The hierarchical representation
of three-item sequences

H. L. ROITBLAT
University of Hawaii, Honolulu, Hawaii

R. A. SCOPATZ
Columbia University and New York City Department of Transportation, New York, New York

and

T. G. BEVER
University of Rochester, Rochester, New York

Pigeons were trained to discriminate one sequence of three colors from other sequences made from the same three colors. At each stage of acquisition, a regression analysis estimated the degree to which the birds' performance was controlled by each of seven hierarchically organized units. Three units represented single elements acting individually (Stimulus 1, Stimulus 2, and Stimulus 3), three represented pairs of stimuli (i.e., the combination of Stimuli 1 and 2, the combination of Stimuli 2 and 3, and the combination of Stimuli 1 and 3), and one represented the triplet of all three stimuli. Responding was initially controlled by the third stimulus in the sequence, but eventually came under the control of higher order units representing combinations of stimuli. These results indicate that pigeons are capable of using coherent, hierarchical representations of sequence information. They also argue against a number of list-processing schemes, such as a retrospective trace-strength discrimination scheme and a prospective conditional sequential discrimination scheme.

Pigeons can perform a number of tasks that require them to be sensitive to the order of stimulus events. They have been trained to produce sequences (Straub, 1979; Straub, Seidenberg, Bever, & Terrace, 1979; Straub & Terrace, 1982), to reproduce single responses from a sequence cued by its serial position (Shimp, 1976), and to discriminate strings of sequentially presented stimuli (e.g., Terrace, 1986; Wasserman, Nelson, & Larew, 1980; Weisman & DiFranco, 1981; Weisman, Duder, & von Konigslow, 1985; Weisman, Wasserman, Dodd, & Larew, 1980). The work described in this paper involves the discrimination of stimulus strings from one another.

In the present experiment, a sequence of three events (colored lights) was presented in order on one key of an operant chamber. Following the sequence, the birds were allowed to peck at a common test stimulus. Pecks to the test stimulus following one sequence (labeled ABC) were reinforced; pecks following the other 26 combinations of the same three stimuli were extinguished. The reinforced string is called the positive sequence; the other strings are the negative sequences. Discrimination of the positive from the negative strings is indicated by higher rates of

pecking at the test stimulus following presentations of the positive string than following the negative strings.

In the present experiment, we adopted a procedure similar to that used earlier, for example, by Weisman et al. (1980), except that we used three-event sequences rather than two-event sequences. Weisman et al. studied the ability of pigeons to discriminate two-item sequences consisting of the four possible orderings of two colors. If A stands for one color and B for another, then the positive sequence can be symbolized AB and the negative sequences can be symbolized AA, BA, and BB. Pigeons rapidly acquired this discrimination: They pecked more at the test stimulus following Sequence AB than they did following the other sequences.

A number of mechanisms have been proposed to describe how pigeons perform this kind of sequence discrimination (see, e.g., Weisman et al., 1985; Weisman, Gibson, & Rochford, 1984; Weisman & von Konigslow, 1984). These proposals all recognize that discriminating among different sequences made from the same events introduces some special problems not present in the discrimination of individual events. Most significant of these special problems is that the organism not only must discriminate each of the events in the sequence, but also must discriminate the order of events. It must represent information about the stimuli in the sequence and additional information about their succession.

Animals could use many different schemes (combinations of representations and processes) to discriminate

among sequences. These schemes can be classified roughly along two dimensions—the completeness and the coherence of the representation. Complete representations contain information about each of the elements in the sequence. In contrast, partial representations contain information about only a subset of the elements in the sequence. For example, a recency rule is a scheme that could use only incomplete representations, reflecting only the final element in the sequences. An animal using a recency scheme could discriminate perfectly Sequence AB from Sequence BA, because these two sequences differ in their terminal elements. Furthermore, by hypothesis, the animal has no information in its incomplete representation about the first element in the sequence. As a result, its behavior is expected to be indifferent to the first element in the sequence. An animal using only this recency scheme should respond identically to Sequences AB and BB, because both sequences match its representation of the positive sequence (the terminal element).

A coherent representation of a sequence is one in which information about items and about their succession is stored explicitly and as a unit in memory. This kind of representation is coherent in the sense that it contains a unified, organized integration of the sequence's elements and their serial position. Coherent representations, which organize the information about the sequence in an integrated unit, can be distinguished from noncoherent representations, which represent information in some more disjointed manner. Schemes employing noncoherent representations compute decisions about the serial order of the events by "list processing" the individual stimuli.

Any memory scheme that is capable of supporting a sequence discrimination must contain information about stimuli and their succession, but the information about the succession of the items could be stored explicitly or could be computed from some unordered list of representational elements. A complete description of an animal's scheme for performing sequence discriminations includes a specification of its representations and the means by which it compares its representation of the positive string with the sequence that is presented.

Partial or incomplete schemes are often adequate for many kinds of sequence discriminations. The general method for detecting incomplete representations is to look for indifference sets—sequences of different elements to which the animal responds identically. The animal should be indifferent to the identity of sequence elements not contained in its representation. For example, if the animal represents only the first stimulus in a two-event sequence, it will fail to discriminate between Sequences AA and AB. Both of these sequences match the positive sequence in the first, and represented, position. A comparison of these two strings with the animal's representation of the positive string yields a match because both strings match the positive string in all represented positions and the animal treats them both identically. In general, if the animal responds systematically to some subsequence (e.g., the first element in a two-

element sequence or the first two elements in a three-element sequence), then information about that subsequence must be contained in the animal's representation. Indifference relative to some subsequence suggests that the subsequence is missing from the animal's representation.

Even if animals use complete representations containing information about each of the stimuli in the sequence, there are still many schemes they could use to discriminate sequences. Some of these schemes involve coherent sequence representations; others involve processing unordered lists of individual events. Many of these schemes were reviewed by Weisman and von Konigslow (1984).

At least three classes of complete-sequence representational schemes would allow an animal correctly to discriminate sequences from one another. These schemes differ in (1) the content of the animal's representation, (2) the coherence of its representation, (3) the time at which the animal is assumed to decide whether a presented string is or is not an example of the positive string, and (4) the decision mechanism the animal uses.

First, by analogy with many experiments on human list-learning, one class of schemes involves the maintenance in working memory of information about each stimulus and its sequence position until a decision is made during the test stimulus. These schemes are retrospective (Honig & Wasserman, 1983; Roitblat, 1980, 1982): On the basis of its recall of the stimulus string and its comparison of the remembered string with the positive string, the pigeon decides following the sequence whether to peck at the test stimulus. Although these retrospective schemes involve complete representations of the sequence, they do not all utilize coherent representations of the positive sequence. For example, each stimulus, as it is presented, could activate some representation of itself in memory. This activated trace then could begin immediately to fade at some predictable rate (e.g., Roberts & Grant, 1976; cf. Roitblat, 1987). The current strength of a stimulus's trace in memory would then provide information about when that stimulus appeared. Each sequence would have associated with it a unique set of trace strengths for each of the stimuli in the sequence. The subject would then have to learn which of these sets represented the positive sequence. In this scheme, information about the order of the stimuli is represented implicitly in the strengths of the individual traces. The animal's reference memory would contain explicit information about only the criterion trace strength or the three stimuli following the positive sequence, not about their order. Although trace strength and time since presentation normally would be correlated, they could each be manipulated separately, for example, by varying the duration of each of the elements (cf. Weisman & DiFranco, 1981). The bird would decide that the presented string is the positive string by comparing the strengths of each trace with its criterial trace strength, not by comparing the sequence as a string with a complete representation in reference memory. The animal would be indifferent to how the trace strength was produced.

As long as the individual traces had the criterion strength, however that strength was produced, the animal would judge the string to be positive.

Although this representation contains implicit sequence information, it does not involve a coherent representation of the sequence. Each element is stored independently, and the order of elements is computed from their strength. This scheme is retrospective in that the decision whether to peck is made at the end of the sequence; the content of its representation is stimulus trace strength; and its decision mechanism involves comparison of individual trace strengths with their respective criteria. Although this scheme does not involve coherent representations of the positive scheme, other retrospective schemes that do utilize coherent representations are possible.

A second class of sequence-discrimination schemes operates by analogy with maze learning. Each successive item in the string is treated as a conditional discrimination. A rat learning to run through a maze need not represent the set of choice points as a sequence of choices (e.g., left turn followed by right turn, . . .). Perfect performance requires only that each choice point be associated unambiguously with a certain response, such as turning right or turning left. Only one response is correct at any choice point, independently of how the rat arrived there. The performance of a correct response at each choice point guarantees errorless traversal of the maze (see Bever, Straub, Terrace, & Townsend, 1980). The order in which choices are made is determined by the order in which choice points are encountered, not by any ordered representation of the sequence of choices.

A similar scheme can be applied to sequence discrimination. Each successive stimulus in the sequence sets the occasion for the conditional discrimination of the next stimulus in the sequence. As each item in the sequence is presented, the subject generates an expectancy regarding the next item in the sequence. Either the subsequent stimulus confirms the expectancy and processing continues, or it disconfirms the expectancy and processing ceases (cf. Terrace, 1986; Weisman, Gibson, & Rochford, 1984). In this scheme, information about the order of the stimuli is represented implicitly in the individual conditional discriminations. The bird decides that the presented string is the positive string by performing each conditional discrimination, not by comparing the sequence as a string with its complete representation in reference memory. The animal's reference memory would contain explicit information only about the unordered set of conditional discriminations, not about the order of the three stimuli. Although the set of conditional discriminations contains implicit sequence information, it does not involve a coherent representation of the sequence, because the conditional discriminations can be represented in memory in any order. The order in which the conditional discriminations are performed is determined by the order in which the stimuli are encountered, not by an ordered representation of the sequence.

Table 1
A Production System for Sequence Discrimination

Working Memory	Stimulus	Memory Production	Action Production
Start	Green	Peck	
Start	Red	No peck	
Peck	Red	Peck	
Peck	Green	No peck	
Peck	Test		Pecking
No peck	Any stimulus	No peck	No pecking

Table 1 shows an example of this kind of sequence-discrimination scheme in the form of a production system (see Hayes-Roth, Waterman, & Lenat, 1978). Each combination of a working memory state (Column 1) and a stimulus (Column 2) produces a change in memory state (Column 3) or an action (Column 4). At the start of the sequence, working memory is in the "start" state. Each stimulus that is presented after the start places working memory either in the "peck" state or in the "no peck" state. If working memory is in the peck state when the test stimulus is presented, then the animal pecks; otherwise, it does not peck. Although sequence information is implicit in the rules, the rules themselves could be represented in any particular order. It is the individual combinations of memory states and stimuli that control the animal's responding, not an ordered representation of the sequence as a whole. The order of the stimuli in the sequence is computed from the correct performance of the conditional discriminations.

This scheme is prospective (Honig & Wasserman, 1983; Roitblat, 1980, 1982), in that the subject is assumed to make decisions about upcoming events during the course of the sequence presentation. A number of recent experiments strongly support the importance of prospective processes in animal memory. For example, in delayed matching-to-sample by pigeons (see, e.g., Roitblat, 1980, 1982) and in serial anticipation learning by rats (see, e.g., Hulse, 1978; Roitblat, Pologe, & Scopatz, 1983; Self & Gaffan, 1983), animals appear to use a prospective code in which they anticipate the next item on the basis of one or more recently presented items. The content of this scheme's representation is the current memory state, and its decision mechanism is the set of state transitions. It does not involve coherent representations, because the sequence is computed from implicit information contained in the state transitions.

Finally, a third class of sequence-discrimination schemes involves hierarchical representations. Some units in the representation correspond to individual stimuli, some to combinations of stimuli, and one to the complete sequence considered as a single unit. As the stimuli in the sequence are presented, they activate these hierarchical units. The bird decides whether to peck on the basis of the units that have been activated, for example, as a function of a weighted sum of the activated units. This scheme involves some prospective coding in that the subject is assumed to decide during the sequence whether cer-

tain elements are to be activated. It is partially retrospective in that the subject need not decide whether to peck until the entire sequence has been presented and the relevant units activated (under some conditions, the bird may not wait until the end of the sequence to make its decision; see Terrace, 1986, and Weisman et al., 1984). It involves a coherent sequence representation, because the higher level units in the hierarchy are coherent representations of stimuli in specific serial positions. They represent simultaneously the stimuli in the sequence and explicit information about their order. They are activated only when their relevant stimuli are presented in the proper serial position.

Three-event sequences were used in the present study to compare these alternative representational schemes. The experiment described in the present paper was not directed at ruling out prospective versus retrospective classes of representations as exhaustive classes. The "temporal direction" of the animals' decision-making procedures is not the primary topic of investigation in the present paper. Rather, the primary concern is with the kind of representations animals form during sequence-discrimination performance—whether they are linear and piecemeal or coherent.

In order to be plausible, a model must predict that the animal, given a sufficient amount of training, will successfully discriminate the positive sequence from other sequences. Distinguishing among plausible models, therefore, is possible only if the models make distinctive predictions about properties of the animal's discrimination performance beyond the prediction that the animal eventually will discriminate positive from negative sequences. We have chosen to analyze response patterns. For this reason, we were forced to use three-event sequences because there are simply not enough different ways in which an animal can respond to a two-event sequence. The animal could respond positively to a negative two-event sequence either because it failed to discriminate the first stimulus or because it failed to discriminate the second stimulus. No other error patterns (responding to a negative sequence as if it were the positive sequence) are possible because all negative sequences, even those with both stimuli different from the positive string, differ from the positive string in either the first or the second position. Sequences consisting of at least three events are necessary to detect the presence of coherent sequence representations, because three-event sequences contain enough potential complexity to compare the alternative schemes described above. With a three-event sequence, there is a possibility that the subject's performance can be controlled by combinations of elements that are greater than one stimulus (consistent with incomplete representations) but less than the entire string (consistent with any plausible account).

Both the trace-strength and the production-system schemes described above suggest that each stimulus in the sequence contributes independently to the animal's discriminative decisions. The trace-strength scheme asserts that peck decisions are made on the basis of the relative trace strengths of the individual stimuli. The production-system scheme argues that discrimination depends on the traversal of a chain of conditional discriminations of individual stimuli. In contrast, the hierarchical scheme argues that peck decisions are made on the basis of hierarchical units representing one or more stimuli in the sequence. We investigated these alternatives using a three-event sequence.

METHOD

Subjects
Four White Carneaux pigeons served as subjects. They were maintained at 75%–80% of their free-feeding weights under a natural light/dark cycle and were housed in individual cages.

Apparatus
Experimental sessions were conducted in one of two single-key, ventilated, sound-attenuating chambers (25×25×30 cm high). Each subject was trained in the same chamber throughout the experiment. The front panel of each chamber (25×30 cm) was equipped with a single pecking key, 2.5 cm in diameter, centered 25 cm above the wire-mesh floor. The key could be transilluminated by an in-line projector (IEE Model 51022-1820) equipped with Kodak Wratten filters (#45 blue, three layers of #102 green, and #92 red). The projector could also project three horizontal black lines along the width of the key (against a white background) and a uniform white field. The chamber contained an illuminated mixed-grain hopper located behind a 5.5×5.5 cm opening in the center of the front panel, 4.5 cm above the wire-mesh floor. A houselight provided constant illumination.

Pretraining Procedure
Pretraining consisted of three phases and was followed by sequence-discrimination training. During the first phase of pretraining, the intertrial interval (ITI) was set to 60 sec. In this first phase, pigeons were autoshaped to peck at the white test stimulus for food reward. A single peck caused the food hopper to be presented for 2.5 sec. If the bird did not peck, the food was presented after 5 sec. When the birds began pecking reliably (i.e., on 140 of 144 daily trials), the second phase of pretraining was begun. During this phase, the warning signal (three horizontal lines) was presented prior to the test stimulus. A single peck to the warning signal caused the trial to advance to the test stimulus, and a single peck to the test then caused the food to be presented. The trial similarly advanced if 5 sec had elapsed without a peck. When the birds pecked reliably (according to the same criterion) at these stimuli, the third phase of pretraining was begun. A 1-sec delay was inserted between the warning signal and the test stimulus. The number of trials per daily session was also reduced from 144 to 72, and the ITI interval was increased to 120 sec. When the criterion was again reached (70 of 72 trials completed with pecks), the experiment proper was begun, and the three-event sequences were inserted into the ITI between the warning signal and the test stimulus.

Sequence Discrimination Training Procedure
The within-trials procedure used in sequence-discrimination training is outlined in Figure 1. Each trial began with a 3-sec presentation of a warning signal consisting of three horizontal black lines on a white surround. A single peck to this stimulus resulted in the presentation of the rest of the trial. If the bird did not peck at this stimulus, then the ITI was presented, followed by a repeat of the same trial. A "ready interval" followed the termination of the warning signal and was itself followed by the first stimulus in the se-
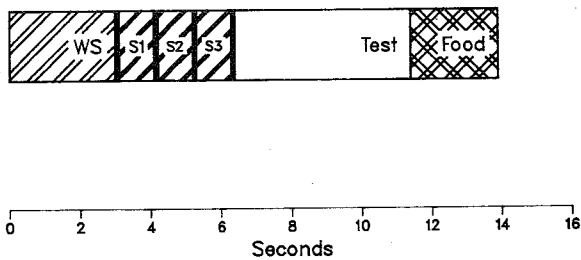
Figure 1. An outline of the general within-trials procedure used. WS = warning signal, S1 = Stimulus 1, S2 = Stimulus 2, S3 = Stimulus 3. Black bars represent interstimulus intervals. Only trials involving positive sequences ended with food. All stimuli were presented on the same key. Trials were separated from one another by 15-sec intertrial intervals.

quence, the first interstimulus interval (ISI), the second stimulus, second ISI, third stimulus, and retention interval (RTI). A white test stimulus appeared after the RTI and remained illuminated for 5 sec. Following the positive sequence, a single peck at any time during this stimulus was sufficient to produce 2.5 sec access to the food hopper. On negative trials and on positive trials without a peck, the termination of the test stimulus was followed by a 2.5-sec delay in place of the food presentation. The ITI was always 15 sec long.

For each bird, the positive sequence consisted of three colors, designated ABC. The identities of Colors A, B, and C were different for each bird. Each session consisted of 144 trials, 40 of which presented the positive color sequence (ABC); the remaining 104 trials presented negative sequences. All possible combinations of A, B, and C were used as negative sequences (e.g., ACB, AAC, CBC, etc). Each of the 26 negative sequences appeared four times within a day. The sequences were presented in a different random order each day.

RESULTS

The 5-sec test stimulus following each sequence allowed the subjects to peck a number of times. We utilized the potential information in this number by calculating discrimination ratios according to Equation 1. An overall discrimination ratio was computed for each session using the average rate of pecking to the test stimulus following positive sequences (ABC) as the rate (positive) and the average rate following the negative sequences as the rate (negative). The ratio was then computed according to Equation 1:

$$DR = \frac{\text{rate (positive)}}{\text{rate (positive)} + \text{rate (negative)}}. \quad (1)$$

The discrimination ratio (DR) has the value of 0.5 when responding is indiscriminate to the positive and negative sequences, and rises to 1.0 when discrimination is perfect (i.e., when no pecks are emitted to the negative sequences). In addition to the overall discrimination ratio for the entire session, which was based on the average rate of pecking to all negative sequences, we also computed discrimination ratios for each individual sequence. In computing these individual ratios, the rate (negative) was the average peck rate following each specific sequence.

Figure 2 shows the discrimination ratios for each bird and each sequence during the 5th, 10th, and 20th six-session blocks of acquisition. The discrimination ratio for Sequence ABC is included for the sake of completeness, but is set to .5 by definition.[1] In Figure 2, a high discrimination ratio for a particular negative sequence indicates a lower rate of pecking to that particular negative sequence than to the positive sequence. In contrast, a low discrimination ratio indicates a high rate of pecking to that particular negative sequence (i.e., a peck rate similar to that observed following the positive sequence).

Not all sequences were equally discriminated from the positive sequence. For example, especially at the beginning of training, sequences ending in Stimulus C were generally discriminated most poorly from the positive sequence. The birds responded more when the sequence ended in C than when it ended in any other stimulus. One possible way of interpreting this generalization is to suppose that the birds' representation included the information that the positive sequence had Stimulus C in the third position. As a result, they pecked frequently when the sequence ended with C, but not when the sequence ended with some other stimulus. If the bird's representation of the positive sequence were only that it ended with C, then it would generalize to all sequences that ended with C but would discriminate all other sequences that ended with one of the other stimuli. It would then show high discrimination ratios for all sequences that ended with A or B and low discrimination ratios for all sequences that ended with C. This pattern was characteristic of all of the birds during the early stages of acquisition. Apparently, they began learning to discriminate among the sequences by first learning that the positive sequence ended with C and then learning about other portions of the positive sequence. Further training eliminated or reduced responding to the negative sequences in this group. Even at the end of the experiment, however, Bird 939 was poor at discriminating AAC, BAC, BBC, and BCC from the positive, ABC, sequence, although it did discriminate other sequences from the positive sequence, and its overall discrimination ratio was above .5.

We used the analytic technique described in the Appendix to examine the pattern by which various stimuli and stimulus conditions came to control responding during acquisition. This analytic method uses regression procedures to estimate the degree of control by units in the sequence. We identified seven possible units in the sequence. Each of these units corresponds to one substring in the sequence.

Three units represent individual stimuli. S_1 represents the first stimulus in the sequence. S_1 is assigned a positive value when the first stimulus in the presented sequence matches the first stimulus in the positive sequence. It is assigned a negative value when the first stimulus in the presented sequence is different from the first stimulus in the positive sequence. S_2 represents the second and S_3 represents the third stimulus in the sequence. Like S_1, S_2 and S_3 are assigned positive values when their respective stimuli match the positive sequence and negative values
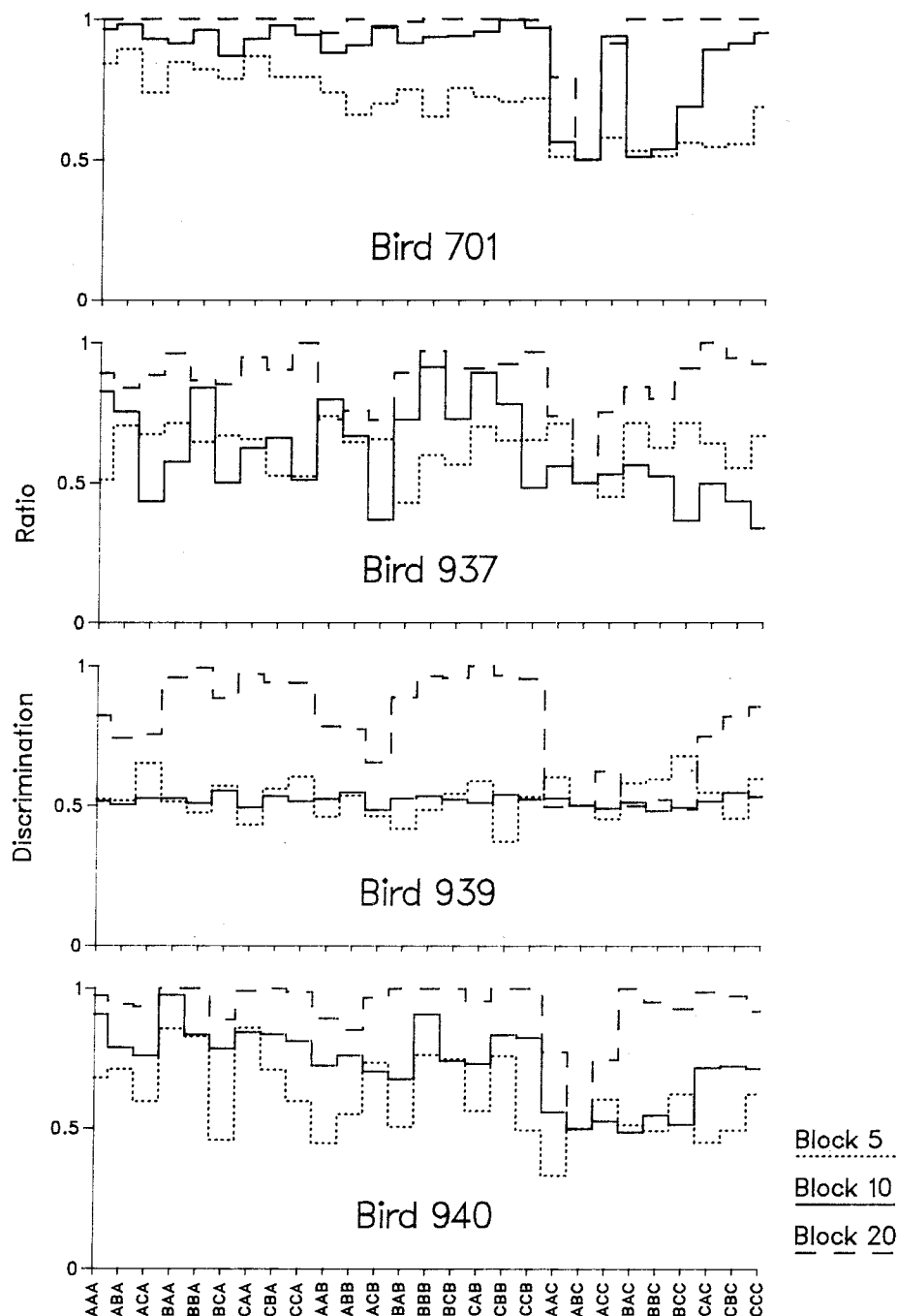
Figure 2. The discrimination ratios of individual sequences at three stages of training. A discrimination ratio of 1.0 indicates perfect discrimination between the indicated sequence and the positive sequence. A ratio of .5 indicates no discrimination. By definition, the discrimination ratio for the positive sequence, ABC, is .5. Each block consisted of six sessions.

when they do not match. For example, Sequence ACB matches the positive string in the first position but not in the second or third position. Therefore, when this sequence is presented, it causes S_1 to have a positive value and S_2 and S_3 to have negative values.

Three additional units were defined to correspond to the three pairs of stimuli in the sequence. S_{12} corresponds to the combination of the first and second stimuli, S_{23} corresponds to the pair of the second and third stimuli, and S_{13} corresponds to the pair of the first and third stimuli. Like the single-stimulus units, these units take on positive values when both stimuli represented by the combination unit match the corresponding stimuli in the presented string and negative values when other stimuli

appear in the presented string. Sequence ACB, described above, would result in negative values for all three of the two-stimulus units because none of the pairs of stimuli (AC-, -CB, A-B) in the presented string corresponds to the positive string. On the other hand, Sequence ACC would cause S_{13} to take a positive value because this combination of the first and third stimuli (A-C) does match the positive sequence.

The final unit, S_{123} , corresponds to the string considered as a whole. This unit takes a positive value only when the positive string is presented. Otherwise, it is set to 0.

The analysis predicts the rate of pecking to the test stimulus following each sequence at each stage of training as a function of these seven units in response to the presented string. The rate of pecking following each sequence j corresponds to a linear weighted sum of the contributions from each of these S_{ij} (substring i in sequence j) units. The weights were estimated by linear regression. A high weight corresponds to strong control by the unit. To the extent that a unit receives a high weight, the bird will increase its rate of pecking when the unit is positive (i.e., when the corresponding substring matches the positive string) and decrease its rate of pecking when the unit is negative (i.e., when the corresponding substring does not match the positive string). To the extent that a unit receives a low weight, it will have little influence on the rate of pecking, and the peck rate will not change as a function of the stimuli appearing in its corresponding substring. Presumably these weights reflect (1) the extent to which the bird included the unit as part of its representation of the positive sequence and (2) the importance (attention) the bird attached to the unit as a discriminative element. Acquisition of the discrimination consists of changes in the weights assigned to one or more of these units.

In many ways, this analysis is analogous to a three-way factorial analysis of variance. The single-stimulus units are analogous to the main effects, the two-stimulus units are analogous to the two-way interactions, and the three-stimulus unit is analogous to the three-way interaction. As in analysis of variance, the higher level units reflect information that is not reducible to additive combinations of the lower level units. For example, if the combination unit S_{13} controls responding, then the presence of Stimulus A in the first position is predicted to cause an increase in responding, as does the presence of Stimulus C in the third position. When both of these stimuli are present, however, in their respective positions, the increase in responding is predicted to be greater (all other things being equal) than the increase due to each element separately (S_1 + S_3). That is, the combination of A-C produces a higher level of responding than the sum of A-- and --C.

Figures 3-6 show for each subject the extent to which each element in the sequence controlled responding. The top panels show the overall discrimination ratios over the course of the experiment and the degree to which the analysis model accounted for the birds' performance (or the
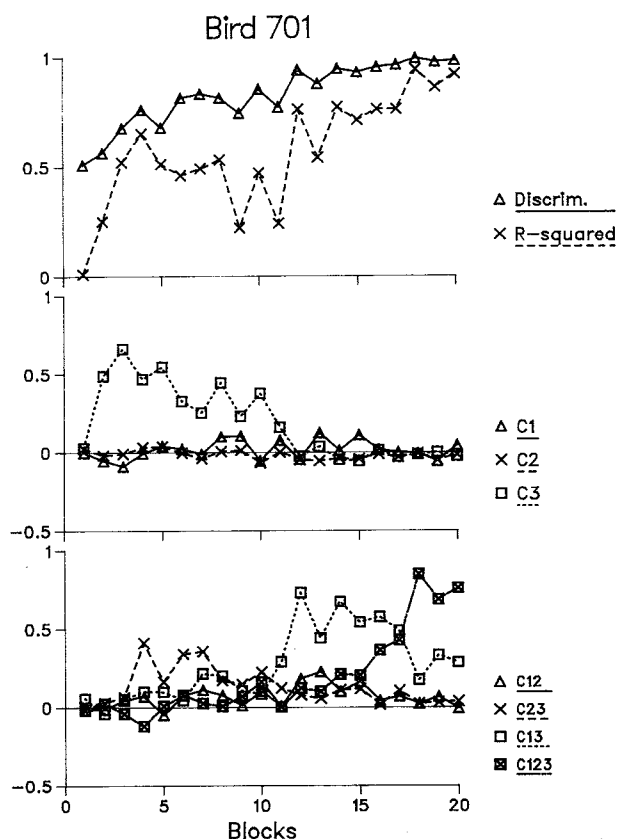

Figure 3. The top panel shows the overall discrimination ratio and the R^2 measure of stimulus control over the course of acquisition for Bird 701. The middle panel shows the control over responding exerted by the first, second, and third stimuli (C1, C2, and C3, respectively) appearing in the sequence. The bottom panel shows the control over responding exerted by combination units representing the first and second (C12), second and third (C23), first and third (C13), and the entire sequence taken as a unit (C123). Magnitude of control is displayed in the form of standardized regression weights. Data are displayed in six-session blocks.

degree to which the pigeons' responding was systematic relative to the model; other factors not explicitly contained in the analysis model, such as hunger, or distractions due to ambient laboratory noise, may have contributed to responding). C1, C2, and C3 in the middle panels are the standardized regression weights. They show the extent to which S_1, S_2, and S_3, the units corresponding to the individual stimuli, respectively controlled responding. High values indicate strong control, values near 0 indicate weak control, and negative values indicate "perverse" control—responding that is systematically related to the stimuli in the string, but in a way that is opposite to that dictated by the reinforcement contingency. In the bottom panels, C12, C23, and C13 show, respectively, the extent to which S_{12}, S_{23}, and S_{13}, the representational units corresponding to the stimulus pairs, controlled responding. Finally, C123 shows the extent to which S_{123}, the string considered as a single unit, controlled responding.
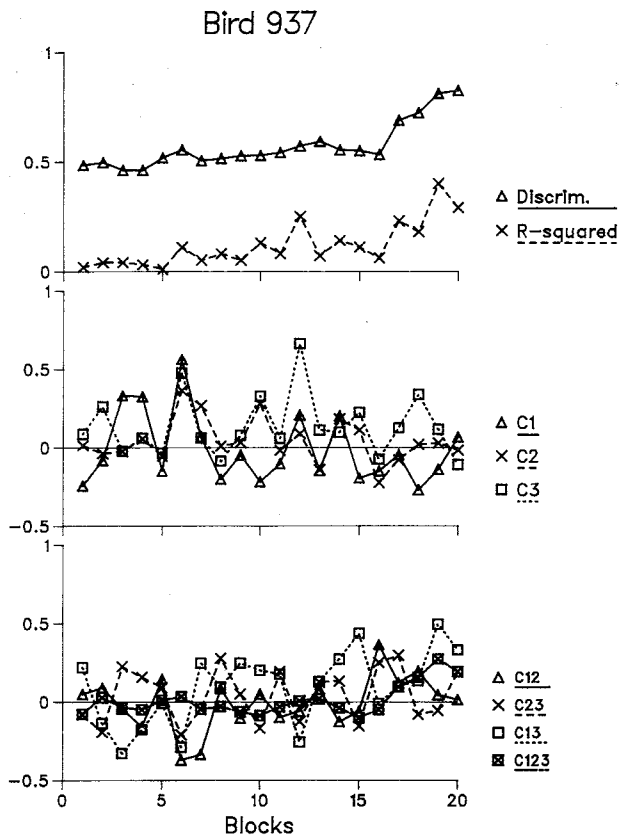
Bird 937

Figure 4. The top panel shows the overall discrimination ratio and the R^2 measure of stimulus control over the course of acquisition for Bird 937. The middle panel shows the control over responding exerted by the first, second, and third stimuli (C1, C2, and C3, respectively) appearing in the sequence. The bottom panel shows the control over responding exerted by combination units representing the first and second (C12), second and third (C23), first and third (C13), and the entire sequence taken as a unit (C123). Magnitude of control is displayed in the form of standardized regression weights. Data are displayed in six-session blocks.

Consistent with the data in Figure 2, the third stimulus in the sequence was generally the first to begin controlling performance. Its contribution, however, soon was replaced by control by higher order combinations of stimuli. For example, the peck rate of Bird 701 (Figure 3) was significantly affected by the identity of the third stimulus (S_3) beginning in the second block of training [$t(160)$ = 4.216, p < .01]. During Block 4, control by S_3 was joined by a significant [$t(160)$ = 5.884, p < .0001] contribution from the unit representing the combination of Stimuli 2 and 3 (S_{23}). During Block 7, pecking was controlled by the third stimulus, S_3 [$t(160)$ = 2.633, p < .01] and by two combination units, S_{23}, representing the combination of Stimuli 2 and 3 [$t(160)$ = 3.414, p < .001], and S_{13}, representing the combination of Stimuli 1 and 3 [$t(160)$ = 2.163, p < .05].

During Block 9, the third stimulus stopped contributing as an individual stimulus. In fact, during this block, none of the stimulus units contributed significantly to

responding, although the regression was still significant [$F(7,154)$ = 6.246, p < .0001, multiple R^2 = 0.2211]. Control by the third stimulus returned briefly in Block 10 [$t(160)$ = 3.833, p < .0005] but was not significant in any later block.

After Block 10, the major combination unit to control responding was that of Stimuli 1 and 3 (S_{13}). This unit contributed significantly in Bock 11 [$t(160)$ = 2.377, p < .02] and in every block after that. Finally, in Block 14, the unit consisting of all three stimuli (S_{123}) began to contribute significantly [$t(160)$ = 3.085, p < .003].

The acquisition pattern for Bird 940 was similar to that for Bird 701. The main difference was in the earlier contribution made by higher order combination units to the performance of Bird 940. The other two birds learned the discrimination more slowly and less systematically.

One can compare the patterns of control illustrated in Figures 3-6 with the patterns of discrimination ratios displayed in Figure 2. Consider, for example, the patterns
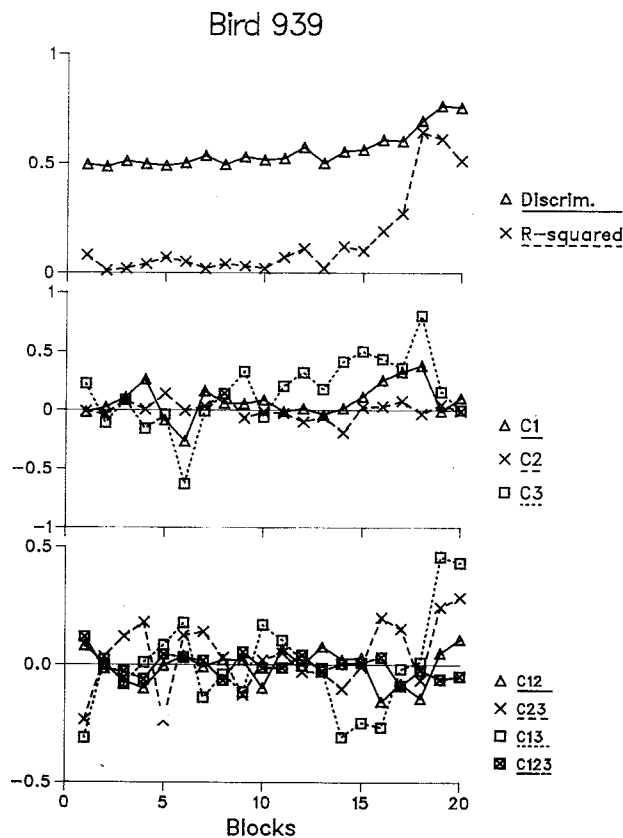
Bird 939

Figure 5. The top panel shows the overall discrimination ratio and the R^2 measure of stimulus control over the course of acquisition for Bird 939. The middle panel shows the control over responding exerted by the first, second, and third stimuli (C1, C2, and C3, respectively) appearing in the sequence. The bottom panel shows the control over responding exerted by combination units representing the first and second (C12), second and third (C23), first and third (C13), and the entire sequence taken as a unit (C123). Magnitude of control is displayed in the form of standardized regression weights. Data are displayed in six-session blocks.
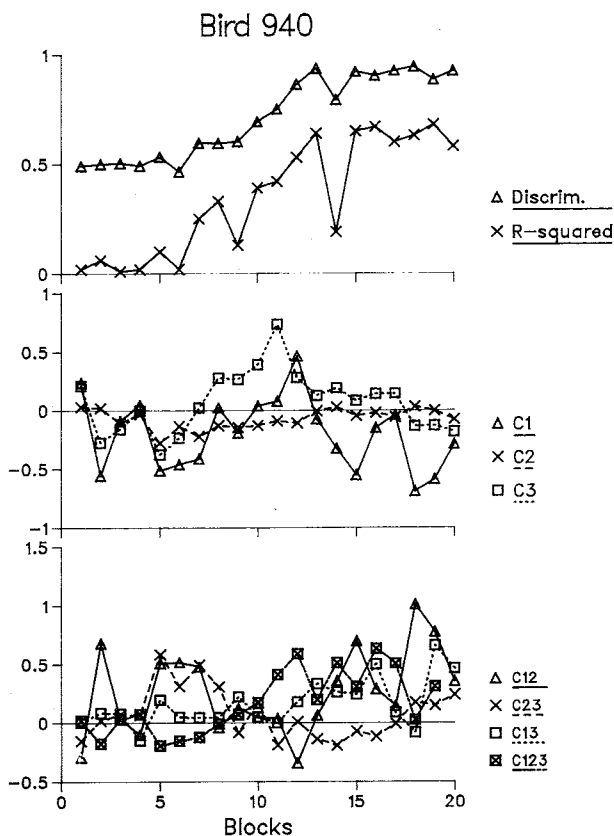
Bird 940

Figure 6. The top panel shows the overall discrimination ratio and the R^2 measure of stimulus control over the course of acquisition for Bird 940. The middle panel shows the control over responding exerted by the first, second, and third stimuli (C1, C2, and C3, respectively) appearing in the sequence. The bottom panel shows the control over responding exerted by combination units representing the first and second (C12), second and third (C23), first and third (C13), and the entire sequence taken as a unit (C123). Magnitude of control is displayed in the form of standardized regression weights. Data are displayed in six session blocks.

shown for Block 20. For Bird 701, significant control was exerted by the combination units S_{13} [$t(160) = 7.42$, $p < .0001$] and S_{123} [$t(160) = 19.28, p < .0001$]. Consistent with control by S_{13}, this bird differentiated sequences that started with A and ended with C from all other sequences; it responded almost exclusively to these sequences. Consistent with control by S_{123}, this bird further differentiated among Sequences AAC, ABC, and ACC; it pecked reliably more to ABC than to the other two A–C sequences.

Only S_{123} exerted significant control over Bird 937's pecking in Block 20 [$t(160) = 2.46, p < .02$]. This bird pecked reliably more to the positive, ABC, sequence than to any other, but did not differentiate systematically among the alternative negative sequences.

For Bird 939 in Block 20, S_{13} [$t(160) = 3.12$, $p < .005$] and S_{23} [$t(160) = 2.35, p < .02$] exerted significant control over responding. Control by these two combination units reflected relatively poor discrimination of sequences that began with A and ended with C or sequences that had B as the second stimulus and ended with C. This

bird pecked reliably if the sequence began with A and ended with C, no matter what was in between. It also pecked reliably if the sequence ended with the subsequence BC. These two independent biases contributed additively to an increase in peck rate if both A–C and –BC were present. This pattern can be seen most clearly in the stepwise increase in discrimination ratio from ABC to CCC in Figure 2. The apparent differentiations in this bird's peck rate among the remaining sequences in Figure 2 were not sufficiently systematic to be captured by the analysis.

For Bird 940 in Block 20, S_{13} [$t(160) = 4.407$, $p < .0001$] and S_{123} [$t(160) = 3.70, p < .0005$] exerted significant control. This is essentially the same pattern of control seen for Bird 701. This bird discriminated better between sequences that started with A and ended with C than between other sequences, and it further differentiated between the positive sequence and the two negative sequences in this subset (ABC vs. AAC and ACC).

DISCUSSION

This experiment found that birds' sequence-discrimination performance was controlled by a changing pattern of representational units. Early in acquisition, the final stimulus, acting by itself, controlled the rate at which the birds pecked following each sequence. The birds pecked more when the final stimulus in the presented string matched the final stimulus in the positive string than when the string ended with a different stimulus. After some amount of training, the pattern of responding for each bird changed. Peck rate was no longer controlled by an individual stimulus in the sequence, but instead was controlled by a higher order unit representing an interactive combination of stimuli. This higher order combination unit was not reducible to any linear or additive combination of contributions from the individual stimuli but, instead, was a true, coherent representation of the stimuli in sequence. This pattern of results is similar to that observed by Weisman et al. (1985). They trained pigeons on a three-event, two-stimulus sequence discrimination (i.e., the positive sequence was BAB). They then analyzed the rank order of discriminability between each sequence and the positive sequence following training. They found that an additive model was inadequate to account for their pigeons' rank orderings and argued instead for a cumulated multiplying model. Although neither their model nor their positive sequence is directly comparable with ours, their model and results are similar to ours in indicating that pecking is determined by an interaction among the elements in the sequence that cannot be reduced to an additive combination of contributions from individual elements.

The control by higher order combination units observed in the present experiment indicates that birds treat the sequences as combinations of hierarchically ordered units representing one or more stimuli, not as unordered sets of conditional discriminations. The effect of a higher order unit is different from the additive effects of the individual component stimuli.

In order for the animal to use hierarchical units of the sort proposed here, it must be able to keep information about the identity and serial position of the unit's stimuli together in memory. In other words, it must maintain a true coherent sequential representation of the stimuli represented by the unit. For example, if the bird's performance is controlled by the higher order unit S_{13}, it must keep in memory information about the identity of the stimulus in Position 1, until the stimulus in Position 3 is presented and identified. Furthermore, it must "know" that Stimulus 3 not only followed Stimulus 1, but also that it followed in Position 3, not in Position 2. Information about earlier stimuli in the sequence must be maintained in working memory until the later stimuli in the unit appear and can be identified. When Stimulus 1 appears, the bird could assign a value to the unit representing Stimulus 1 (S_1) and maintain a sensory representation of the first stimulus. Then, when the third stimulus appears, the bird could translate the sensory information remaining from Stimulus 1 and combine it with that derived from Stimulus 3 to produce a more abstract representation of the combination of Stimuli 1 and 3.

Although the higher level units were coherent representations of parts of the sequence, they were often incomplete in the sense that they represented only a subset of the stimuli in the sequence. For example, pecking for Bird 701 in block 10 was controlled by S_{13}, which is a coherent representation of the combination of Stimuli 1 and 3, but contains no information about Stimulus 2. With extended training, however, the completeness of the representation increased, either by including overlapping combination units or by increasing the size of the unit. Given sufficient training, there is every reason to expect that the birds' performance would eventually be controlled largely or exclusively by the highest order combination unit representing the entire string as a single coherent unit.

The patterns of discriminative performance observed in the present experiment argue strongly against both the trace-strength retrospective and the production-system prospective schemes described in the introduction. The trace-strength hypothesis proposed that each of the stimuli is retained in working memory in a basically sensory form until the test appears. On the basis of such an account, there is no reason to expect that these stimulus memories should show any kind of structure, except that structure resulting from the loss of strength with the passage of time. Therefore, according to the trace-strength scheme, pecking should be controlled either by an additive combination of individual stimuli or by a multiplicative combination of all three stimuli as a single unit (cf. Weisman et al., 1985; Weisman & von Konigslow, 1984). Neither of these patterns was observed. Responding to the test stimulus frequently was controlled by combination units that represented only subsets of the stimuli in the sequence. Most importantly, responding was often controlled by subsets representing nonadjacent stimuli.

If we had tested the birds only after acquisition was complete, and they had all achieved complete coherent representations, we would not have been able to discriminate between the trace-strength and the hierarchical models. Both models are compatible with a complete coherent representation at the end of training. They make different predictions, however, about the representations used at earlier stages of acquisition. The trace-strength model predicts that the representation used at all stages of training will be either individual stimulus units or a complete coherent representation. The presence of incomplete hierarchical units is incompatible with the predictions of the trace-strength model but consistent with the hierarchical model.

Although the data from the present experiment are incompatible with the trace-strength model described in the introduction, they do not rule out all retrospective schemes. The feature of the trace-strength model that generated predictions incompatible with the present data was not the time at which the comparison was made (during vs. following the sequence); rather, it was the independent way in which the elements of the sequence were represented. It is possible, for example, to construct a hierarchical model that maintains information about the identity of each stimulus in working memory and then uses this information at the end of the sequence to activate units in the hierarchy. The important point is not whether the sequence-processing system is retrospective or prospective, but whether it involves list processing of individual stimuli or coherent sequence processing. Many retrospective memory schemes can be ruled out by the present experiments, but only because they make inappropriate assumptions about the nature of the representation that is retrospectively formed.

The production-system scheme is also inconsistent with the data reported in the present experiment. The production-system model predicts that pecking is controlled either by units representing individual stimuli or by units consisting of adjacent pairs of elements. Control by higher order units consisting of nonadjacent stimuli or of the entire string as a single unit is inconsistent with the kind of production-system model outlined in Table 1. Furthermore, it is doubtful that any purely prospective scheme involving stimulus-by-stimulus comparison of the ongoing sequence with the representation of the positive string could be made compatible with representations of nonadjacent stimuli. Stimulus-by-stimulus comparison is not possible because the unit representing the combination of Stimuli 1 and 3 cannot be assigned a value when Stimulus 1 is presented. Instead, it must wait until Stimulus 3 appears before it can be compared with memory. Its value is a noncumulative product of the identities of Stimuli 1 and 3. The identity of the intervening Stimulus 2 plays no role (cf. Weisman et al., 1985).

Two experiments (Terrace, 1986; Weisman et al., 1984) found that under certain conditions animals can report before the sequence is complete whether it is a positive or a negative string. Weisman and his associates provided pigeons with an "advance" key that allowed the pigeon to terminate an ongoing sequence and bring on the

next trial. They found that pigeons frequently responded to terminate the presentation of negative sequences before the sequences were completed. Terrace trained his pigeons to respond to one key if the positive sequence had appeared and to respond to another key if a negative sequence had appeared. Toward the end of acquisition, Terrace observed that on about 58% of the trials his birds pecked at one of the two decision keys before the sequence was finished and before the keys were illuminated. Hence, both experiments showed that animals do not always need to wait for the end of a sequence before making a decision about the identity of the sequence. Terrace interpreted these results as evidence for a purely prospective scheme and as evidence against the use of higher order sequence representations. Procedural and measurement differences make comparison of these experiments with the present one difficult, but three comments can be made.

First, the representation an animal forms in a task is likely to depend on the way the task is presented to the animal. Representations useful in one domain may be less useful in another domain (Roitblat, 1982). As a result, the animals studied by Weisman et al. (1984) and Terrace (1986) may have formed representations that were appropriate to the tasks on which they were trained and different from the representations apparently formed by the animals in the present experiment. For example, the birds trained by Weisman et al. may have formed representations that were specialized for recognizing a negative sequence as quickly as possible because they were rewarded for rejecting an ongoing negative sequence by speeding the presentation of the next, possibly positive, sequence. Although it seems less plausible, Terrace's two-key (positive/negative) response procedure may also have encouraged his pigeons to form specialized representations. Evidence is available, however, that suggests that Terrace's animals employed representations and decision rules that were similar to those apparent in our own birds' performance.

Second, the use of a hierarchical representation does not preclude making decisions on the basis of incomplete information. For example, the use of two response keys to which only a single peck was allowed may have encouraged Terrace's (1986) birds to apply a sharp criterion to their decisions about which key to peck. In contrast, our birds could respond any number of times to the single test stimulus, which allowed them to employ a graded decision mechanism cumulating information from a variety of units. In support of the notion that decision pecks during the sequence are generated on the basis of incomplete information, Terrace (1986, Figure 2) observed that approximately 20%–30% of these pecks were followed by incompatible choices at the end of the trial. That is, the birds frequently changed their decisions after more complete information had become available with the appearance of the rest of the sequence. Because a response to the advance key terminated the trial, similar data are not available from the study by Weisman et al. (1984).

Third, consistent with our pattern of acquisition, Terrace (1986) found that discrimination was poorest when the sequence ended with C. Although his animals were capable of deciding prior to the terminal stimulus that some sequences were not the positive sequence, they still had the most difficulty with sequences that ended with the same stimulus as the positive sequence, regardless of the stimuli that preceded it (Terrace, 1986, Table 1). Early in acquisition, sequences ending in C were the most poorly discriminated set of sequences. At the end of acquisition, they were the second most poorly discriminated set of sequences (improving from 47.1% correct to 60%). This curious finding suggests that the appearance of negative stimuli early in the sequence (i.e., stimuli that did not match the corresponding stimuli in the positive sequence) did not terminate processing of the sequence and that Terrace's pigeons were not performing the sequence discrimination in a wholly prospective manner, as he suggested. If the birds had decided that the sequence was negative as soon as they encountered a negative stimulus, then they should have stopped processing subsequent stimuli because, by the time they made this decision, they already would have complete information about the sequence and the outcome of the trial (a sequence beginning with Stimulus B can never be the positive sequence). They should have been indifferent to the stimuli that appeared in the sequence after the first negative element. They should readily have discriminated BCC, CBC, and ACC from the positive ABC. Instead, these sequences were among the most difficult.

According to the prospective model described by Terrace (1986), decisions to peck at the positive or the negative test are generated during the sequence presentation. This model predicts that 2/3 of the presented sequences should be rejected on Stimulus 1 in the sequence, 2/9 should be rejected on Stimulus 2 in the sequence, and the remaining 2/27 should be rejected on Stimulus 3 in the sequence. Therefore, by the time Stimulus C appears in the third position, a prospective model predicts that 24/27 of the sequences should already have been rejected. Thus, a prospective model predicts that accuracy with Stimulus C in the third position should be very high, because most of the sequences ending with C should already have been rejected by a decision process earlier in the sequence. In contrast, both Terrace and we observed very poor discriminative performance on sequences ending with C. Even if the decision process were less than perfectly efficient, sequences ending with Stimulus C should be relatively easy to discriminate, rather than be the most difficult or the second most difficult. Therefore, Terrace's pigeons, as were ours, were engaging in decision processes more complicated than a simple prospective model would suggest. They were doing more than comparing individual stimuli in working memory with a representation in reference memory.

Little can be said about the representation used by Terrace's (1986) pigeons. Choice accuracies are given for classes of sequences characterized by individual stimuli

(e.g., sequences beginning with C, sequences with B in the middle position, etc.). As a result, we cannot know about the effects of combinations of stimuli and possible interactions among them. We cannot know, therefore, whether or not Terrace's animals were employing hierarchical representations. His data suggest, nonetheless, that his animals were employing higher order representations. The finding that sequences beginning with A and sequences ending with C were the two most difficult suggests that his animals' representations contained the A–C S_{13} combination unit.

The evidence from our experiment clearly favors a hierarchical scheme for sequence discrimination. We can next turn to consideration of some of the factors that might control acquisition of the representation. For example, the data are incompatible with a chaining or spread-of-reinforcement account (cf. Terrace, 1986). By this hypothesis, for example, the third stimulus in the sequence initially controlled performance because it was closest to the reinforcer. It then became a conditioned reinforcer controlling discrimination of the second stimulus, and so forth. The strongest form of this hypothesis would predict that acquisition would proceed in the following order: S_3 would be the first to gain control over responding, followed by S_2, and then by S_1, as each stimulus served as a conditioned reinforcer for the conditional discrimination of the preceding stimulus in the sequence. In this form, the hypothesis clearly predicts that discriminative performance is controlled by individual stimuli, not by higher order units. In two ways, then, it fails to account for the observed performance. First, acquisition did not follow the predicted order, and, second, performance was controlled by higher order units.

A weaker form of this spreading-reinforcement hypothesis might argue that conditioned reinforcement causes the preceding element to be integrated into a combination unit with the conditioned reinforcer. We would then expect acquisition to proceed in the order S_3, followed by S_{23}, followed by $S_{12}+S_{23}$. By this hypothesis, discriminative performance is controlled by higher order units made from adjacent pairs of stimuli. This hypothesis is also inconsistent with the observed pattern of responding. Performance frequently was controlled by combination units involving nonadjacent stimuli and, toward the end of the experiment, by the unit integrating all three stimuli in Sequence S_{123}. These data show that the birds were processing the stimulus strings in ways that were more complex than would be consistent with either a stimulus-by-stimulus account (either the trace-strength or the production-system schemes discussed above) or a conditioned-reinforcement account, because these accounts prohibit significant control by nonadjacent combinations of stimuli.

The birds differed in the extent to which higher order units controlled their responding. This variation is consistent with the hierarchical representational scheme we are advocating. This scheme permits both lower level and higher level units to control responding. In contrast, the other schemes we have considered assert only a single level of representation and generally prohibit any control over responding by higher level units. They assert, instead, that response rate is controlled solely by additive contributions from the discrimination of individual stimuli.

We conclude from this experiment that birds are capable of hierarchically organized representations of sequences. The higher level units in the representation are separate from the lower level units that correspond to smaller subsequences. Thus, they involve emergent properties that cannot be predicted solely from the reductive analysis of their components.

REFERENCES

BEVER, T. G., STRAUB, R. O., TERRACE, H. S., & TOWNSEND, D. J. (1980). The comparative study of serially integrated behavior in humans and animals. In P. Juscyk & R. Klein (Eds.), *The nature of thought: Essays in honor of D. O. Hebb.* Hillsdale, NJ: Erlbaum.

COHEN, J. (1968). Multiple regression as a general data-analytic system. *Psychological Bulletin, 70,* 426-443.

DWYER, J. H. (1983). *Statistical models for the social and behavioral sciences.* New York: Oxford University Press.

HAYES-ROTH, F., WATERMAN, D. A., & LENAT, D. B. (1978). Principles of pattern-directed inference systems. In D. A. Waterman & F. Hayes-Roth (Eds.), *Pattern-directed inference systems* (pp. 577-601). New York: Academic Press.

HONIG, W. K., & WASSERMAN, E. A. (1983). Performance of pigeons on delayed simple and delayed conditional discriminations under equivalent training procedures. *Learning & Motivation, 12,* 149-170.

HULSE, S. H. (1978). Cognitive structure and serial pattern learning by animals. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 311-340). Hillsdale, NJ: Erlbaum.

ROBERTS, W. A., & GRANT, D. S. (1976).Studies of short-term memory in the pigeon using the delayed matching to sample procedure. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), *Processes of animal memory.* Hillsdale, NJ: Erlbaum.

ROITBLAT, H. L. (1980). Codes and coding processes in pigeon short-term memory. *Animal Learning & Behavior, 8,* 341-351.

ROITBLAT, H. L. (1982). The meaning of representation in animal memory. *Behavioral & Brain Sciences, 5,* 353-406.

ROITBLAT, H. L. (1987). *Introduction to comparative cognition.* New York: Freeman.

ROITBLAT, H. L., POLOGE, B., & SCOPATZ, R. A. (1983). The representation of items in serial position. *Animal Learning & Behavior, 11,* 489-498.

SELF, R., & GAFFAN, E. A. (1983). An analysis of serial pattern learning by rats. *Animal Learning & Behavior, 11,* 10-18.

SHIMP, C. P. (1976). Short-term memory in the pigeon: Relative recency. *Journal of the Experimental Analysis of Behavior, 25,* 55-61.

STRAUB, R. O. (1979). *Serial learning and representation of a sequence in the pigeon.* Unpublished doctoral dissertation, Columbia University, New York.

STRAUB, R. O., SEIDENBERG, M. S., BEVER, T. G., & TERRACE, H. S. (1979). Serial learning in the pigeon. *Journal of the Experimental Analysis of Behavior, 32,* 137-148.

STRAUB, R. A., & TERRACE, H. S. (1982). Generalization of serial learning in the pigeon. *Animal Learning & Behavior, 9,* 454-468.

TERRACE, H. S. (1986). Positive transfer from sequence production to sequence discrimination in a nonverbal organism. *Journal of Experimental Psychology: Animal Behavior Processes, 12,* 215-234.

WASSERMAN, E. A., NELSON, K. R., & LAREW, M. B. (1980). Memory for sequences of stimuli and responses. *Journal of the Experimental Analysis of Behavior, 34,* 49-59.

WEISMAN, R. G., & DiFRANCO, M. P. (1981). Testing models of delayed sequence discrimination in pigeons: Delay intervals and stimulus durations. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 413-424.

WEISMAN, R. G., DUDER, C., & VON KONIGSLOW, R. (1985). Representation and retention of three-event sequences in pigeons. *Learning & Motivation*, **16**, 239-258.

WEISMAN, R. G., GIBSON, M., & ROCHFORD, J. (1984). Testing models of delayed sequence discrimination in pigeons: The advance key procedure. *Canadian Journal of Psychology*, **38**, 256-268.

WEISMAN, R. G., & VON KONIGSLOW, R. (1984). Order competencies in animals: Models for the delayed sequence discrimination task. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal Cognition* (pp. 199-214). Hillsdale, NJ: Erlbaum.

WEISMAN, R. G., WASSERMAN, E. A., DODD, P. W., & LAREW, M. B. (1980). Representation and retention of two-event sequences in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, **6**, 312-325.

NOTE

1. For the positive sequence, the discrimination ratio must be .5 because it compares the rate of pecking to the positive sequence divided by twice the rate of pecking to the positive sequence.

APPENDIX

We analyzed the rate of pecking to the test stimulus Y_j following each of the $J=27$ sequences as a function of seven different units. Three units represent individual stimuli. S_{1j} represents the stimulus in the first position of Sequence j, S_{2j} represents the second stimulus in Sequence j, and S_{3j} represents the third stimulus in Sequence j. Three units represent pairs of stimuli in the sequence. S_{12j} represents the combination of the first and second stimuli in Sequence j, S_{23j} represents the combination of the second and third stimuli in Sequence j, and S_{13j} represents the combination of the first and third stimuli in Sequence j. The seventh unit, S_{123j}, represents the entire sequence considered as a single unit. The predicted peck rate following each sequence is given by Equation A1:

$$Y_j = k + C_1 S_{1j} + C_2 S_{2j} + C_3 S_{3j} + C_{12} S_{12j} +$$

$$C_{23} S_{23j} + C_{13} S_{13j} + C_{123} S_{123j} \qquad \text{(A1)}$$

The S_{ij} values were estimated on the basis of terminal performance (i.e., the last three sessions). For example, Equations A2 through A5 show how some of the S_{ij} values were calculated. These are the values for S_1 when the sequence begins with A (S_{1a}), B (S_{1b}), and C (S_{1c}). Each value is a deviation from an average peck rate (E_1). In the following equations, *ABC* represents the rate of pecking following the positive sequence ABC, *BBC* represents the rate of pecking following BBC, and *CBC* represents the rate of pecking following the sequence CBC.

$$E_1 = (ABC+BBC+CBC)/3 \qquad \text{(A2)}$$

$$S_{1a} = (ABC-E_1)/E_1 \qquad \text{(A3)}$$

$$S_{1b} = (BBC-E_1)/E_1 \qquad \text{(A4)}$$

$$S_{1c} = (CBC-E_1)/E_1 \qquad \text{(A5)}$$

With the exception of S_{123j}, each of the S parameters was estimated in the same way—as a deviation from the average of all sequences containing the specified substring. For example, S_{12j} was based on each of the sequences that differed from the positive ABC sequence only in the first two positions. One S_{12j}

parameter was estimated for each combination of stimuli in the first two positions by comparing the peck rate following each sequence with the average peck rate for all of the sequences in the set (i.e., one parameter was derived for each of the following sequences that differed from the positive sequence in the first two places *AAC, ABC, ACC, BAC, BBC, BCC, CAC, CBC,* and *CCC*). Similarly, the S_{23j} parameters were estimated on the basis of those sequences that differed from the positive in the last two positions (i.e., *AAA, AAB, AAC, ABA, ABB, ABC, ACA, ACB,* and *ACC*).

In setting the S_{ij} values, we simply made the weak assumption that some units could differ more than others from the corresponding units in the positive string. For example, there seems no a priori reason to believe that the difference between strings ABC and BBC is the same as the difference between strings ABC and CBC. By estimating these differences on the basis of terminal performance, we allowed the birds to "tell" us how different the various units were and avoided making strong arbitrary assumptions or using up excessive degrees of freedom. The same S_{ij} values were used throughout the experiment. With the exception of S_{123j}, the units for each position or combination of positions sums to 0 (e.g., $S_{1a}+S_{1b}+S_{1c} = 0$).

The S_{123j} parameter was set in a different manner. This parameter was set to 1.0 for the positive string and to 0.0 for all other strings. When the string is considered in its entirety, it is clear that this string is either the positive string or the negative string. Intermediate values for this string, then, do not make sense.

Each of the C_i values in Equation A1 is a weight representing the degree to which the corresponding unit controlled responding. Although seven possible units are available for the control of responding, they need not all be equally effective. Some of the units could be very influential in controlling pecking, while others would not be effective at all. The C_i values were estimated for each block of six sessions by the BMDP regression program. The magnitude of the C_i coefficient indicates the degree to which a substring controlled responding. The C_i values correspond, in ANOVA, to the magnitude of the effect specified by the S_{ij} values. In typical ANOVAs, the magnitude of the effect, represented by its mean square value, is used only to decide whether the effect is significantly different from 0, that is, whether or not the value could have been obtained through sampling error alone. In addition to determining whether an effect was present or absent, we were also interested in measuring the size of effect, namely, the degree to which the unit controlled responding. These values are displayed in the paper in the form of standardized regression coefficients. A standardized regression coefficient is the normalized form of the ordinary regression coefficient. The standardized regression coefficient is transformed from the units of the variables (e.g., the number of pecks) into standard units. It predicts the Z-score of the dependent variable on the basis of the Z-score of the independent variable. As a result, the standardized regression coefficient is insensitive to the magnitude and the variance of the variables. Each standardized regression coefficient represents a portion of the explained variance in the peck rate, that is, the control over responding exerted by the corresponding unit.

The use of standardized regression coefficients makes the analysis insensitive to the way in which the parameters of the model were estimated. Any consistent rule that assigns higher values to the parameters denoting positive substrings than to the parameters denoting negative substrings will give approximately equivalent results.

Every statistical analysis requires the investigator to make certain theoretical assumptions (Cohen, 1968; Dwyer, 1983). It should be apparent from the above discussion, however, that Equation A1 presents a means of analysis that is relatively free of theoretical commitments to any particular representational system. The assumptions of the present analysis are similar to the assumptions underlying ANOVA. The main assumption is that variability of the peck rate can be partitioned into units corresponding to the serial positions and combinations of serial positions in the sequence. The analysis, but not the results, is consistent with any of the alternative views of sequence discrimination we have considered. The main difference between our analytic technique and that represented by ANOVA is that we are interested not only in determining whether or not a given effect is reliably present, but also in determining the size of the effect. Our interest in magnitude is the primary impetus for using the analysis represented by Equation A1 rather than a standard ANOVA.

This analysis, like its corresponding ANOVA, reflects the assumption that it is appropriate to analyze the sequence being discriminated into three factors and their interactions (the stimuli and the higher order units). Neither analysis assumes that higher order interactions are necessarily present, but both analyses provide mechanisms for discovering these interactions if they are present. In both analyses, these interactions appear only if they account for variance that cannot be explained on the basis of additive combinations of lower level units. Neither analysis, therefore, presumes the presence of a hierarchical representation for the sequence.