CHAPTER
31

The Organization of Sequential Behavior

Conditioning, Memory, and Abstraction


Abstract
Recent work suggests that sequential behavior depends on multiple concurrent processes including conditioning, memory, and abstraction. To determine whether multiple behavioral processes act concurrently to produce sequential behavior, we study forms of sequential behavior that are sufficiently complex that they are likely to recruit multiple behavioral processes concurrently. Evidence from both behavioral and neurobehavioral studies is presented that supports the view that, in addition to encoding information about the nature of sequential items, rats also concurrently abstract rules from hierarchical patterns and interleaved patterns, employ chunking processes that are sensitive to phrasing cues, and encode information about the serial position of sequential events. Thus, evidence supports the view that rats concurrently track multiple interoceptive, exteroceptive, and cognitive sources of information to organize their behavior through time.

Keywords: sequential behavior, sequential learning, serial-pattern learning, rule learning, associative learning, chunking, phrasing, hierarchical pattern learning, interleaved pattern learning, concurrent cognitive processing, MK-801, dizocilpine, adolescent nicotine exposure, rats, mice

A fundamental feature of behavior is that it occurs sequentially in time. Likewise, events are often arranged so that an organism must learn ordered relationships among them if meaningful predictions are to be made about the occurrence of future events. Thus, a classic problem has been the nature of the mechanisms mediating sequential learning and behavior. In recent years, it has become clear that humans and other animals have much in common in terms of sequential behavior and the processes that seem to be responsible for sequential learning and behavior (Fountain & Rowan, 1995b; Kesner, 2002; McGonigle & Chalmers, 2002; Sands & Wright, 1980; 1982; Terrace & McGonigle, 1994). However, rather than a consensus emerging regarding the nature of the mechanisms responsible for sequential behavior, competing hypotheses and theories to describe sequential behavior have multiplied. No general process theory is currently adequate to describe all sequential learning phenomena. We propose that, in sufficiently complex paradigms, sequential learning involves multiple, concurrently active behavioral and brain processes.

Current Models of Sequential Behavior
For decades, beginning with Lashley’s (1951) celebrated paper or even earlier with the work of Ebbinghaus and others (Hunter, 1920; Skinner, 1934), sequential learning and memory research has repeatedly aroused debate over the fundamental nature of sequential learning, memory, and representation in humans and animals. In recent years, debate has centered on whether nonhuman animals can use nonassociative symbolic processes such as rule induction to learn about the structure of patterned sequences (“serial patterns”) of events. The work by Hulse et al. on rats’ serial pattern learning of food reward magnitude in runways, in particular,
supported a rule-learning theory (Hulse, 1978, 1979; Hulse & Campbell, 1975; Hulse & Dorsky, 1977; Fountain & Hulse, 1981; Fountain, Evensen, & Hulse, 1983). Rule-learning theory proposed that rats learned some representation of the abstract rules that described organized sequences (Fountain, 1986; Hulse & Dorsky, 1977; 1979; Roitblat, Pologe, & Scopatz, 1983; Wathen & Roberts, 1994). The implication was that rats did not have to rely on chaining or remote associations alone to master sequences. Later work in our laboratory involved a somewhat different serial pattern learning paradigm employing patterns of stimuli drawn from another stimulus dimension (viz., a spatial dimension rather than food reward magnitude). This work also supported the rule-learning view of serial pattern learning in both rats and mice (e.g., Fountain, 1990; Fountain, Krauchunas, & Rowan, 1999; Fountain & Rowan, 1995a, b).

Very early on, however, Capaldi et al. contested the rule-learning view, proposing as a competing view that serial pattern learning in rats could be accounted for by associative mechanisms alone (Capaldi & Molina, 1979; Capaldi, Nawrocki, Miller, & Verry, 1985; Capaldi, Nawrocki, & Verry, 1982; Capaldi, Verry, & Davidson, 1980a, b; Capaldi, 1986). Capaldi likened sequential learning to other discrimination learning problems. However, as Capaldi himself admitted at the time (Capaldi et al., 1980a), his "sequential memory" model lacked sufficient parametric power to make strong predictions regarding the outcome of sequential learning tasks involving long patterns like those used by Hulse et al. Adding to the complexity of the issue, the rule-learning view frequently failed to predict outcomes for shorter patterns; even Hulse (1980) conceded that Capaldi's sequential memory theory was a better account in that domain. According to Hulse, short patterns approximate the paradigm of paired-associate learning rather than that of sequential learning.

Roitblat, Pologe, and Scopatz (1983) proposed yet another view: namely, that serial position played an important role in rat serial pattern learning. Serial position models assume that sequence elements become associated with their position in the sequence, not with other sequence items (Roitblat et al., 1983; Chen, Swartz, & Terrace, 1997; Burns & Gordon, 1988; Burns, Hulbert, & Cribb, 1990). Although serial position models may have much in common with Capaldi's sequential memory view, the critical difference is that sensitivity to serial position, the central construct here, seems to imply the additional cognitive ability of counting or timing the serial pattern events.

Although the foregoing models were originally proposed by some as mutually exclusive accounts of sequential learning and memory, recent behavioral and psychobiological work has produced evidence to support all of these positions, sometimes suggesting that rats may employ more than one strategy or type of information to encode and reproduce a single complex behavioral sequence (Fountain & Rowan, 1995a, b, 2000). Recent work in our laboratory involving behavioral, computational (Wallace & Fountain, 2002, 2003), and psychobiological approaches to understanding serial pattern learning support this view. Much of that evidence will be reviewed below. Our conclusion, for now, is that no general process theory is currently adequate to describe all serial pattern learning phenomena. Although this could be due to weaknesses in the available theories, we believe the evidence supports the idea that serial pattern learning actually involves multiple, concurrent behavioral and brain processes.

Consistent with this view is research from behavioral neuroscience and related fields providing neurobehavioral evidence for more than one process and more than one brain area mediating sequential learning. Nissen and Bullemer's (1987) paper on sequence learning has been particularly influential. Nissen's work on brain correlates of her human serial reaction time task (Knopman & Nissen, 1987; Nissen, Knopman, & Schacter, 1987) has supported the idea that serial learning is subserved by both declarative and nondeclarative memory systems. Nissen et al. showed that human Alzheimer's patients and scopolamine-treated experimental participants could improve their reaction times for a repeating 10-element pseudo-random response sequence (Knopman & Nissen, 1987; Nissen et al., 1987). Both groups showed no recognition that they were learning a repeating sequence, thus suggesting that learning could occur implicitly (or "procedurally"). Huntington's disease patients, however, exhibited no improvement; that is, they showed a deficit in serial learning described as a procedural memory deficit (Knopman & Nissen, 1991). Given that basal ganglia are severely affected by Huntington's disease, the results contribute to the growing evidence implicating the basal ganglia in sequence learning. For example, Huntington’s disease and Parkinson’s disease patients who suffer basal ganglia dysfunction have motor learning deficits that are characterized as deficits in sequencing and sequence.
learning independent of general motor performance deficits (Heindel, Butters, & Salmon, 1988; Willingham, 1998). In animal studies, injections of the dopaminergic antagonist haloperidol into the caudate nucleus of rats produces disruptions of “nonexteroceptively directed” motor sequences that were not cued by exteroceptive stimuli, but did not affect cued “exteroceptively directed” sequences of behavior (Jaspers, Schwarz, Sontag, & Cools, 1984). Lesions of hippocampus and medial caudoputamen produce a double dissociation of processes relevant to serial pattern learning (DeCoteau & Kesner, 2000), and hippocampal lesions have also been shown to cause deficits in rats’ ability to disambiguate sequences (Agster, Fortin, & Eichenbaum, 2002). In Fountain and Rowan (2000), we showed that MK-801 (also known as dizocilpine), an N-methyl-D-aspartate (NMDA) receptor antagonist, caused severe deficits in serial pattern learning in rats, but only for elements of serial patterns that interrupted or violated pattern structure. More recent work in our laboratory has indicated that these effects of MK-801 were not likely the results of impairment of dorsal hippocampus or medial prefrontal cortex, thus perhaps implicating subcortical structures. Heindel, Butters, and Salmon (1988) stated that “knowledge of the anatomical substrate underlying the acquisition of motor skills and other memory capacities preserved in amnesia is extremely limited.” Although this situation has improved in recent years (Graybiel, 1997, 1998; Willingham, 1998, 1999), it is possible that identifying the “critical” brain circuits involved in sequential learning and behavior has been difficult because, as the foregoing suggests, sequential behavior may be mediated by multiple brain systems and multiple concurrently active behavioral processes.

A Serial Multiple-Choice Task for Rats
If we are to determine whether multiple behavioral processes act concurrently to produce sequential behavior, then we will have to study forms of sequential behavior that are sufficiently complex that they would likely recruit multiple processes concurrently. Our serial pattern learning task for rats seems to be well suited for this purpose. The method we developed (Fountain et al., 2006) is a functional analog of nonverbal human pattern learning tasks that require participants to learn to choose items from an array in the proper sequential order (Hartman, Knopman, & Nissen, 1989; Knopman & Nissen, 1991; Reber, 1973, 1989; Restle, 1970; Restle & Brown, 1970b, c; Willingham, 1998; Willingham, Nissen, & Bullemer, 1989). In our task, rats learn to choose from a circular array of eight levers (or, in some studies, nose poke receptacles) in the proper sequential order to obtain reinforcement. The levers are designated Levers 1 through 8 in clockwise order with Lever 8 adjacent to Lever 1 (Fig. 31.1). All the levers are presented at the beginning of each trial, and the rat may press any of the eight levers. If the correct lever is pressed, then the rats receive hypothalamic brain-stimulation reward. If an incorrect lever is pressed, then all of the levers except the correct lever are removed from the box and the animal is not reinforced until the correct response is emitted. This method is easily learned by the rat without pretraining procedures other than leverpress shaping. It is an improvement over earlier methods used with rats because it allows us to study how rats learn long, elaborate serial patterns and because it provides measures of correct-response rates, error rates, and “intrusion” rates (i.e., the number of specific kinds of errors produced at particular locations in the pattern) on a trial-by-trial basis throughout the serial pattern.

With this method, we can create serial patterns with many items that could be associated, with spatial and temporal cues that could be relevant, with particular pacing or rhythmic structures, and with patterns of movements that could potentially be coded internally as motor patterns or as rule-based structures. Typically, many of these cues and features are concurrently available to the rat as the
sequence training takes place, and, as we shall show, it appears that rats concurrently make use of multiple sources of cues and behavioral processes to learn to navigate these serial patterns. Evidence to support this latter assertion comes from behavioral and drug studies to be presented below.

Rats Abstract Rules from Hierarchical Patterns

One prediction from the rule-learning view is that a highly organized, hierarchically structured sequence should be easier to learn than a sequence having little or no higher-order structure. We have already reported several studies that explored whether pattern structure determines the ease or difficulty of learning long and elaborate patterns. In one experiment (Fountain & Rowan, 1995a), we tested whether pattern structure described as “runs” (e.g., 1-2-3-4-5-...), or “trills” (e.g., 1-2-1-2-1-...), would determine the ease or difficulty of rats’ anticipating a final sequence item that either conformed to the implied structure of the sequence or violated pattern structure. Rats received patterns having either perfect structure or one sequence element (the last in the series) that violated an otherwise perfect structure:

Perfect Runs: 123 234 345 456 567 678 781 812
Violation Runs: 123 234 345 456 567 678 781 812
Perfect Trills: 121 232 343 454 565 676 787 818
Violation Trills: 121 232 343 454 565 676 787 812

Violation elements are underlined. A 1-second intertrial interval (ITI) was used except where spaces indicate 3-second phrasing cues. Rats experienced 20 patterns per day. We observed high error rates in acquisition on the violation trial (the last trial of the pattern) for both Violation Runs and Violation Trills patterns (Fountain & Rowan, 1995a), despite the fact that one view might predict that the generalization of associations from other parts of the pattern should have predisposed the animals to learn the violation patterns easily. For example, in the Violation Trills pattern, a correct response on Lever 1 should always predict that the next response should be to Lever 2; yet rats had great difficulty learning to respond on Lever 2 on the last trial of the pattern, but not on the second trial of the pattern. No comparable errors were observed for the Perfect Runs or Perfect Trills patterns. We concluded that rats abstracted the rules that describe the highly organized structure of the sequence even when doing so produced errors at the violation element and even though those errors might have been avoided by adopting another “simpler” associative strategy. CF1 mice show the same pattern of results as rats when learning the perfect and violation runs patterns described here (Fountain et al., 1999).

In another set of studies (Fountain & Rowan, 1995b), we tested whether pattern structure would determine the ease or difficulty of pattern learning by developing patterns with hierarchical structure, then reordering chunks of the pattern to produce “linear” structure, that is, a sequence of unrelated chunks. The Hierarchical (H) and Linear (L) patterns were:

H Pattern:
123 234 345 456 567 678 781 812
L Pattern:
123 234 345 456 567 678 765 654 432

For both groups, the digits indicate the clockwise position of the correct response on successive trials, the spaces indicate brief pauses, and underlined digits in the L pattern indicate where two chunks were exchanged.

As shown in Figure 31.2, the completely nested H pattern is described by a simple hierarchical rule structure: elements within three-element chunks are related by first-order rules, chunks within the first and second halves of the pattern, respectively, are related to each other by second-order rules, and the first half of the pattern is related to the second half of the pattern by a third-order “mirror” rule.

The incompletely nested L pattern was generated by exchanging the two underlined three-element chunks in the H pattern. In so doing, however, it should be noted that all pairwise associations were maintained; rats were always required to press a lever immediately to the left or right of the last correct response in both patterns, and the number of transitions from a given lever to any other was conserved across patterns. In this structure, elements within any chunk are related by a rule, but chunks are not related to each other in any systematic way.

The results showed that, for rats, pattern complexity predicted pattern learning difficulty (Fountain & Rowan, 1995b). The formally simpler H pattern was easier to learn than the formally more complex L pattern. In addition, as shown in Figure 31.3, rats in H were sensitive to the hierarchical structure of their pattern. For rats in the H pattern groups, as in humans in an analogous task (Fountain & Rowan, 1995b), the difficulty of learning to respond appropriately on any trial was a function of the hierarchical level of the rule required to predict the item. Rats produced significantly more errors on the first
of sequential behavior.

598  THE ORGANIZATION OF SEQUENTIAL BEHAVIOR

1 trial of Chunks 1 and 6 than on all other trials. These trials corresponded to the highest-order rule transitions in the pattern structure (i.e., third-order rule transitions). Fewer errors were observed on the first trial of other chunks; these trials corresponded to second-order rule transitions. The fewest errors occurred within chunks, where trials corresponded to first-order rule transitions. Thus, in the completely hierarchical pattern, the difficulty of learning to respond appropriately on any trial was a function of the hierarchical level of the rule required to predict the item. Rats in L did not show the three-level hierarchical pattern of errors observed for H rats.

In the hierarchical versus linear structure experiment just described, rats demonstrated sensitivity to multilevel hierarchically organized pattern structure. Rats found learning completely nested hierarchical patterns easier than learning less organized patterns, even when pairwise associations and pattern length were conserved across patterns. In another study from the same series (Fountain & Rowan, 1995b), a three-level hierarchy was easier to learn than a four-level hierarchy when pattern length was conserved across patterns. As a rule, then, pattern complexity was a better predictor of acquisition difficulty in these studies than was pattern length. These acquisition results alone are strong evidence that pattern organization, that is, pattern complexity, was the primary determinant of pattern difficulty, as argued by a rule-learning view of sequential learning.

Rats Abstract Rules from Interleaved Patterns

One question of significance for animal sequential learning research is whether animals are constrained to learn sequences on the basis of pairwise associations between successive elements, for example, as in chaining (Skinner, 1934). A significant body of evidence suggests that animals are able to be more flexible in representing sequential events, conceivably by coding hierarchical representations characterized by relations for nonadjacent events (Roitblat, Bever, Helweg, & Harley, 1991; Roitblat, Scopatz, & Bever, 1987; Roitblat et al., 1991). The mechanisms of learning involving nonadjacent events are not well-understood. Terrace (1987), for example, indicated that little evidence existed that animals are able to spontaneously reorganize sequentially presented items into chunks not presented by the experimenter. As Terrace noted, such processes are readily observed in human free recall (Tulving, 1962). Additionally, it should be noted that chunking of nonadjacent items in human serial pattern learning has been studied extensively using patterns of letters and digits (Hersh, 1974). We have previously shown that rats, when presented a sequence of
reward quantities, can spontaneously sort quantities from nonadjacent serial positions into chunks to facilitate learning (Fountain & Annau, 1984). A comparable strategy in humans would be to learn the pattern 255545565558 by sorting pattern elements into 555 chunks and a 2468 chunk. Other work also supports the view that rats have this capacity. For example, Capaldi and Miller (1988) have shown that rats can keep count of different kinds of rewards by chunking nonadjacent items in series into different food categories. In two recent studies in our laboratory (Fountain, Rowan, & Benson, Jr., 1999), rats learned either a structured (ST) or unstructured (UNST) sequence interleaved with elements of a repeating (R) sequence in one experiment or an alternation (A) sequence in another experiment. The question was whether rats would learn the interleaved subpatterns at different rates as a function of subpattern complexity.

A first experiment sought to determine whether rats would show signs of being sensitive to the organization of nonadjacent items from interleaved subpatterns when one subpattern was composed of simple, repeating elements and the second subpattern...
was either highly structured or not. For rats in the structured (ST) subpattern condition, a 123 234 345 456 567 subpattern was interleaved with a repeating (R) subpattern, 888 888 888 888 888, resulting in the ST-R pattern that rats were required to learn:

$$182838 283848 384858 485868 586878$$

For rats in the unstructured (UNST) subpattern condition, a 153 236 342 547 subpattern was interleaved with the same R subpattern to create the UNST-R pattern in the same manner. For both patterns, integers represent the clockwise position of correct levers in the octagonal chamber on successive trials and spaces represent pauses that served as phrasing cues.

Acquisition of the interleaved structured pattern (i.e., ST-R) was significantly faster than for the interleaved unstructured pattern (i.e., UNST-R). The unstructured pattern was generated by exchanging only two pairs of elements in the structured pattern, as described above. In so doing, however, all pairwise associations in the interleaved patterns were maintained because all of the relocated items were preceded by “8” trials. Nevertheless, as shown in Figure 31.4, the effects of disrupting pattern structure were apparent throughout the pattern.

UNST-R rats’ performance was poorer even in the third (middle) chunk that was not altered in producing the unstructured pattern; rats found this chunk, 38458, harder to learn in the context of the UNST-R pattern than in the ST-R pattern.

In the second experiment, rats learned two interleaved sequences, where both were created from sets composed of more than one element. As before, longer patterns were composed of two interleaved subpatterns; either a structured or unstructured subpattern was interleaved with a subpattern of two alternating elements. For one group of rats, the structured (ST) subpattern, 123456, was interleaved with the alternating (A) subpattern, 787878, to create the ST-A pattern, 172837485768. For another group of rats, the unstructured (UNST) subpattern, 153426, was likewise interleaved with the same alternating subpattern to produce the UNST-A pattern. Note that the unstructured subpattern was generated by exchanging two items of the structured subpattern. Rats learned the subpatterns of their interleaved patterns at different rates both within and between pattern groups. As predicted based on subpattern structure, in the case of the UNST-A pattern, the A subpattern was acquired faster than the UNST subpattern. The A subpattern would be expected to be acquired faster because it is

![Figure 31.4](image-url)

**Figure 31.4** Group mean element-by-element errors for the interleaved structured-repeating (ST-R) and unstructured-repeating (UNST-R) patterns averaged across week 3 of training (from “Rule learning in rats: Serial tracking in interleaved patterns,” by S. B. Fountain, J. D. Rowan, & D. M. Benson, Jr., 1999, *Animal Cognition*, 2, pp. 41–54. Copyright 1999 by Springer-Verlag. Adapted with permission).
formally simple, whereas the UNST subpattern has little structure (Hulse, 1978; Hulse & Dorsky, 1977; Jones, 1974). Based on similar reasoning, it was expected that the ST subpattern should be easier to learn than the UNST subpattern; this result was obtained. In the case of the ST-A pattern, because both subpatterns were structured, it might be difficult to predict in advance based on subpattern structure alone whether rats should find either the ST or A subpattern easier to learn than the other. However, if structural complexity is equated (i.e., if the same number of rules is needed to describe subpattern structure), then rats might show the same predisposition that humans do (Kotovsky & Simon, 1973) to detect repeating items before other structural features of patterns. In fact, evidence for the latter assertion was obtained in this experiment. Rats in the ST-A pattern group showed better acquisition for A with its repeating “7” and “8” elements than ST subpatterns of their interleaved pattern despite the fact that both ST and A subpatterns have simple structure that can be described by a single rule (viz., a “+1” rule for the 123456 ST subpattern vs. an “alternate” rule for the 787878 A subpattern). The results of differential acquisition of ST and UNST subpatterns support the notion that accurate performance on these interleaved subpatterns was dependent on a mnemonic representation characterized by relations for nonadjacent events (Roitblat et al., 1987, 1991). The results indicate that rats are sensitive to the organization of nonadjacent elements in serial patterns and that they can detect and sort structural relationships in interleaved patterns. Pattern and subpattern structure appear to drive how animals sort, chunk, and represent nonadjacent pattern elements that are related by common rules or features.

Phrasing Cues, Chunking, and Overshadowing

Fountain, Henne, and Hulse (1984) argued that phrasing cues speed learning by facilitating chunking vis-à-vis pattern structure. Rats were trained to anticipate food quantities presented in a series of trials in a runway, such as in the five-element pattern, 14-7-3-1-0. Structurally, this pattern is highly organized; that is, it is formally simple, because successive elements can be described by a single “less than” rule (Hulse, 1978). A longer pattern was produced by presenting a 14-7-3-1-0 subpattern of food quantities five times in succession. Fountain et al. (1984) reported that, when spatial or temporal cues were placed congruent with boundaries between subpatterns, the cues facilitated pattern learning (Fountain, Henne, & Hulse, 1984). This result tended to support the notion that phrasing facilitated learning by highlighting salient features of pattern structure rather than by cueing specific responses. However, one distinction between spatial and temporal cues was noted. When phrasing cues were removed, rats trained with spatial cues alone showed savings relative to controls that had never experienced cues. In contrast, rats trained with temporal cues alone or in combination with spatial cues showed no savings after cue removal, suggesting that spatial and temporal cues may have facilitated learning via different mechanisms (Fountain et al., 1984).

Capaldi, Verry, Nawrocki, and Miller (1984) replicated the effects of spatial cues reported by Fountain et al. (1984), but argued that phrasing effects, like other aspects of serial pattern learning, should be explained by appealing to traditional associative processes rather than rule-learning processes.

Echoing these earlier ideas, in Stempowski et al. (1999), we argued that discriminative cues might serve at least two functions in serial learning. One possibility is that temporal intervals signal or guide responses as traditional discrimination learning theory suggests (Capaldi, Verry, Nawrocki, & Miller, 1984). In the case of temporal intervals inserted into patterns at chunk boundaries, the intervals presumably would overshadow the naturally occurring interim associations and would become the principal signal of the correct response or the next item in the trial following the temporal pause (Capaldi et al., 1984). If this were the effect of inserting temporal cues into sequences, then rats should become dependent on these cues for producing the cued responses and should fail to respond correctly if the cues are removed. A second hypothesis is that temporal cues bias the perception of pattern organization (Fountain et al., 1984; Fountain, Rowan, & Carman, 2007). According to this view, temporal cues identify salient features of pattern structure, for example, by indicating transitions between chunks. According to the hypothesis that intervals serve as cues for structure, the differential intervals located at chunk boundaries simply facilitate encoding pattern structure and should thus result in savings in pattern tracking even after phrasing cues are removed (for a similar argument, see Fountain, Henne, & Hulse, 1984).

Consistent with the views of Capaldi et al., two experiments in our laboratory on the effects of temporal phrasing in rat pattern learning showed that temporal intervals positioned at chunk boundaries...
facilitated pattern learning by serving as discriminative cues that overshadowed associations between sequence elements (Stempowski, Carman, & Fountain, 1999). One experiment showed that rat serial pattern learning could be facilitated when distinct temporal intervals preceded chunk boundaries regardless of whether the intervals were longer or shorter than the intervals within chunks. A second experiment replicated the acquisition results of the first with a different, more difficult serial pattern. In addition, after both 14 and 35 days of acquisition with phrasing cues, cue removal produced severe deficits in tracking the first element of chunks, the element directly after the phrasing cues during acquisition (Stempowski et al., 1999). The results indicated that rats used both short and long temporal phrasing intervals as discriminative cues, and that facilitated learning due to phrasing is not the result of additional processing time provided by longer intertrial intervals at chunk boundaries (Terrace, 1991). Furthermore, many of the finer details of the results could be accounted for by the additional assumption that phrasing cues overshadowed interitem associations. These results support the associative view that phrasing cues serve as discriminative cues that control responses on the trial after the cue on the first trial of chunks. However, this conclusion does not rule out the possibility that phrasing cues might play other roles in sequential learning under other conditions.

**Serial Position, Item Memory, and “Instrumental Blocking”**

As we have indicated, violation elements (elements that violate pattern structure) are particularly difficult for rats to learn, although eventually they can be learned to a high level of accuracy. In a recent series of studies (Muller & Fountain, 2010), we investigated what cues contribute to accurate anticipation of violation elements in serial patterns. Groups of rats learned two-level hierarchical sequences containing a violation. For some groups, sequences started on different levers of the chamber on a pattern-by-pattern basis, and patterns were separated by a long (9-second) interpattern interval to mark the beginning of patterns. For the Serial Position (SP) group the violation element (underlined in patterns that follow) was always the last (24th) element of the pattern no matter which lever was randomly chosen as the starting lever:

Starting Lever 1: 123 234 345 456 567 678 781 81
Starting Lever 2: 234 345 456 567 678 781 812 12
Starting Lever 3: 345 456 567 678 781 812 123 23 etc.

In the Location (L) group, the violation always occurred at the same position in the chamber but occurred at different serial positions depending on the randomly chosen starting lever:

Starting Lever 1: 123 234 345 456 567 678 781 81
Starting Lever 2: 234 345 456 567 678 781 812 12
Starting Lever 3: 345 456 567 678 781 812 123 23 etc.

Performance for these groups was compared to that of a group where both serial position and location cues were relevant and consistent (SP+L):

Group SP+L pattern: 123 234 345 456 567 678 781 81

A No Cues (NC) group learned patterns where neither serial position nor location cues were relevant:

Starting Lever 1: 123 234 345 456 567 678 781 81
Starting Lever 2: 234 345 456 567 678 781 812 12
Starting Lever 3: 345 456 567 678 781 812 123 23 etc.

Rats experienced 24 pattern repetitions a day (3 each from each possible starting location in groups NC, SP, and L). Patterns were presented with 1-second ITIs for elements within chunks and 3-second ITIs (phrasing cues) at chunk boundaries (indicated by spaces in the above patterns) in the manner described by Fountain and Rowan (1995a). For groups that met a criterion of fewer than 10% errors on the violation, a series of 1-day transfers were conducted to determine what controlled responding to the violation element.

Acquisition results under the foregoing conditions were very clear. Rats in the two conditions with relevant item/place cues (viz., SP+L and L) learned the sequence equally quickly, with all rats reaching criterion for the violation element in 14 to 22 days. Rats with serial position cues alone (viz., group SP) never reached criterion. Rats in SP, like those with no cues in NC, mastered other elements of their pattern, but they continued to make approximately 100% errors on the violation element throughout 40 days of acquisition. We found no evidence that rats used serial position cues, either when presented alone or in combination with other cues, to anticipate the violation element when it was located at the end of the pattern in serial position 24.

In subsequent transfer phases, rats in SP+L and L groups experienced transfer conditions presented on separate days; transfer days were separated by retraining to criterion on the original pattern. On one day, rats in both conditions received serial
position cues only (i.e., 1 day of training under group SP conditions). Under these conditions, all rats failed to anticipate the violation element (making nearly 100% errors), indicating that training with serial position cues in SP+L did not give them an advantage in transfer over the group that never experienced consistent serial position cues for the violation element. On another transfer day, rats were trained in their usual chamber that had been rotated 180 degrees. Under these conditions, rats’ performance on the violation did not change, indicating that anticipation of the violation element was controlled in large measure by intra-chamber cues (perhaps cues from the floor, walls, levers, or associated apparatus) rather than by extra-chamber spatial cues. On another transfer day, rats were transferred to a new chamber in a different testing room. With both extra-chamber and intra-chamber cues removed, rats made nearly 100% errors on the violation element, but their performance was unchanged on all other elements of the pattern. Thus, intra-maze cues were important for anticipating the violation element, but they otherwise played little role in rats’ performing the sequential behavior. Finally, rats in the SP+L group were transferred for one day to L-only conditions (i.e., serial position cues were removed). No effects were observed, providing additional evidence that serial position cues were not important for anticipating the violation element.

It should be noted that having identified the relevant cues for the violation element as involving intra-chamber cues, we have still not identified the specific cue or set of cues that control the behavior. A close examination of the serial pattern that rats learned reveals that no single exteroceptive or interoceptive cue, except the cues that define serial position 24, reliably predicts the violation element. Rats encounter all of the intra-chamber and interoceptive cues surrounding the violation element more than once in the course of producing each serial pattern. Clearly, rats are faced with the problem of the same cues signaling different outcomes at different points in the sequence, creating a difficult “branching” problem (Restle & Brown, 1970a, b). Evidence from other studies indicates that stimuli from multiple trials leading up to the violation trial serve as compound or configural cues that come to control the violation response.

In a separate study, other groups of rats were trained with the violation element in serial position 6 or 12 (SP6 and SP12, respectively, where the violation element appears as the third element of Chunks 2 and 4, respectively, of the eight-chunk pattern). The results indicated that rats in SP6 learned to anticipate the violation, but only with great difficulty, requiring approximately 50 days to reach asymptote at approximately 25% errors. SP6 rats were further tested until Day 70, but they never improved. Rats in SP12 never diverged from 100% errors in 70 days of training (24 patterns per day). Clearly, rats have great difficulty using serial position as a cue for anticipating important sequential events such as violation elements, even when serial position is the only reliable cue for the event. Indeed, rats may not be able to do so at all for serial positions 12 or beyond.

In a recent study by Kundey and Fountain (2010) employing the pattern learning paradigm described above (Fountain & Rowan, 1995a, 2000), we tested for “instrumental blocking” of cues involved in violation element learning. Throughout the experiment, we eliminated periodicity (serial position) of the violation element as a cue. In Phase 1, rats were trained with a noise cue (stimulus A) that was the only consistent cue predicting one violation element per pattern, which could occur anywhere in the spatial array of levers. In Phase 2, spatial cues (stimulus B) were added when two violation elements per pattern were permanently located at two separate spatial locations within the chamber. Spatial cues were presented alone (i.e., B only) as a cue for one violation or in combination with the noise cue experienced in Phase 1 (i.e., A+B) as discriminative cues for another violation element.

To determine if noise cues in Phase 2 blocked learning to use spatial cues to anticipate violation elements when they were presented as compound cues with the noise cue, in Phase 3 we tested for instrumental blocking by presenting spatial cues alone. Phase 2 contained two violations in fixed spatial locations within the chamber. In this phase, one violation element in each pattern received blocking training, that is, it was signaled by the previously learned noise cue and the newly added spatial cues, whereas the other violation in each pattern did not receive blocking training, that is, it was signaled by new spatial cues alone. Figure 31.5 shows that during the last session of Phase 2, rats made few errors at either violation location. In Phase 3, rats continued to make few errors for the violation element that did not receive blocking training in Phase 2, but they made many errors at the violation element that did receive blocking training in Phase 2.

The results suggest associative learning mediated cuing of violation elements. Taken together with other evidence reported earlier implicating rule...
learning when rats learn this pattern or “runs” in this paradigm (e.g., Fountain & Rowan, 1995a), these results implicate at least two concurrent learning processes in rat serial pattern learning, namely, multiple-item memory and rule learning.

Concurrent Processes in Sequential Learning
Phrasing Cues Bias Rule Learning
In another octagonal chamber study, we examined the effects of phrasing a structurally ambiguous pattern as either a series of “runs” or “trills” (Fountain et al., 2007). A 123434567878 pattern phrased as runs (1234-3456-7878) was easier to learn than when it was phrased as trills (1212-34-3456-7878), a result that resembles a similar “run bias” reported in the human sequential learning literature and in earlier studies of rats’ response to “run” versus “trill” structures (Fountain & Rowan, 1995a). Whereas rats learning the runs-phrased pattern showed rapid learning and little tendency to make trills errors, rats learning the trills-phrased version of the pattern produced inflated rates of both trills and runs errors (Fig. 31.6; Fountain et al., 2007). The results show that rats represented the runs- and trills-phrased versions of the pattern differently. These results add to the evidence that, in addition to serving as discriminative cues as demonstrated by Stempowski et al. (1999), phrasing cues can bias pattern perception in rat serial pattern learning, resulting in memorial representations characterized by multiple interpretations of the same pattern.

Phrasing Cues: Serial Position Revisited
In a study designed to determine whether rats were sensitive to the serial position of phrasing cues, we trained rats on a 24-element sequence containing a violation element, then exposed rats to probe patterns with the position of some phrasing cues shifted (Muller & Fountain, 2010). The training pattern was: 123-234-34567-678124. Rats received 20 training patterns per day. When performance on the violation element reached a criterion of 10% or fewer errors in a daily session, rats were transferred to a phase where they experienced three probe patterns daily along with a “probe” instance of the original training pattern. The three probe patterns were randomly distributed in a normal day’s testing for 10 days. The transfer patterns manipulated both chunk length and overall pattern length. One transfer pattern was an 18-element pattern made up of chunks of varying lengths: 123-23-2345-4-34567-67-678-78-78124. Another transfer pattern was the same length as the training pattern, but contained chunks ranging in length from one to five elements: 12-123-2345-56-5678-7-678124. Yet another transfer pattern was a 30-element pattern made up of chunks of varying lengths from one to five elements: 123-23-2345-4-34567-67-678-78-78124.
two-element chunks, that is, shorter than three elements as trained, errors were high on the trial following the cue (the chunk boundary). When phrasing cues followed chunks that were three elements or longer, errors were low on the following trials. Thus, phrasing cues and serial position cues formed compound or configural cues. It is interesting to note that none of the transfer patterns resulted in poor performance on the violation element despite the fact that the violation element appeared in different serial positions in the 18- and 30-element probes. The latter result is consistent with work discussed earlier (Muller & Fountain, 2010) showing that serial position cues are not used by rats to anticipate a violation element positioned in serial position 24.

**Interleaved Patterns: Concurrent Item Memory and Rule Induction**

Fountain and Benson (2006) examined the influence of pattern structure on learning interleaved patterns containing violation elements. In one group, the first interleaved subpattern was a formally simple sequence, whereas in two other groups the first subpattern was formally more complex, containing two or four violation elements, respectively. One group learned an interleaved S-S serial pattern: 1526374851627384. This interleaved pattern is based on two formally simple (S) subpatterns, 12345678 and 56781234. Both subpatterns are considered formally simple because they can be described by a single rule, namely, a “+1” rule that indicates that on successive trials the rat must choose the lever to the right of the last correct lever. In fact, the subpatterns are structurally the same, differing only in terms of where the pattern begins in the array (Lever 1 vs. Lever 5 as starting lever). This interleaved pattern was called “S-S” to signify that it was composed of two subpatterns of simple structure. A second group learned an interleaved 2V-S serial pattern: 1526473851627384. This interleaved 2V-S pattern was composed of two subpatterns, one with 2 violation (2V) elements and the other with simple structure: 12435678 and 56781234. Note that the 2V subpattern contains two violation elements that break the +1 rule that describes the elements constituting the rest of this subpattern. The 2V subpattern was created by exchanging the “4” and “3” elements of the first subpattern of the S-S pattern. The second subpattern of 2V-S is identical to the second subpattern of the S-S pattern. A third group learned an interleaved 4V-S serial pattern: 1526473861527384. This interleaved 4V-S pattern

---

1 Results from this experiment (Fig. 31.7) indicated that when transfer chunks were longer than three elements, error rates generally increased on elements beyond the third. Thus, rats showed some evidence of anticipating a chunk boundary after three elements even on trials that did not follow phrasing cues. When phrasing cues followed one- or
was composed of two subpatterns, one with 4 violation (4V) elements and the other with simple structure: 12436578 and 56781234. The 4V subpattern was created by exchanging the “3” and “4” elements and the “5” and “6” elements of the first subpattern of the S-S pattern. The second subpattern of 4V-S is identical to the second subpattern of the S-S and 2V-S patterns.

The results showed that rats were sensitive to the relationships between elements of interleaved...
patterns even though the elements were not adjacent in series. Figure 31.8 shows rats’ group mean acquisition curves as day-by-day percent errors. The top and bottom panels parse the data by first and second subpatterns of the interleaved patterns, respectively. The data show that rats chunked patterns into component subpatterns, learned the "+1" rule relating nonadjacent elements of simple (S) subpatterns, and treated violation elements as structural imperfections. Evidence to support these claims for chunking and rule learning comes from (1) comparisons of pattern and subpattern acquisition rates and (2) intrusion error analyses, particularly for violation elements. Acquisition results

![Graph showing acquisition curves for the component first and second subpatterns (top and bottom panels, respectively) of S-S, 2V-S, and 4V-S over the 35 days of the experiment. The top panel shows acquisition for the first subpattern of each interleaved pattern and the bottom panel shows acquisition for the second subpattern. Percentages of daily mean errors were averaged across all elements of the patterns. Error bars: ± SEM (from “Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning,” by S. B. Fountain & D. M. Benson, Jr., 2006, Learning and Motivation, 37, pp. 95–112. Copyright 2006 by Elsevier. Adapted with permission).]
showed that the interleaved serial pattern, S-S, composed of two subpatterns of simple structure, was learned faster than interleaved serial patterns containing violation elements, as predicted by pattern structure. Further, in each interleaved pattern, subpatterns were acquired at different rates. Even though the two S subpatterns in the interleaved serial pattern S-S were essentially identical, rats nonetheless chunked elements with respect to constituent subpatterns, learning the first S subpattern faster than the second S subpattern. Similarly, with reference to the 2V-S and 4V-S interleaved patterns, the first subpatterns, which contained violation elements, were acquired faster than the second subpatterns, S, in both cases. When first subpatterns are considered, the formally simple S subpattern was easier to learn than formally complex 2V and 4V subpatterns that contained violations, as predicted by rule-learning theory. Taken together, these results strongly support the notion that rats were chunking patterns into their constituent subpatterns and that subpattern structure was a determinant of subpattern difficulty.

Evidence for chunking and rule learning also came from results of the intrusion error analysis. Figure 31.9 shows rats’ group mean element-by-element percent error rates collapsed across all days of the experiment. The top and bottom panels parse the data by first and second subpatterns of the interleaved patterns, respectively. In violation patterns (top panel), the highest-frequency errors on violation elements were responses consistent with the +1 rule describing simple subpattern structure (indicated by circles in the top panel for trials where this was true). Such responses fit well with predictions of rule-learning theory but are not easily explained by common associative mechanisms of discrimination learning.

If we ignore subpattern structure and treat the patterns as an undifferentiated series of events, we find strong evidence that sequential adjacent-element associations and remote associations also controlled rats’ sequential choices. Several of the easiest and most difficult elements for rats to learn could be accounted for by multiple adjacent-item associations (indicated by circles in the bottom panel of Fig. 31.9 for trials where this was true). When multiple cues consistently signaled the same correct outcome, the target elements were among the easiest to learn. In contrast, when the same cues signaled different outcomes at different points in the sequence, creating the difficult “branching” problem already discussed earlier, the target elements were among the most difficult to learn due to generalization of errors across instances where the same cues predicted different outcomes. This associative view, however, cannot account for the foregoing phenomena implicating chunking and rule learning. Instead, the results taken together fit better with the idea that rats used rule learning and associative learning concurrently.

Rats appeared to learn about the formally simple rules that described the underlying interleaved structures of the sequences and governed long strings of elements. They also used associative memory of multiple items as compound cues to anticipate specific pattern elements where those cues were particularly distinctive. Thus, the behavioral evidence suggests that rats used chunking, rule learning, and interitem association learning concurrently to master these complex patterns.

**Neurobehavioral Studies: Dissociations of Concurrent Processes**

Recently, we began to explore the neural basis of sequential learning and behavior in a systematic manner. Our first approach was to investigate the effects of drugs thought to have effects on learning by their impact on specific receptor systems. This approach had the advantage that we could choose drugs that could be injected systemically and thus avoid the problem of producing lesions in animals to be implanted with bipolar electrodes. Thus, in these studies rats learned leverpress response patterns for brain-stimulation reward. In one series of experiments (Fountain & Rowan, 2000), rats were trained on two patterns, one which was structurally “perfect” and a second virtually identical to the first but containing a single element that violated the otherwise simple structure. The Perfect (P) and Violation (V) patterns were:

- P pattern: 123 234 345 456 567 678 781 812
- V pattern: 123 234 345 456 567 678 781 81

As before, the digits indicate the reinforced lever for successive trials and spaces indicate the location of phrasing cues. The last “8” item of the V pattern (underlined) was the violation element. Rats from one group for each pattern condition were injected with 0.0625 mg/kg MK-801 daily before training. MK-801 (dizocilpine) is a systemically administered NMDA receptor antagonist that blocks neuronal plasticity in the hippocampus and other areas of the brain, including amygdala and basal ganglia. As shown in Figure 31.10, MK-801 had little effect on learning to respond to rule-governed elements.
within chunks. However, it did impair responding for the violation element and for cued responses at chunk boundaries. As-yet-unpublished work in our lab also shows that cholinergic blockade by systemic administration (i.p. injection) of atropine or scopolamine produces greatly elevated response rates on chunk boundaries and the violation element but has almost no effect on within-chunk elements, similar to the effects of MK-801. Other related work shows that exposure to nicotine during...
adolescence in rats causes later cognitive deficits in adult serial pattern learning with much smaller effects but with a pattern similar to those reported for MK-801 and anticholinergic drugs (Fountain, Rowan, Kelley, Willey, & Nolley, 2008). In our study investigating the effects of adolescent nicotine exposure on adult serial pattern learning, adolescent rats received daily i.p. injections of either 1.0 mg/kg nicotine or saline for 5 days per week for 5 weeks beginning on postnatal day 25 (P25), and then nicotine exposure was terminated and rats were allowed 35 days drug-free. Rats then began training on P95 as adults on a 24-element serial pattern composed of eight 3-element chunks with no violation element: 123-234-345-456-567-678-781-812. The results, shown in acquisition curves in terms of daily mean correct response rates in Figure 31.11, showed that adolescent nicotine exposure transiently retarded later adult learning for elements 2 and 3 of chunks (within-chunk elements) of the pattern that represented rule-governed elements of their pattern, but adolescent nicotine exposure did not affect asymptotic levels of acquisition for these elements of the pattern by the end of the experiment. However, the results clearly showed that adolescent nicotine exposure retarded later adult learning (i.e., it decreased daily mean correct responses in the figure) for element 1 of chunks (chunk-boundary elements) of rats' pattern that represented transitions between “chunks” of the pattern. The principal conclusion from this study was that nicotine that was experienced only in adolescence produced significant deficits in adult rat serial pattern learning even after five intervening weeks of nicotine-free recovery time, and the observed deficits were consistent with the view that multiple concurrent brain and behavioral systems are involved in sequential learning.

From both a learning theory perspective and a neural systems perspective, it is important to (1) determine whether multiple neural systems are recruited in sequential learning, as evidence seems strongly to indicate, (2) identify the neural systems involved, and (3) determine whether multiple neural systems are recruited concurrently, as we have proposed above for the concomitant behavioral processes. As for the behavioral processes, it is also important to identify factors that determine when these neural systems come into play, whether or not they are recruited concurrently. This information is critical for developing a coherent general model of behavioral and neural function in sequential learning.

Fig. 31.10 Effects of systemically injected MK-801 (dizocilpine) on serial pattern learning in rats. Mean pattern tracking errors averaged across 50 presentations of the violation pattern on Day 7, the last day of the experiment, for groups that were injected daily with either saline or MK-801. MK-801 profoundly retarded learning (observed here as increased errors) on elements characterized as transitions between phrases of the pattern and an “exception-to-the-rule,” namely, chunk-boundary elements (element 1 of each three-element chunk) and the violation element (the last element of the pattern), respectively. MK-801 also produced a small and persistent effect on asymptotic levels of learning rule-governed within-chunk elements, but initial acquisition of within-chunk elements was spared (from “Differential impairments of rat serial pattern learning and retention induced by MK-801, an NMDA receptor antagonist,” by S. B. Fountain & J. D. Rowan, 2000, Psychobiology, 28, 32–44. Copyright 2000 by the Psychonomic Society. Adapted with permission).

Animals as Multimodality

Sequential Processors

The general goal of the foregoing studies was to begin to identify which behavioral processes contribute to control of rats' responses in serial patterns of behavior on an element-by-element basis in order to develop a more complete picture of how sequential behavior is acquired, represented, and produced. The approach was to determine the extent to which rats' sequential behavior is controlled by extra-sequence stimuli through associative processes, such as temporal or spatial cues acting as discriminative cues; by some other mechanisms, such as internal motor patterns or cognitive structures; or by multiple processes acting concurrently. The behavioral evidence that we have accumulated strongly supports the view that rats can be described as multimodality sequential processors. Evidence supports the view that rats concurrently track multiple interoceptive, exteroceptive, and cognitive sources of information to organize their behavior through time.
The outcomes of these studies are relevant to current research problems in several domains. First, the results are important for building a better understanding of the mechanisms of serial pattern learning, because they begin the process of cataloging the conditions under which item associations, extra-sequence cues, and pattern structure individually or collectively control sequential behavior. This information will be critical in developing an integrative model of serial pattern learning. Second, if rats are concurrently using multiple cues to anticipate upcoming pattern elements, as suggested by

Fig. 31.11 Adult rats’ acquisition curves for elements 1, 2, and 3 of chunks (panels a, b, and c, respectively) over 21 days of training. Rats received prior adolescent exposure to either 1.0 mg/kg nicotine or saline during adolescence only. Error bars: ±SEM. Asterisks indicate significant differences relative to controls (p < 0.05) (from “Adolescent exposure to nicotine impairs adult serial pattern learning in rats,” by S. B. Fountain, J. D. Rowan, B. M. Kelley, A. R. Willey, & E. P. Nolley, 2008, Experimental Brain Research, 187, 651–656. Copyright 2008 by Springer-Verlag. Adapted with permission).
several studies reported here, then this research is also potentially relevant to more generally understanding how animals use cues concurrently either in stimulus compounds that remain “elemental” or that become “configural” (Pearce, 1987; Pearce & Bouton, 2001). The results could help clarify whether cues in complex serial pattern learning tasks are treated the same or differently by the organism compared to when they are encountered in simpler instrumental or classical conditioning settings. Third, the results of these studies will be very useful in informing behavioral neuroscience research that is attempting to understand the cortical and subcortical circuits responsible for sequential behavior. For example, if sequential behavior routinely includes temporal coding for serial position of pattern elements even when other processes are also coding item associations (as indicated in the foregoing studies), then neural models will benefit from that information. This type of information could be particularly useful to those attempting to understand the sequential processes mediated by the hippocampal system, the medial prefrontal cortex, the caudate-frontal cortical loop, and the putamen-supplementary motor area cortical loop, especially as these relate to processes in Alzheimer’s, Parkinson’s, and Huntington’s disease, and to deficits caused by adolescent exposure to drugs. Finally, this line of work can be potentially relevant to understanding how multiple behavioral processes are selected and controlled to produce efficient and accurate behavior through time, a problem central to both serial pattern learning and to the concept of “executive function.”

Acknowledgment
Preparation of this chapter was supported in part by Award Number R15DA023349 from the U.S. National Institute on Drug Abuse to S. B. Fountain. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Drug Abuse or the National Institutes of Health.

References

612 THE ORGANIZATION OF SEQUENTIAL BEHAVIOR


