groups. For the other two groups, the signals in the training session occurred in a five-element structure. For each of these two groups, there were six

subgroups, each with a different ordering. The six structured sequences were 12323, 13232, 21313, 23131, 31212, and 32121. Each of the structures can be characterized as having one unique position and two repeated positions. Ignoring starting position and allowing no immediate repetitions of an element, these constitute all possible structures of five elements with three elementary components. For each block of 100 trials, the structured sequence was repeated 20 times without break. Each block of trials started at a randomly selected place in the structure. For example, the structured sequence we have labeled 12323 on some trials might begin 32312.... This manipulation was adopted to further obscure the fact that there was a repeating structure. No feedback was given on the primary reaction time task because this would interfere with the presentation of the next signal.

The training phase of the experiment, therefore, involved a 2×2 factorial design, with one factor being structured sequence versus random sequence training and the other being tone-task difficulty. Different subjects were assigned to each of the four resultant conditions.

The second phase of the experiment involved the generate task. The *X* mark in the unique position in the structured sequence was presented first. The subjects' task was to press the key indicating which position would come on *next*. Instructions explicitly emphasized accuracy, and the subjects were told that reaction times would no longer be recorded. They were *not* informed that the stimuli would occur with a structure, but the fact that the task was to predict the next stimulus made implicit that they were to learn a structured sequence. Following a key press, the *X* in the next position came on, the subjects had to predict the next position of the *X* mark. Those subjects who had a structured sequence during the tapping task continued with exactly the same structure in the generation phase. Of those who had random sequence, 2 subjects from each difficulty level were assigned to each of the six structures during generation. The generation procedure continued for a single block of 100 trials—that is, 20 cycles through the structured sequence.

Results and Discussion

To ensure that subjects paid sufficient attention to the tone task, a criterion was adopted to eliminate from the data analysis any subjects whose tone counts were in error, on average, by more than 10%. For example, if the number of target tones in a given block was 60, the acceptable range of response was 54 to 66. One subject exceeded the criterion and was replaced. Although all the remaining subjects averaged

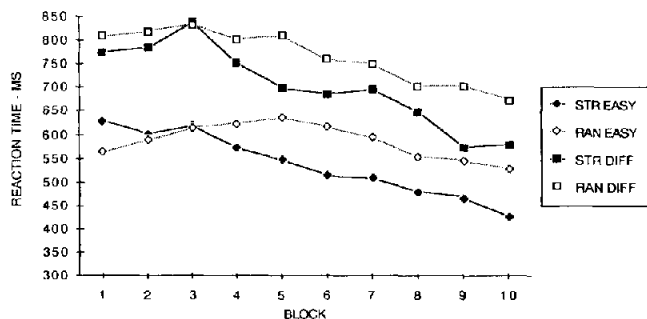


Figure 1. Experiment 1: Reaction time on the button-pressing task as a function of the amount of practice, of whether the stimuli appeared at random (RAN) or in structured (STR) sequence, and of the difficulty of a tone-counting distraction task. (DIFF = difficult.)

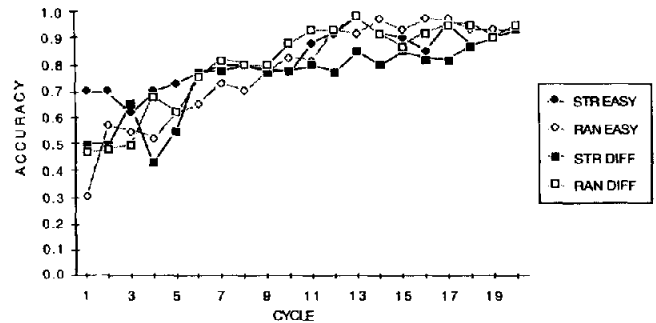


Figure 2. Experiment 1: Proportion correct on the generation task in predicting the next item in a structured sequence as a function of the cycle number of the structure. (During this phase, signals occurred in structured sequence and without the distraction task, but in the preceding reaction time phase signals occurred either at random [RAN] or in structure [STR] and with either the easy or difficult [DIFF] distraction task.)

within 10% of the correct number of tones, it still was the case that on most trials they did not get the number of tones exactly correct. However, the percent trials in which errors were within the accepted range did not differ between the random and structured sequence conditions, being 77% in both cases. More trials were in error on the difficult tone-task conditions (84%) than on the easy ones (68%), reflecting the greater opportunity for error with more target tones.

Figure 1 shows the effect of practice on reaction times in the tapping task. With practice, reaction time improved for all of the groups, but the amount of reduction was larger by about 100 ms for the structured conditions than for the random conditions. The Block of Practice $\times$ Condition (random vs. structured) interaction was significant, $F(9, 396) = 3.50$, $p < .001$, $MS_e = 7,065$. These results indicate that the structured sequence was learned. Distraction task difficulty had no bearing on the amount of learning of the structured sequence: The interaction of Random versus Structured Sequence $\times$ Tone-Task Difficulty was not significant, $F(9, 396) < 1$, $MS_e = 7,065$. Importantly, tone-task difficulty had a main effect: $F(1, 44) = 15.92$, $p < .001$, $MS_e = 228,329$, indicating that the attentional manipulation was effective. Nonetheless, increasing the attentional demands of the tone task had no influence on the amount of structured sequence learning. Error rates on the tapping task ranged from 2% to 3% and did not differ significantly in the various conditions.

Given the evidence for structured sequence learning in the presence of attentional distraction, the next question concerns whether subjects were aware of the structure. The results of the generation task are shown in Figure 2. Performance for three groups started out around 50% and gradually increased as the structured sequence was memorized. However, the mean accuracy of the fourth group, the structure subjects with the easy tone task, was higher than for the random conditions over the first few cycles through the sequence. The difference disappeared with practice. Because the main interest concerns awareness at the beginning of the generation task, an analysis of variance was conducted on the average of the first four cycles. The interaction of Tone-Task Difficulty $\times$ Random

versus Structured Sequence was significant, $F(1, 44) = 5.74$, $p < .025$, $MS_e = 0.023$). Thus, some of the subjects in the easy tone task appeared to be partially aware of the structured sequence. For the subjects who learned the structure concurrently with the difficult tone task, however, subjects appeared to be unaware of the structure. The generation performance was similar for these subjects and the two random conditions.

Taken together, the two phases of Experiment 1 suggest that, as measured by a performance criterion, a structured sequence can be learned even when subjects are not aware of it. These results correspond with data obtained on a similar task by Nissen and her colleagues (Nissen & Bullemer, 1987; Nissen, Knopman, & Schacter, 1987) and with data from other paradigms (e.g., Lewicki, Hill, & Bizot, 1988). Such results can be accounted for by postulating the existence of a memory system that operates outside awareness. A number of theorists have referred to such memory systems as *procedural memories* and have argued for their independent status from *declarative memories* (see Mishkin & Appenzeller, 1987; Mishkin, Malamut, & Bachevalier, 1984; Squire, 1986; and Schacter, 1987, for a discussion of these separate kinds of memories). One cautionary note is in order: Possibly subjects were aware of the sequence during learning, but such awareness was not reflected in the generation task because they were not apprised of a relation between the two phases of the experiment. The issue will be discussed in Experiment 4, where it is concluded that awareness, if it exists, is minimal and does not affect learning of the structured sequence.

Our results differ from those of Nissen and colleagues in one important respect. They showed that although structured sequence learning without awareness was possible, a secondary task prevented sequence learning altogether. In contrast, we find evidence of structured sequence learning not only in the absence of awareness but also with attentional distraction. Because both our study and Nissen's used similar distraction tasks, presumably the different results are due to some difference in the structured sequences learned. This issue will be explored further in Experiments 3 and 4.

Two qualifications should be added to our claim that the structured sequence can be learned with attentional distraction. First, it is obvious from the low error rates on the reaction time task that subjects were able to attend to the individual stimuli. Our claim is that the attentional manipulation distracted subjects from attending to the *relations* between successive stimuli. Despite this distraction, subjects were able to learn the structured sequence. Second, the distraction tone task is very different from the primary tapping task. It is possible that a distraction task more similar to the tapping task might have prevented learning of the structured sequence. We will discuss this possibility and its theoretical implications in the General Discussion.

Experiment 2

The question of interest in Experiment 2 concerns whether sequence memory established in the dual-task situation is specific to the motor system in which learning occurs. As mentioned in the introduction, some motor programs appear partially independent of the effectors that produce the action.

However, the issue has received little formal investigation with respect to rapid sequences, particularly in a manner that allows quantitative assessment. Moreover, it remains an empirical question whether a motor program established outside awareness, and presumably dependent on procedural memory, is effector independent. Had a single-task method been used, it might be argued that any transfer of sequence knowledge from one effector system to another might be based on declarative knowledge (because subjects could have become aware of the presence of a structured sequence).

Method

Subjects. Twenty-four new subjects, taken from the same pool as the previous experiment, participated in the experiment.

Procedure. The first phase of Experiment 2 was identical to the group in Experiment 1 which received the structured sequence in the tapping task under the difficult tone-task condition. The subjects first performed two practice blocks, with a random sequence of the X marks, followed by 10 blocks with a structured sequence. In this initial training phase, subjects responded by pressing the three keys of the keyboard with *three* fingers, each finger corresponding to a different stimulus location. After this phase, the subjects were subdivided into two transfer groups with 12 subjects each. One group transferred to a situation in which the structured sequence remained unchanged, whereas for the other group stimulus selection became random. The critical change in transfer involved the effector system. Instead of using three fingers, *all* of the subjects were instructed to hit the three keys with *only* the index finger. This requirement made the arm the primary effector because the subjects had to move the arm up and down and back and forth to hit the target key. Not only do the arm and finger movements use different muscles, but they are quite different in nature. This effector manipulation is much like switching from touch typing to "hunt and peck" typing. Moreover, the middle and ring fingers no longer participate in the motor production. In this transfer phase an additional 10 blocks of 100 trials were presented. The distraction task was continued as well. The transfer of structured sequence learning can be assessed by comparing the structure and random groups in the transfer phase. Note that the division into structure and random groups occurred only in the transfer phase. Both groups initially trained under structure conditions.

Because of the transfer phase, it was not feasible to use the generation task to assess awareness at the end of the training phase. However, given that Phase 1 was the same as in Experiment 1 and that only the more difficult distraction task was used, there is no reason to assume any difference from Experiment 1, in which subjects appeared unaware of the structure at the end of initial training.

Results and Discussion

The tone task was used solely for distraction purposes, and results on it are of little interest. The number of errors for all subjects was within criterion levels. That is, the average response to the number of tones was within 10% of the actual number of tones.

The reaction times on the tapping task are shown in Figure 3. As stated above, both groups practiced with the fixed structure of signals during the training phase. Therefore, no difference in performance should be expected. An analysis of variance (ANOVA) confirmed that there was no reliable difference between groups ($F < 1$) and no interaction between

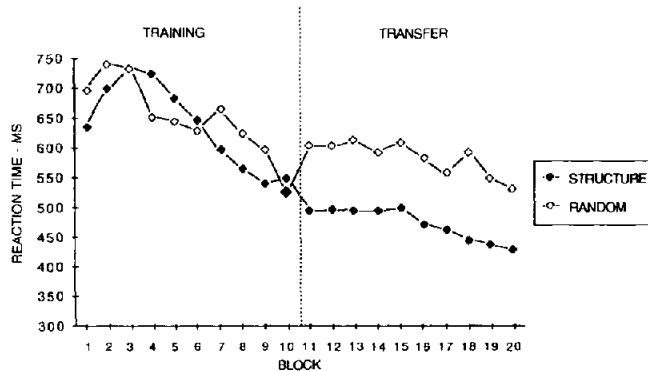


Figure 3. Experiment 2: Reaction time on the button pressing task as a function of practice. (During training, both groups were treated identically by practicing the sequence by pressing buttons with finger movements. In transfer, both groups responded with the arm. One group maintained the same structure, and the other transferred to random signals.)

groups and blocks of practice, $F(9, 198) = 1.64$, $p > .05$, $MS_e = 9,176$.

During the transfer phase, the reaction times for the group that shifted to randomly occurring stimuli increased dramatically. The group that maintained the structure showed no deficit when transferred to Phase 2. This Group $\times$ Phase interaction was significant, $F(1, 22) = 6.62$, $p < .02$, $MS_e = 42,671$. The advantage of the structured over random conditions apparent on the first transfer block was unaltered over successive transfer blocks. An ANOVA showed no interaction between block of transfer with structured versus random sequence ($F < 1$).

The fact that the transfer group that maintained the structure showed superior performance despite novel response requirements indicates that at least a portion of the memory for the structured sequence is independent of the effector system. It is not possible to determine definitively whether there was some residual structured sequence learning that was specific to the effectors, but any such residual component appears to be small, if existent at all. On the initial transfer trials, the random group responded to the signals about 100 ms more slowly than did the structured group. This value, which is an index of learning, can be compared with the difference between the random and structured groups of Experiment 1 near the end of the initial practice (see Figure 1). The difference scores in Experiment 1 were also about 100 ms. Thus, it appears that the advantage of the structured sequence representation transfers completely when the effectors are changed. However, this comparison is across experiments, and a more conclusive answer awaits further study.

Error rates on the tapping task were nearly identical for the two groups, averaging 3% for the random group and 2% for the structured group. There were no noticeable changes in error rates from training to transfer in either group. For the group that maintained the structure, error rates were 1% and 2% on Blocks 10 and 11. For the group shifted to random, the error rates were 2% and 2%.

The reaction time results indicate that the sequential learn-

ing investigated in these studies is not represented at the level of the nervous system that selects and activates specific effectors. Rather, structured sequence information appears represented at a level that describes either where in the environment successive movements are to occur or where successive stimuli are to occur. Presumably, a subsequent system determines how specific movements are implemented. The findings of Experiment 2 do not imply that learning never occurs at the effector-specific level. It is likely that subjects learned how to properly move effectors to designated locations, but this type of learning did not differ between the random and structured groups. Moreover, the individual actions required in both the finger and arm conditions are quite simple and presumably well experienced. Thus learning of the key pressing actions themselves might have already been near an asymptote (see MacKay, 1982).

Experiment 3 and 4

Experiments 1 and 2 demonstrated structured sequence learning with attentional distraction. Those results conflict with results of Nissen and Bullemer (1987), who used a very similar paradigm. They used four different signal portions embedded in a 10-item structure. Each element occurred two or three times in a structure, with no association being unique. A structured sequence of 10 may be too long to be learned without attention. However, another possibility is that structures in which there are no unique associations between individual elements may favor hierarchic coding. Such hierarchic representation may require attention for its formation. To test this idea, the present experiments varied structured sequence type.

Method

Subjects. Thirty-six subjects participated in Experiment 3, and 72 subjects participated in Experiment 4. The subjects were selected from the same pool as in the previous experiments. Five subjects in Experiment 3 failed to meet the criterion on the tone task of averaging within 10% of the correct number of high-pitched tones. They were replaced by others to maintain 12 subjects per group. In Experiment 4, 16 subjects were replaced, 8 for making too many errors on the tone task and 8 for failure to respond appropriately on the generation task. The failures on the generation task involved pressing keys corresponding to the current signal rather than the keys for the predicted signal. This may have occurred because a large number of non-native speakers of English were used, and they may have had trouble with the instructions.

Procedure. As in Experiments 1 and 2, the stimuli were *X* marks that appeared in varying horizontal positions on a monitor. There were three structure types, which we will call *unique*, *ambiguous*, and *hybrid*.

The unique type of structure involved five signal positions, none of which were repeated in a structure cycle. An example is 15243, where the numbers correspond to the five positions from left to right. The five fingers of the right hand were placed on the five keys of a keyboard, and subjects responded to the signals by pressing the key corresponding to the signal position. It is possible for a structure of this type to be coded as a set of linear associations in which 1 is associated with 5, 5 with 2, 2 with 4 and so on. In other words, every element has a unique association with the next element in the

structure. To ensure that the results were not due to the particular structure, four different versions were used. The four versions are 12354, 13425, 14532, and 15243. Each block in this condition had 100 trials, resulting in 20 cycles of the structured sequence.

A second type of structure, called ambiguous, was six items in length and was composed of three different signal positions, each of which occurred twice. The subjects responded with the index, middle, and ring fingers of the right hand. An example of such a structure is 132312. In such a structure there are no unique associations between elements of the structure. For instance, Position 1 is followed by Position 2 at some point in the structure and by Position 3 at another point. Thus, simple, pairwise linear associations would be unable to code this structure usefully. Such structures might be coded hierarchically by parsing the structure into 2 or 3 substructures. Note that the constraint of no unique association for the ambiguous condition mandates that the structure must be six items in length. In contrast, the unique structures we selected were five elements in length. For the unique condition to have six elements, both hands would be needed to implement six unique responses. It was deemed more important to keep all responding to one hand rather than to equate structure length. Three ambiguous versions were used: 123132, 123213, and 132312. Given our constraints (no unique association, and starting position of the block is random), the three versions constitute all possible structures. Each sequence cycled 20 times for a total of 120 key presses per block.

An example of the third sequence type, called hybrid, is 142312. Subjects used all fingers except the thumb of the right hand. The structures here are like those of Experiments 1 and 2 but have six items rather than five. Each structure has two unique associations (Position 4 is followed by Position 2, and Position 3 is followed by Position 1 in the example above) and two ambiguous associations (after Position 2 and after Position 1). Six versions were used: 123243, 123134, 143132, 142312, 132412, and 423213. Each structure cycled 20 times for a total of 120 key presses per block. One reason for including this condition was to match the length of the ambiguous structures. In addition, the hybrid structures provide a close replication of Experiments 1 and 2. One might suppose that the repetition of elements in the hybrid condition would also force hierarchic coding. However, it is possible that unique elements provide a peg around which linear associations could operate.

Our hypothesis is that the unique and hybrid structures can be learned under attentional distraction, whereas ambiguous structures cannot. This hypothesis was examined in Experiment 3, in which the tapping task was performed with the tone task. Three groups with 12 subjects each were assigned to the three structured sequence types. The number of high-pitched target tones ranged between 50 and 75 for the unique condition with blocks of 100 trials and between 60 and 90 for the other two conditions with blocks of 120 trials. These ranges were comparable to the difficult distraction condition of Experiment 1.

Our hypothesis rests in part on the assumption that when attention is not distracted, all three structured sequence types can be learned. This assumption was explored in Experiment 4, in which learning of the three structured sequence types was examined with and without the distracting tone task. Three groups of 12 subjects, assigned to one of the three structure types, performed both the tapping and the tone task. This is a direct replication of Experiment 3. Three additional groups of 12 subjects were tested on the tapping task alone.

To assess structured sequence learning, a method developed by Nissen and Bullemer (1987) was used. After several training blocks with the structured sequence, the subjects received two blocks, with random events followed by an additional two with the initial structure. To the degree that the structure has been learned, reaction times should increase when the events switch to random sequence and decrease with the switch back to the structured sequence.

Subjects began the experiment with two blocks of 30 practice trials involving random selection of the stimulus locations. The number of horizontal positions of the X marks in the random sequence blocks was comparable for each structure type condition to the number of positions in the structured sequence blocks. This was followed by 14 blocks of either 100 or 120 trials, depending on the structure type: 2 blocks with random signal presentations, 8 blocks with the structure, 2 more random blocks, and 2 final structure blocks. There was no change in the interblock interval when switching back and forth from random to structured conditions. Thus, subjects had no external cues regarding the changes.

Following the 14 test blocks, subjects performed on the generation task in a manner similar to Experiment 1. One change, however, involved the first generation cycle. The first two signals of the structured sequence were shown prior to the first required response in order to reduce any ambiguity about the correct next item. This change was especially important for the ambiguous condition, in which after only a single stimulus it is ambiguous which stimulus is next.

Results and Discussion

Experiment 3: Dual-task analysis. Reaction times on the tapping task are shown in Figure 4. With practice, reaction times improved. The degree to which the improvement is due to learning the structured sequence can be assessed by examining changes in reaction time when the stimuli became random on Blocks 11 and 12 and returned to the structure on Blocks 13 and 14. For the unique and hybrid conditions, performance slowed about 90 ms with the shift to random orders and immediately rebounded on the shift back to the structure. For the ambiguous condition, the shift from structured to random events and back resulted in changes of only about 20 ms, suggesting that for this group much less had been learned about the structure. For purposes of statistical analysis, the two sequence blocks just prior to the shift were averaged. Likewise, the two random blocks and the following two sequence blocks were averaged. This procedure yielded preshift, shift, and postshift scores for each subject to be entered into an analysis of variance along with the factor of sequence type. The interaction of the two variables was sig-

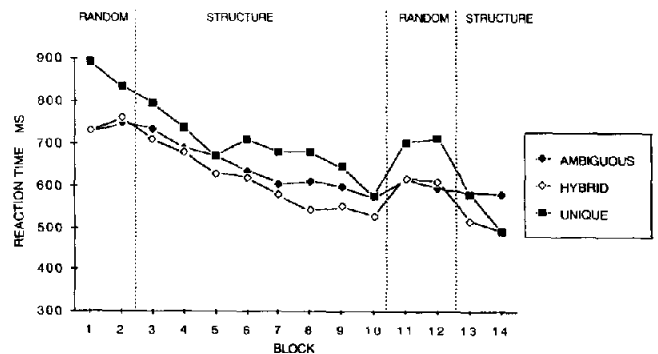


Figure 4. Experiment 3: Reaction time as a function of practice for unique, hybrid, and ambiguous structured sequences. (On Blocks 11 and 12 the signals switched from structured to random order, and on Blocks 13 and 14 the signals again occurred in structure. Performance was concurrent with a tone-counting distraction task.)

nificant, $F(4, 66) = 3.50$, $p < .01$, $MS_e = 4,745$. Post hoc analyses of variance showed a significant difference between the shift reaction times and the average of both the pre- and postshift reaction times for the unique ($F[1, 11] = 27.02$, $p < .01$, $MS_e = 3,729$) and hybrid ($F[1, 11] = 29.66$, $p < .01$, $MS_e = 1,755$) groups. The differences were not significant for the ambiguous group ($F[1, 11] = 3.15$, $p > .05$, $MS_e = 798$).

Error rates on the tapping task showed a similar pattern. For the preshift, random, and postshift blocks, the error rates for the unique group were 2%, 6%, and 4% respectively. For the hybrid group they were 2%, 4%, and 3%. For the ambiguous group, they were 1%, 2%, and 2%. Thus, the two groups who appeared to learn the sequence also made the most errors during the random blocks. However, neither the difference in error rates for the three sequence types nor the interaction of structured sequence type with shift position was reliable.

Experiment 3: Generation analysis. The generation data are shown in Figure 5. It is apparent that, as a whole, none of the groups showed much awareness of the structured sequences. The ability to specify the stimulus to follow the current one was poor on the first few cycles and improved only gradually. A statistical comparison of groups is inappropriate because they have different numbers of elements, resulting in different a priori probabilities of guessing the succeeding stimulus. More appropriate comparisons are available in Experiment 4, where the general problem of a priori probabilities is further discussed. The implication of the generation data at this point is simply that the greater success of the unique and hybrid groups in structured sequence learning in comparison with the ambiguous group appears not to be the result of marked awareness of the structured sequences.

The results for the unique and hybrid groups are in agreement with results of the first two experiments. Performance measures of sequential learning show improvement in the presence of distraction. For the ambiguous group, however, structured sequence learning was minimal or absent. These results suggest that a fundamental difference exists between memory representations for the ambiguous and nonambiguous conditions. We hypothesize that the ambiguous condition forces hierarchic coding and that such a code requires attention for its formation. This hypothesis rests in part on the assumption that ambiguous sequences can be readily learned when attention is focused on the task. Experiment 4,

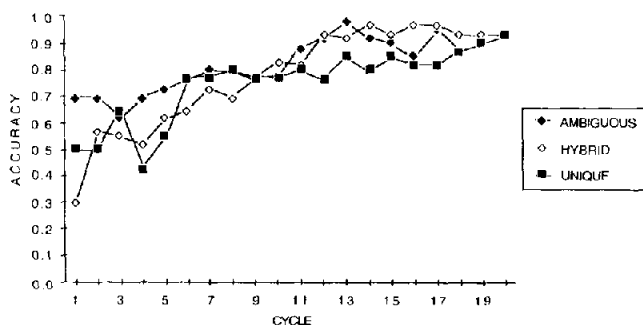


Figure 5. Proportion correct on the generation task as a function of structured sequence type: Experiment 3.

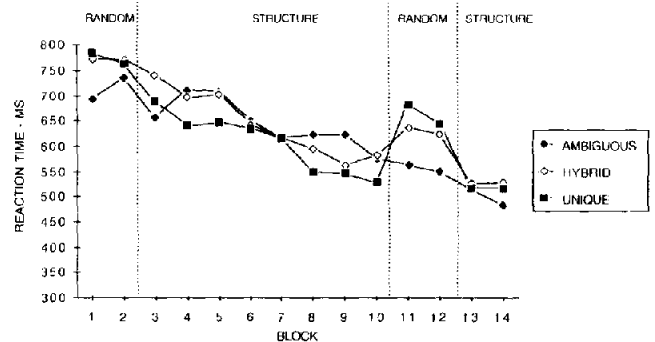


Figure 6. Experiment 4: Reaction time as a function of practice and with the tone-counting distraction task. Blocks 11 and 12 involved random signals. (On other blocks the signals occurred in structured sequence.)

in which performance in dual task conditions was compared with performance in single-task conditions, was designed to test this assumption.

Experiment 4: Dual-task analysis. Reaction time results for the dual task condition are shown in Figure 6. The results largely replicated those of Experiment 3. For both the unique and hybrid groups, reaction times decreased with practice on the structured sequence, increased when switched to random events, and recovered when switched back to structured ones. The ambiguous group showed little evidence of structured sequence learning. Switching from the structured to the random signals actually produced a slight *improvement* in reaction time, presumably as a result of nonspecific practice. With the shift back to structure, reaction times continued to improve about 50 ms. Again, this improvement may be due to nonspecific practice effects. At any rate, the effects of changing from structure to random and back are much less than for the other two groups.

Analysis of variance confirmed these observations. The interaction of structured sequence type with the block type (preshift, shift, and postshift) was significant, $F(4, 66) = 9.24$, $p < .01$, $MS_e = 2,403$. The nonspecific practice effect can be partially eliminated in a statistical analysis by averaging the preshift and postshift blocks because they bracket the shift blocks. Post hoc analysis revealed that both the unique ($F[1, 11] = 42.69$, $p < .01$, $MS_e = 2,612$) and hybrid ($F[1, 11] = 48.26$, $p < .01$, $MS_e = 761$) groups showed significant effects. The difference was not reliable for the ambiguous group ($F < 1$). Though the error rates showed a similar pattern, the interaction of shift position by structure type was not reliable for these data. Overall error rate was 3%. In summary, the pattern of results under dual-task conditions was quite similar to that for Experiment 3: Both the unique and hybrid groups showed structured sequence learning, but the ambiguous group showed none.

Experiment 4: Single-task analysis. Figure 7 shows the reaction time results for the three groups that did not have the accompanying tone task. Here all three groups exhibited structured sequence learning, though to differing degrees. The unique and hybrid groups showed a reaction time alteration of 250–350 ms with the change from structure to random and

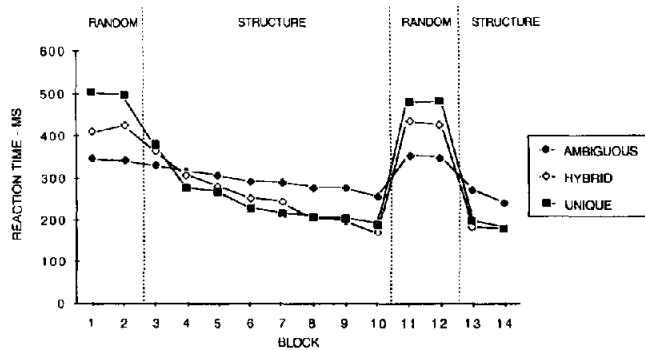


Figure 7. Experiment 4: Reaction time as a function practice without the tone-counting distraction task. (Blocks 11 and 12 involved random signals. On other blocks the signals occurred in structured sequence.)

back. Of critical importance, the ambiguous group also exhibited evidence of structured sequence learning. The shift from structured to random increased reaction time about 80 ms; the shift back to structure produced an improvement of almost 100 ms. Post hoc analyses of variance compared reaction times on the two random blocks with the average of the two preshift and two postshift blocks. The comparison was highly significant for all three structure types (unique, $F[1, 11] = 283.00$, $p < .01$, $MS_e = 1,775$; hybrid, $F[1, 11] = 235.16$, $p < .01$, $MS_e = 1,602$; ambiguous, $F[1, 11] = 70.3$, $p < .01$, $MS_e = 686$).

Although the ambiguous group showed clear evidence of structured sequence learning in the absence of the distraction task, the change in reaction time from structured to random conditions was less than for the other two groups. This could indicate less structured sequence learning, but a strict comparison of reaction time changes is problematic. The main problem is that reaction times for the three groups were not equal during the random blocks because the number of stimulus-response alternatives was not equal. The random conditions in the unique group constituted a five choice-reaction time task because there were five stimulus locations. There were only four choices for the hybrid group, and the number of alternatives was only three for the ambiguous group. As would be predicted from the Hick-Hyman law (Hick, 1952; Hyman, 1953), the reaction times under the random conditions were fastest when the number of stimulus-response alternatives was smallest. (This same effect can be seen in Figures 4 and 6 for the dual-task conditions of Experiments 3 and 4.) As a consequence, the range in which the ambiguous group could exhibit an improvement in reaction times was reduced. However, the fact that reaction times in the structured sequence blocks were slowest for the ambiguous group does suggest lesser learning of the structure. Alternatively, if the ambiguous sequence was hierarchically coded, such a code could result in slower retrieval times because more than one level is involved.

Despite the ambiguity concerning the amount of learning, the crucial point remains that the ambiguous structured sequence was learned when attention was available, but learning was blocked when attention was diverted. This stands in sharp

contrast to the results for the other two structure types in which learning occurred in both the single- and dual-task situations. The results are consistent with the hypothesis that frequent repetition of events in differing contexts favors the establishment of a hierarchic code and that such coding requires attention. On the other hand, because structured sequences with at least some unique associations are learnable under distraction, there also appears to be a second kind of learning system that is based on sequential associations. This type of coding may not require attention for learning the relations between successive events.

Experiment 4: Generation analysis. The generation paradigm was employed at the conclusion of Experiment 4 after the last of the performance blocks with the structured sequence. The issue, as before, was to assess the degree of awareness of the structure once it is learned. Figure 8 shows the results for the six groups. Although there was a small tendency for greater accuracy on the initial cycles for structured sequences learned without the distraction task, it is striking that performance was so poor for all six groups. Analyses of variance comparing single- with dual-task generation performance were based on the first four cycles of generation. Each analysis compared only the single and dual task conditions of a single structure type because the base-rate guessing probabilities were different. Only the hybrid task showed reliably better performance under single-task conditions than under dual-task conditions ($F(1, 22) = 5.23$, $p < .05$, $MS_e = 0.032$). Thus, though subjects in all three structured sequence types showed a substantially larger amount of structure learning in the performance phase when there was no distraction, the lack of distraction only partly led to an increase in awareness of the structure. It is especially noteworthy that removal of distraction allowed learning of the ambiguous sequences even though there was no significant change in the generation task.

The results of the generation task may serve to illuminate two related issues regarding awareness. The first issue concerns whether subjects became aware of the structured sequence. Secondly, if some subjects became aware of the structured sequence, did it enhance their learning of the structure as measured by their reaction time to the tapping task? In particular, is it possible that the unique and hybrid structured sequence types were learned in the dual-task conditions be-

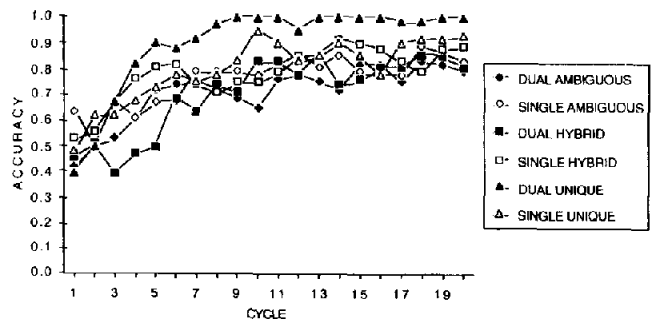


Figure 8. Proportion correct on the generation task as a function of structured sequence type and training under dual- or single-task conditions: Experiment 4.

cause subjects in these conditions were more aware of the structures?

The data were not conclusive with regard to the first issue. It is clear from the data that most subjects had poor or nonexistent awareness of the structured sequence. It is possible, however, that some subjects became partially aware of the structure. One feature of the generation data that may suggest partial awareness of the sequence even under dual-task conditions was the apparently higher than chance probability of guessing the correct stimuli on the first cycle of a structured sequence. However, deeper consideration suggests that this interpretation is problematic. Consider the unique data under dual-task conditions of Figures 5 and 8. The percent correct was around 40%–45% on the first cycle of the generation task. Given that there were four alternative signals besides the currently displayed one, guessing probability might be 25%. There are two reasons, however, why this figure may be too low. First, on the initial generation cycle, subjects were given the first *two* stimuli of the structure. Second, there is no reason to suppose that subjects guessed with a strategy of strict replacement. Suppose that given the first two stimuli, subjects guessed at random one of the remaining three. After the third was presented, they guessed one of the remaining two, and after the fourth stimulus, they selected the remaining alternative. Given such assumptions, the chance probability on the first generation cycle would be considerably in excess of 25%. Indeed in other unpublished work that we have conducted in which subjects start immediately on the generation phase, performance on the initial cycle of the hybrid and unique sequences averaged about 40% correct.

Nevertheless, it is possible that a subset of the subjects became aware of the structured sequences. Perhaps only these subjects showed learning, and this would account for the learning of the unique and hybrid structured sequences in the dual-task conditions. To test this possibility, a post hoc analysis was conducted which divided subjects for each structure type into those who did best during the first four cycles of the generate task and those who did poorly. The percent correct of the "aware" and "nonaware" groups in the hybrid condition of the dual task over four cycles of the structure was 0.57 and 0.35, respectively. The difference was even larger in the unique condition of the dual task: 0.80 and 0.38 for the aware and nonaware groups, respectively. Despite this large difference on the generate task, the learning of the structured sequence as measured by reaction time was virtually identical. Thus, even if some subjects were aware of the structured sequence, it appears that this awareness did not enhance the learning of the structure.

One potential problem with the generation task used in the previous experiments is that subjects were never informed of a structure during either the reaction time phase or the generation phase. Thus, it might not reflect explicit awareness of the subjects. To investigate further the issue of awareness, a supplementary experiment was conducted. There were four groups with 12 subjects each. One group never received the initial reaction time task and instead started off in the generation phase. The sequences used were of the hybrid variety of Experiment 4. The purpose of this group was to establish a baseline of generation performance with no previous exposure

to the structure. A second hybrid group performed on the reaction time task concurrently with the tone-counting distraction. After the performance phase those subjects were explicitly asked whether the signals had occurred at random or in structure, and they rated their confidence and knowledge. Then the generation phase began. They were told that there had been a structure during the initial phase, and the structure they were to learn during the generation phase was identical. A third group was treated identically except without the distraction task during the performance phase. The fourth group received the ambiguous structured sequences of Experiment 4 under single-task reaction time conditions and followed by awareness assessment. Thus, in contrast to the earlier experiments, although subjects were not told about a structure during the initial phase, they were clearly informed about its presence prior to testing on the generation task. Specifically, they were told that it was the same structure as had occurred in the initial phase whether they had noticed it or not.

The three groups that performed the reaction time phase showed reaction time results similar to those of the earlier experiments. Each group learned the structured sequence, as indicated by the deterioration in performance when shifted from structured sequence stimuli to random stimuli ($F[1, 11] = 13.57, p < .01, MS_e = 1,006$ for the dual-hybrid conditions; $F[1, 11] = 171.12, p < .01, MS_e = 1,847$ for the single-hybrid condition; $F[1, 11] = 41.2, p < .01, MS_e = 2,113$ for the single-ambiguous condition).

When asked immediately after the performance phase whether the preceding signals had occurred at random or in structure, only 4 of the 12 hybrid dual-task subjects reported a structure. Under single-task conditions, 10 of the 12 hybrid subjects reported awareness. For the ambiguous structure, which also was conducted under single-task conditions, all 12 subjects reported awareness. Thus, more subjects do report awareness of a sequence following single-task conditions than following dual-task conditions. The fact that more subjects reported a structure under single-task conditions suggests that removing the distraction task increased awareness. Moreover, the 4 subjects who under dual-task conditions reported a structured sequence exhibited less confidence on a rating scale (3.8 on a scale of 5 vs. 4.2 and 4.0 for the single-hybrid and single-ambiguous conditions, respectively) and less knowledge (2.5 on a scale of 5 vs. 3.8 and 3.2 for the single-hybrid and single-ambiguous conditions) than did subjects in single-task conditions who reported a structure.

These qualitative assessments all suggest that, at best, most subjects have very limited awareness of the structures after dual-task conditions.

Interestingly, reporting awareness did not translate into strong performance on the generation task (Figure 9). The group with no prior exposure to the sequence established a baseline. The group's performance was slightly worse on the first generation cycle than for the hybrid dual-task group, but on subsequent cycles, generation performance differed little between the two groups and was substantially below perfect. The hybrid group receiving single-task training performed somewhat better on generation but was still less than perfect. Of perhaps greatest interest were results for the group receiving ambiguous structures under single-task conditions. Although

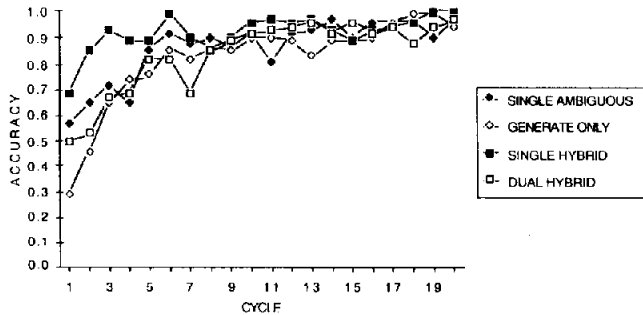


Figure 9. Proportion correct on the generation task for subjects who had never seen the structured sequence or for subjects trained on either hybrid or ambiguous sequences under dual- or single-task conditions. (Supplement to Experiment 4).

all of these subjects reported awareness of the structure, on the whole their generation performance was poor and little different from that of either the dual-task or single-task groups receiving the same sequence types in Experiment 4 (see Figure 5 and 8). The poor generation performance occurred despite the fact that explicit instructions emphasized the continuity of structured sequence from the reaction time phase to the generation phase. Thus, lack of explicit instructions in the previous experiments apparently did not affect the results. These results lead us to conclude that especially under dual-task conditions, subjects have very little explicit knowledge about the sequence. Even under single-task conditions, claims of awareness appear based on scanty knowledge because generation performance was so poor. Such results are consistent with those of Willingham, Nissen, and Bullemer, 1989. They report that some subjects under single-task conditions learn sequences but show no evidence of awareness of them.

Thus, although attention may be necessary for some types of learning, namely of repeat-type sequences, clear awareness may not be. It appears that lack of attention rather than the absence of awareness blocks learning of structured sequences with all repeated items.

General Discussion

The present experiments make two primary contributions. One concerns effector independence. Once a structured sequence is learned by the use of one particular effector system, that memory can be accessed by different effector systems. The other contribution concerns attention and learning: Some types of structured sequences can be learned in the presence of an attentional distractor; others, in which there are no unique associations between any two elements in the structured sequence, cannot. A variety of issues may be raised regarding both of these primary contributions.

Effector Independence

How can one determine whether representation of a sequence of activity is independent of the effector system that typically implements it? One argument refers to the similarity

of motor production across different effectors, as in the similarity of arm writing to handwriting. A problem here is that similarities tend to be subjective. Moreover, despite similarities, distinct features occur with different effectors, and it is difficult to know whether those arise from differences in sequence representation or differences in implementation (Wright, in press). What is needed is a method of separating sequence knowledge from effector competence.

The method we have selected makes use of familiar subtractive logic. Subjects were trained on a structured sequence using one effector system, the fingers. They were then transferred to a new effector system, the arm, not previously trained on the structure. To assess the degree to which sequence knowledge transfers, performance with the new effector must be compared with a situation that is methodologically identical but without the possibility of sequence knowledge. This is done by subtracting performance with the structured sequence from performance under random events. When this comparison is made across Experiments 1 and 2, it appears that sequence knowledge transferred in its entirety to the new effector. The conclusion that structured sequence knowledge of the sort studied here is effector independent has been bolstered by other similar studies we have conducted and reported in preliminary form (Keele, Cohen & Ivry, in press).

Attention and Learning

In our experiments the distraction task was inserted in the intervals between responses and subsequent stimuli of the primary task. Such a manipulation allows attention to the stimuli and the responses of the primary task but appears to block attention to the relation between successive events. This was demonstrated in the first experiment in which difficulty of the distraction task was varied. When the proportion of tones which were targets was increased, reaction time increased on the primary task, indicating that the attentional manipulation was effective. Increased attentional demand left structured sequence learning unaffected, suggesting that such learning can occur with less attention being available between the successive stimuli. Despite lack of attention to the relations, subjects are able to associate successive stimuli as long as some of the pairwise associations are unambiguous.

The secondary tone task used to distract attention is very different from the primary tapping task. In particular, it is processed by a different modality and, unlike the tapping task, it does not involve spatial stimuli. We used such a different task because previous studies have shown that even a dramatically different secondary task prevents learning (Fisk & Schneider, 1984; Nissen & Bullemer, 1987). It is possible that a secondary task more similar to the tapping task would prevent learning of the structured sequence. For example, a secondary task that requires processing of spatial stimuli might block the learning of spatial structured sequence. Such a finding would be of great importance because it might shed light on the processes involved in learning the structured sequence. Experiments in which the similarity between the secondary and the primary tapping task is manipulated are currently being conducted in our lab.

Awareness and Learning

Another indicant of the effectiveness of the distraction task was whether or not subjects were able to explicitly generate the structured sequence after the learning phase. Overall, the results suggest that most subjects were not aware of the presence of the structured sequence. Furthermore, to the extent that some subjects were aware of the structured sequence, this awareness had no effect on learning the structure. Subjects in Experiment 1 who got the structured sequence under the difficult distraction condition learned the structure. Their performance on the generation task, however, was no better than those who had previously responded to random events. Moreover, there was no difference in learning the structure as measured implicitly by reaction time between the easy and the difficult tone conditions of Experiment 1, but there was a difference between the two groups in the generation task. Experiment 4 and the supplement to that experiment showed that when the secondary task was removed during training, there was at least a marginal increase in awareness of the structured sequence. These results imply that awareness of the structure was impaired under dual-task conditions. Indeed, when subjects were directly asked whether they had noticed a structure or not, the majority who had had dual-task conditions indicated no awareness. In addition, when subjects in Experiment 4 were split into two groups on the basis of percent correct on the generation task, there was no difference between the groups in the magnitude of sequence learning as expressed in reaction times. Overall, the results appear quite convincing that explicit knowledge did not play a role in learning the structured sequence.

Why Does Distraction Block Learning of Ambiguous Structures?

At the descriptive level, the distraction task prevents learning of structured sequences in which there are no unique pairwise associations within a cycle. Distraction does not, however, block learning of other structures in which at least some associations are unique. What can account for such results? Here we entertain three possibilities. The first is that the ambiguous structures are generally more difficult than the others. The second is that a purely associative mechanism is able to learn structures with unique associations, but ambiguous structures require hierarchic coding. The third is that all structure types are learned by the same basic mechanism, a hierarchic mechanism, but that distraction makes it more difficult to find an appropriate parsing for ambiguous structures. Let us deal with each of these in turn.

Difficulty. The argument might be raised that the unique and hybrid structured sequences are in some way less difficult than the ambiguous structures. There are at least three problems with this argument. First, the label *difficult* in no way specifies what it is about the ambiguous structures that makes them difficult. The structures are of approximately equal length and share some of the same irregularity of hybrid structures. To apply the criticism, a theory of what makes particular structures difficult is needed. The two following hypotheses attempt to provide such an account. Second, there

is no indication under dual-task conditions that subjects became more aware of unique and hybrid structures than of ambiguous structures. Third, Figures 5 and 8 indicate that, after dual-task performance, subjects' rate of explicitly learning the structured sequences is similar for the unique, hybrid, and ambiguous conditions. The difficulty in learning ambiguous structures occurs under conditions of distraction. A theory must provide an account of why difficulty differs between explicit and implicit learning.

Associational and hierarchic coding. The theory motivating the present experiments posits two distinct learning mechanisms. One mechanism, which presumably can operate under distraction, forms associations between adjacent items. It would allow learning of structures with unique associations, but it would have difficulty when all of the items repeat in a sequence with different pairwise associations. The latter ambiguous structures may require a different mechanism for their learning, one which builds a hierarchic code. The essential idea is that ambiguous structures need to be parsed, with a higher level description provided for each part. The higher level description provides an additional cue that constrains the order of items that it controls. An explicit computational model of such a process has been proposed by Jordan (1986a, 1986b; in press) and Jordan and Rosenbaum (in press).

A critical issue concerns how the sequence becomes parsed and why the distraction task interferes with parsing. The parsing may depend on a short-term memory process that temporarily retains the preceding items in a sequence. The short-term memory may provide a basis for recognition of sequence parts that occur in the same order on other occasions. Such recognition constitutes a parse of the structured sequence because the recognition portion has a start and an end that separates it from adjacent parts. The parse then allows a control code to be established for the subsection of the structure. If the conjecture is correct, a possible reason that a secondary task interferes with learning the ambiguous structures is that distraction interferes with the short-term memory that is responsible for recognizing recurring subsections. The system that recognizes repetitions of identical structure parts need not involve *explicit* awareness, however. Experiment 4 makes the point that although removing a distraction task allows learning of ambiguous structures, such removal does not necessarily make people aware of that structure, although such awareness becomes more likely.

Although this theory can account for the results, there is no direct evidence supplied by the present studies for hierarchic coding of the ambiguous structures when learned under single-task conditions and for associational coding of the unique and hybrid structures. A large body of evidence, partly reviewed in the introduction, does indicate that structures similar to the ambiguous ones studied here are indeed coded hierarchically under conditions of explicit learning (e.g., Povel & Collard, 1982; Rosenbaum et al., 1983). Hierarchies are inferred in those studies from the patterns of errors and reaction times that demonstrate that the structured sequences have been parsed into chunks.

In principle, the error and reaction time analyses could be applied to the sequences of the present study, but in practice our procedures make such analysis difficult. In this study,

structured sequences started at random positions on different trials, and the structure continued for 20 cycles. Under such circumstances it is very likely that different subjects will parse the structures in different ways and thus obscure any consistent parsing pattern. One coding solution for a structure like 132312 is to decompose it into chunks of 132 and 312; other decompositions could be 121 and 323, or 13 and 2312, or any of a variety of others. Furthermore, in our studies there was always a pause of 200 ms between the response and the next item. Such a pause may obscure effects of chunking on reaction time and error patterns.

Likewise, there is no direct evidence from the present study that the unique and hybrid sequences, when learned concurrently with the distraction task, are coded associatively. In similar work by Keele and Summers (1976), evidence was reported for associative mechanisms even when some items in a structure were repeated. However, it is difficult to rule out the possibility that hierarchic coding, as well as associational coding, was employed for the unique and hybrid sequences under conditions of distraction.

Thus, for the present the theory of separable learning systems remains a hypothesis that is supported by the outcome of a priori predictions, but it lacks direct test.

Codes based on unique associations. Another possible theory that might explain the present results is based on the premise that all structured sequences, whether learned under distraction or not, are coded in hierarchic form. The basis for parsing the structures differs in the different cases, however, and distraction has a differential effect on those bases.

Parsing is based on recognition, not necessarily conscious, of passages that recur in the structure. The presence of a unique association provides a salient cue that the sequence will follow the same order as before. That is, the unique associations may serve as a sort of flag that defines the start of a structure. The flag also defines the end of the preceding series. Such flagging is equivalent to a parse in a model like that of Jordan (1986a, 1986b; in press). When a structure has only ambiguous associations, then the parsing mechanism must note that series of events occur in the same order on different occasions. Such recognition again defines the start and end of substructures to which a control code can be applied.

The present experiment provides no direct evidence for the hypothesis of two modes of hierarchic parsing. Because previous research has indicated the presence of two learning mechanisms (Keele & Summers, 1976), associative and hierarchic, we favor this hypothesis over the hypothesis of hierarchical parsing for all structured sequences. Nevertheless, both hypotheses may be viable, and they are in need of further testing.

Two approaches to the problem are currently under investigation in our laboratory. One approach makes use of neurologically impaired patients. If it is the case that there are separable systems responsible for hierarchic and associational coding, we may be able to dissociate them in patients. A second approach makes use of more formal theorizing. We have suggested that some structured sequences—unique and hybrid—can be learned by an associational mechanism, and other structures—ambiguous—require hierarchic representa-

tion. To formally test the view requires a precise computational model of both association and parsing. The outlines of such a model have been provided by Jordan (in press). We are currently examining the degree to which extensions of his model duplicate the performance on unique, hybrid, and ambiguous structures that we have found with human subjects.

References

- Attneave, F., & Benson, B. (1969). Spatial coding of tactual stimulation. *Journal of Experimental Psychology*, 81, 216–222.
- Bernstein, M. (1947). [On the formation of movement]. Published in Russian by State Medical Literature Publishers (Medgiz).
- Fisk, A. D., & Schneider, W. (1984). Memory as a function of attention, level of processing, and automatization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 181–197.
- Gordon, P. C., & Meyer, D. E. (1987). Control of serial order in rapidly spoken syllable sequences. *Journal of Memory and Language*, 26, 300–321.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4, 11–26.
- Hillis, A. E., & Caramazza, A. (1988). *The graphemic buffer and attentional mechanisms* (Reports of the Cognitive Neuropsychology Laboratory: Rep. No. 30). Baltimore, MD: Johns Hopkins University.
- Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 45, 188–196.
- Jordan, M. I. (1986a). Attractor dynamics and parallelism in a connectionist sequential machine. *Eighth Annual Conference of the Cognitive Science Society* (pp. 531–546). Hillsdale, NJ: Erlbaum.
- Jordan, M. I. (1986b). *Serial order: A parallel distributed processing approach* (ICS Report 8604). La Jolla, CA: Institute for Cognitive Science, University of California at San Diego.
- Jordan, M. I. (in press). Learning to articulate: Sequential networks and distal constraints. In M. Jeannerod (Ed.), *Attention and performance XIII*. Hillsdale, NJ: Erlbaum.
- Jordan, M. I., & Rosenbaum, D. A. (in press). Action. In M. I. Posner (Ed.), *Handbook of cognitive science*. Cambridge, MA: MIT Press.
- Keele, S. W. (1981). Behavioral analysis of movement. In V. Brooks (Ed.), *Motor control: Handbook of physiology: Section 1, Volume 2* (pp. 1391–1414). Bethesda, MD: American Physiological Society. (Distributed by Williams & Wilkins, Baltimore.)
- Keele, S. W. (1986). Motor control. In L. Kaufman, J. Thomas, & K. Boff (Eds.), *Handbook of perception and performance* (pp. 30.1–30.60). New York: Wiley.
- Keele, S. W., Cohen, A., & Ivry, R. I. (in press). Motor programs: Concepts and issues. In M. Jeannerod (Ed.), *Attention and performance XIII*. Hillsdale, NJ: Erlbaum.
- Keele, S. W., & Summers, J. G. (1976). Structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.
- Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading, reconsidered. *Journal of Experimental Psychology*, 100, 368–374.
- Lewicki, P., Hill, T., & Bizot, E. (1988). Acquisition of procedural knowledge about a pattern of stimuli that cannot be articulated. *Cognitive Psychology*, 20, 24–37.
- MacKay, D. G. (1982). The problem of flexibility and fluency in skilled behavior. *Psychological Review*, 89, 483–506.
- Margolin, D. I. (1984). The neuropsychology of writing and spelling: Semantic, phonological, motor, and perceptual processes. *Quarterly Journal of Experimental Psychology*, 36A, 459–489.

- Mishkin, M., & Appenzeller, T. (1987). The anatomy of memory. *Scientific American*, 256, 80-89.
- Mishkin, M., Malamut, B., & Bachevalier, J. (1984). Memories and habits: Two neural systems. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 65-76). New York: Guilford Press.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19, 1-32.
- Nissen, M. J., Knopman, D. S., & Schacter, D. L. (1987). Neurochemical dissociation of memory systems. *Neurology*, 37, 789-794.
- Peterson, L. R., & Peterson, M. (1959). Short-term retention of individual items. *Journal of Experimental Psychology*, 58, 193-198.
- Pew, R. W. (1974). Levels of analysis in motor control. *Brain Research*, 71, 393-400.
- Posner, M. I. (1984). Selective attention and the storage of information. In G. Lynch, J. L. McGaugh, & N. M. Weinberger (Eds.), *Neurobiology of learning and memory* (pp. 89-101). New York: Guilford Press.
- Povel, D.-J., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica*, 52, 107-123.
- Raibert, M. H. (1977). *Motor control and learning by the state space model*. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General*, 113, 372-393.
- Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 86-102.
- Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 501-518.
- Squire, L. R. (1986). Mechanisms of memory. *Science*, 232, 1612-1619.
- Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117-152). New York: Academic Press.
- Tolman, E. C. (1984). Cognitive maps in rats and men. *Psychological Review*, 55, 189-208.
- Wallace, R. J. (1971). S-R Compatibility and the idea of a response code. *Journal of Experimental Psychology*, 88, 354-360.
- Wickens, D. D. (1943). Studies of response generalization in conditioning: I. Stimulus generalization during response generalization. *Journal of Experimental Psychology*, 33, 221-227.
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1047-1060.
- Wright, C. E. (in press). Generalized motor programs: Reevaluating claims of effector independence. In M. Jeannerod (Ed.), *Attention and performance XII*. New York: Erlbaum.

Received January 3, 1989

Revision received March 16, 1989

Accepted April 10, 1989 ■

Call for Nominations for *Psychological Bulletin*

The Publications and Communications Board has opened nominations for a new editor of *Psychological Bulletin*. John C. Masters is the immediate past editor. Candidates must be members of APA and should be prepared to begin receiving manuscripts by September 1990. Please note that the P&C Board encourages participation by members of underrepresented groups in the publications process and would particularly welcome such nominees. To nominate candidates, prepare a statement of one page or less in support of each candidate. Submit nominations to

Walter Mischel, PhD
Psychological Bulletin Editor Search Committee
 American Psychological Association
 1200 Seventeenth Street, N.W.
 Suite VA-710
 Washington, DC 20036

First review of nominations will begin March 1, 1990, although the nomination deadline is March 31, 1990.