

# Rate and structure in memory for auditory patterns

MARI RIESS JONES, DARYLE J. MASER, and GARY R. KIDD  
 Ohio State University, Columbus, Ohio 43210

Musically knowledgeable listeners heard auditory patterns based on sets of six (Study 1) or eight tones (Study 2). In the first study, listeners ordered events from patterns generated by hierarchical rule trees and which possessed different pitch space and time structures: one type (nondistance nested) was more likely to produce auditory streaming than the other (distance nested). In the second study, different listeners reconstructed pitch intervals contained in one of eight patterns. Patterns differed according to (1) levels of pitch distance (four), (2) levels of pattern contour (two), and (3) rate (two). In both studies, fast patterns with many large pitch distances were more difficult to recollect. Listeners in the second study tended to "telescope" pitch distances. Most difficult were those rapid sequences with large pitch intervals combined into a changing contour (nondistance nested); these patterns streamed. A third study replicated effects due to differences in pitch distances observed in Study 2. Results were interpreted in terms of a rhythmic theory of memory.

As we listen to a music-like sequence unfold, we generate certain expectancies about future pitch and time relations. When these expectancies are fulfilled, pattern recollection should be more accurate, according to the theory of the present research. Expectancy, used in this context, is defined in a recent theory of auditory pattern perception (Jones, 1976b).

To understand this new approach, consider traditional theories of serial pattern memory in which event-event relations were conceived as abstract rules (Greeno & Simon, 1974; Jones, 1974; Restle, 1970; Simon, 1972). In Jones' (1974) earlier system, rules formed mathematical subgroups, and this property provided powerful descriptions of invariant pattern properties. Nonetheless, in this and other systems, rules had little meaning apart from their potential for changing one serial event into another. Imagine, for example, the set, *S*, of six digits (1, 2, 3, 4, 5, 6). A digit pattern, 12324565, generated from this set by hierarchically combining rules defined by Jones (1974) is shown in Figure 1a. This tree structure involves a "next" rule (*N*), a "reflection" rule (*Rℓ*), and a "transpose" rule (*T*), defined in the box of Figure 1. The hierarchical formula of this tree is  $T(Rℓ(N(1)))$ . This formula states that the unit 12 occurs by applying a next rule to the argument (1) and concatenating the resulting digit (i.e., 2) with the argument. Similarly, the reflection

This article was prepared with the support of National Science Foundation Grant BMS 74-21492. The authors are indebted to James Morgan, Pamela Howard, and Robin Wetzel, who assisted in data collection, and to Lewis Hinkle and Harvey Shulman, who contributed technical support. Requests for reprints should be sent to Mari Riess Jones, Department of Psychology, Ohio State University, 404-C West 17th Avenue, Columbus, Ohio 43210.

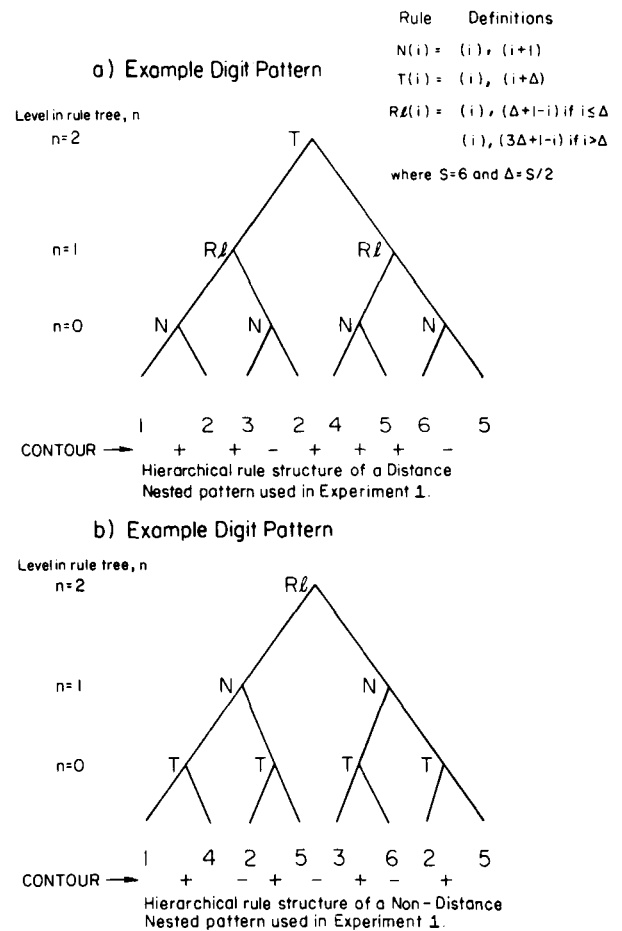


Figure 1. (a) Hierarchical rule tree for digit pattern based on rules, defined for a set of ordered digits  $S = 6$ , conforming to the formula  $T(Rℓ(N(1)))$ . (b) A second hierarchical rule tree conforming to rearrangement of the same rules as in Figure 1a in the formula  $Rℓ(N(T(1)))$ .

rule operates on the digit set but with reference to its midpoint. Thus,  $R\ell(1) = 13$  and  $R\ell(3) = 31$ , while  $R\ell(4) = 46$ , and so on. In combination, rearrangements of rules generate different patterns. The hierarchical combination  $R\ell(N(T(1)))$  yields 14253625, for example, shown in Figure 1b.

A majority of serial pattern research has focused on differences between hierarchical patterns, such as those in Figure 1, and nonhierarchical patterns. Hierarchical patterns are assumed to be psychologically easy because rules are strictly nested, with the result that a simple combination of rules economically describes an entire sequence (Jones, 1974; Restle, 1976; Simon, 1972). When rules are strictly nested, one rule applies uniformly to a given hierarchical level throughout a pattern. In a nonhierarchical pattern, this is not always true. For example, the sequence 12324564 is not strictly rule nested. All digit pairs at the lowest hierarchical level (e.g., 12 and 32), except the final pair (i.e., 64), are related by the next rule: The final pair is related by reflection. In contrast, the sequence 12324565 possesses a next rule throughout the lowest tree ( $n = 0$ ) level, and so is strictly rule nested.

While both patterns of Figure 1 are hierarchical, intuitively, that of Figure 1a seems simpler still than that of Figure 1b. Why should this be so? Differences in difficulty could arise from (1) intrinsic differences in rules themselves (e.g., N and T span different intervals), and (2) differential effects of certain rule combinations (e.g., N at level  $n = 0$  in combination with  $R\ell$  at  $n = 1$  in the pattern of Figure 1a). Kotovsky and Simon (1973) consider the first possibility by ordering rules according to their difficulty. The real issue, however, is not that the rules differ in difficulty, but why they differ. Furthermore, how are these differences modulated by context? A more explicit interpretation of individual rules holds some promise for clarifying both problems.

Such an interpretation is found in Jones' (1976b) recent approach. Rules become distances along a dimension in cognitive space. In this view the digit set represents elements from a number dimension given cognitively as a "number line"<sup>1</sup>, wherein digits are situated certain distances apart. Involvement of cognitive distance introduces a new nesting problem called distance nesting. Note that in the digit sequence 12324565, distances at lower hierarchical levels (e.g.,  $n = 0$ ) are small in value (i.e., one unit), while those at higher levels are increasingly greater (i.e., 2 at level  $n = 1$  and 3 at level  $n = 2$ ), where distance is reckoned from the initial event. The pattern of Figure 1a is distance nested. To distinguish distance nesting from strict rule nesting, a capital N denotes the former. In a distance-nested pattern, relatively small distances occur at lower hierarchical levels and larger distances occur at correspondingly higher levels. In contrast, in the pattern 14253625, relatively large distances (three units) occur at the lowest level, while the distance

at the highest level is small (two units). This pattern is not distance nested. Development of a distance-nesting scheme sets the stage for a differentiation of patterns in Figure 1, in terms of psychological difficulty.

Patterns with hierarchical rule nesting may be either distance nested or not. Differences in nesting are indexed by a distance ratio based on adjacent tree levels (e.g., between levels  $n = 0$  and  $n = 1$ ). The simplest distance-nested structure has a constant ratio of +2 over levels. This yields the digit pattern 12345678 (assuming a set of eight digits). Note that the distance at level  $n = 0$  (e.g., between 1 and 2, or 3 and 4, etc.) is one unit, while at levels  $n = 1$  (e.g., between 1 and 3) and  $n = 2$  (e.g., between 1 and 5) the distances are 2 and 4, respectively, leading to corresponding ratios of  $2/1 = 2$  and  $4/2 = 2$ . This straightforward ascension is one of our simplest patterns. Clearly, the pattern of Figure 1a approximates this simple scheme, although two directional changes make the pattern slightly more complex. Other, less simple, distance-nested patterns are those with a constant integer ratio of distances as  $n$  increases, but the ratio exceeds 2. Non-distance-nested patterns are fractional ratios based on the rule distance of level  $n + 1$  divided by that of level  $n$ . Thus, in the pattern of Figure 1b, the distance of Level 1 divided by that at  $n = 0$  is  $1/3$ . This interpretation of rules as space-like distances provides a simple way of assessing the whole pattern's rule structure in terms of distance ratios. Furthermore, this approach paves the way for a more encompassing representation of serial structure, one that includes not only space-like dimensions but also the time dimension.

Serial patterns are often experienced as patterns in time. A second advantage of tying rules to distances arises if we consider that temporal structure is also given by distances along the time dimension,  $\Delta t_n$ . Not only does the digit 2 following digit 1 (in the pattern of Figure 1a) define a spatial extent, it also defines a time interval,  $\Delta t_0$  (level  $n = 0$ ). Other time intervals span various hierarchical levels, namely,  $\Delta t_1$  and  $\Delta t_2$ , as shown in the temporal hierarchy of Figure 2. Since events occur over time and small time intervals are nested within larger ones, time is nested. Therefore, a change in space and a corresponding time change exist at each level in a pattern's tree structure. Since velocity is distance/time, the space-time structure of a sequence may be a set of velocities (one velocity value for each level). Patterns with constant velocity over hierarchical levels are simplest. If pattern presentation rate is constant, the uniform velocity case corresponds to patterns with a uniform distance-nesting ratio of +2. Patterns with changing velocities are more difficult in proportion to their disparate velocities. These conceptions not only elaborate what is meant by simple patterns, but they also allow predictions about nesting and timing in serial patterning. A person encountering events patterned over time generates simple expectancies

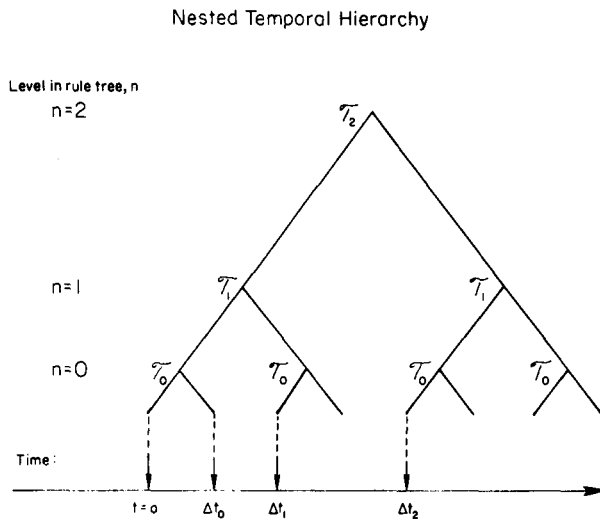


Figure 2. Temporal rules,  $\tau_n$ , corresponding to a hierarchically defined serial pattern delimit time intervals,  $\Delta t_n$ , that each initiate with the onset of the first event at  $t = 0$ . Thus, at  $n = 0$ , the interval between adjacent events is  $\Delta t_0$ , but at  $n = 1$ , the interval is larger,  $\Delta t_1$ , while  $\Delta t_2$  spans the time from  $t = 0$  to the onset of the fifth serial event.

about “where” in space and “when” in time forthcoming events should occur. Simply put, based on initial space-time differences, these simple expectancies mean that people anticipate constant velocities. To the extent that a pattern’s actual velocity deviates from this uniform schema, serial relations within the pattern will be difficult to anticipate and recall. Thus, failure to anticipate future relations can occur either because patterns, such as nonnested ones, are based upon complicated schemes or because patterns (nested or nonnested) occur too fast (or both). Our tracking velocities are limited. Let this limiting speed be identified as the velocity of an expectancy signal which travels along simple paths to “prime” appropriate cognitive locales. If incorrect locales are primed either because the pattern deviates greatly from the expected or because it occurs too quickly, some event-event relations will not be detected, and so the initial scheme cannot be revised. It is a revised expectancy scheme that governs a person’s memory for the event sequence during recall. In this view, perception of an event-event relation determines its recall. It is important to note that perception here refers to an event-event relation (i.e., a velocity value) at some level, and not to detection of individual event occurrences.

In the present research we study effects of pattern structure upon perception of temporal order in auditory tone patterns. While research in serial pattern learning has grown up in the milieu of digit, light, and letter patterns (e.g., Jones, 1976a; Kotovsky & Simon, 1973; Restle, 1970), we know surprisingly little about the application of formal serial order theories to auditory events. In view of commonplace justifications of such

theories in terms of music perception, the dearth of formal applications is disconcerting.

Moreover, auditory patterns present interesting possibilities for reconciling formal theories of serial structure with general auditory perception (for a discussion, see Jones, 1976b, in press). Briefly, tone height and chroma pitch dimension can be taken to be space-like; here we focus upon tone height distances (Jones, 1976b). With respect to the time dimension, work with auditory strings presented at different rates has revealed an interesting and, undoubtedly, pervasive attentional phenomenon known as “auditory streaming” (Bregman & Campbell, 1971; Dannenbring & Bregman, 1976; Dowling, 1973; Jones, 1976b), which acquires meaning in a velocity-oriented approach to serial order. Streaming refers to the phenomenal experience listeners describe with rapid auditory sequences that seem to break apart into cooccurring subparts or “streams.” If, for example, a set of six tones composed of three high tones (H) and three low tones (L) and denoted as  $S = H_1 H_2 H_3 L_4 L_5 L_6$  is the basis for the sequence  $H_1 L_4 H_2 L_5 H_3 L_6$ , a listener may not actually “hear” the veridical temporal sequence. If the sequence is rapid, listeners report hearing either  $H_1 H_2 H_3$  or  $L_4 L_5 L_6$  or both as separate pitch-related “streams” occurring simultaneously. As a result, there is little overall temporal order retention.

Theoretically, streaming is interesting because it is more likely to occur with (1) greater pitch distances between tones (Dannenbring & Bregman, 1976; Miller & Heise, 1950; van Noorden, 1975) and with (2) increments in presentation rate (Dannenbring & Bregman, 1976; van Noorden, 1975). Little systematic work relates streaming to pattern structure, although Heise and Miller (1951) have established that pattern context is important. One question addressed in the present research asks if different lawful rule-governed arrangements of tones from the same tone set vary in susceptibility to streaming. For example, patterns that systematically place tones of greatest pitch distance close together in time could be more susceptible to streaming than those which do not. Patterns with more changes in pitch direction might be more likely to stream than those with few or no directional changes. Directional changes relate to pattern contour, which is the series of ordinal (+, -) pitch differences between adjacent tones. (Throughout this discussion, a “+” indicates a pitch increment and a “-” a decrement.) In distance-nested patterns, smaller pitch distances occur relatively close in time and possess few directional changes. In contrast, nonnested patterns invert distance/time relationships and have more complex contours as well.

One explanation of streaming involves pitch differences ( $\Delta p_n$ ), pattern contour, and presentation rate (Jones, 1976b). For a listener to serially relate pattern tones at any hierarchical level, the pitch velocity

(i.e.,  $k_n = \Delta p_n / \Delta t_n$ ) at that level must not exceed some critical threshold that is determined by the listener's expectancy speed. This limiting velocity is the finite preparation or "prime" speed. If some pitch velocity,  $k_n$ , within the pattern is less than the threshold speed,<sup>2</sup> then a listener is able to anticipate and ultimately recall that event-event pair. This means that recall difficulty in general depends upon how a pattern's structure (e.g., nested vs. nonnested) forces the listener close to such a limiting speed. Patterns with many directional changes effectively enhance the cognitive distance a person must mentally traverse to prime appropriate space-time locales. If events shift back and forth across large distances in pitch space and/or fast rates, listeners simply cannot anticipate. The result is cognitive breakdown. This breakdown is often not manifest in percepts of noise, but as cooccurring auditory patterns, or streams. Theoretically, attentional focusing reverts to several restricted space-time regions (e.g., high or low tones), wherein event-event relations are more likely to possess lower pitch velocities. To the extent that tones within these pitch-related sets conform to simple uniform velocity schemes, it is event-event relations within separate streams that listeners will detect. Consequently, order retention will depend upon within-stream relations (see Jones, 1976b, pp. 336-339, for formal predictions). Because nonnested patterns contain changing pitch velocities, they will be more difficult for listeners than nested ones. And at rapid rates, one consequence of this greater difficulty is that enough deviant relations break threshold speeds that stream-determining schemes take hold. Thus, streaming is one manifestation of perceptual breakdown.

Because other approaches to serial patterning fail to incorporate time as a part of pattern structure (e.g., Kotovsky & Simon, 1973; Restle, 1970), they offer no parsimonious basis for predicting differences between nested and nonnested patterns, nor for conceiving of the phenomenon of streaming at all. Our purpose in Study 1 was to show that two strictly rule-nested auditory patterns differ in difficulty, and so in resistance to streaming, because one is distance nested (N) and the other is not (U). Both patterns and their rule formulas, shown in Table 1, were constructed from a common tone set ( $c^1 e^1 g\#^1 e^2 g\#^2 c^3$  in musical notation shown in Table 1) that spanned two octaves. Tone Patterns N

and U conform, respectively, to the digit patterns illustrated in Figures 1a and 1b, but the N rule now covers a unit distance of 4 semitones<sup>3</sup> (ST) and the T rule transposes 16 ST. Thus, the tone set for these patterns contains three lower tones spaced 4 ST apart and three higher tones similarly spaced. A listener heard each pattern repeated three times at one of two rates (either 5 or 10 tones/sec) and was required to recall each, in writing, immediately after its presentation. While both patterns are hierarchical according to a rule formulation approach, a space-time analysis suggests that the U pattern should be more difficult, and at fast rates, parallel percepts of auditory streams should result.

In Study 2 eight patterns of eight tones were presented to listeners nine times each. In this study, distance-nesting values were parametrically varied to decouple effects of pattern contour and pitch distance while holding constant distance-nesting ratios over pattern levels. These constraints required that the four nested patterns and corresponding nonnested controls be constructed from different sets of eight tones. Study 3 verifies the results of Study 2.

In all studies, of interest were effects of (1) pitch distance, (2) pattern contour, and (3) rate. Patterns with larger pitch distances will be more difficult to remember than those with smaller distances at a given rate. Distance-nested patterns with few directional changes should provide a simpler contour than nonnested ones with many directional changes. And finally, rapid patterns will be more difficult to track and so to recollect; detrimental effects of rate may be most apparent in patterns containing larger pitch distances.

**STUDY 1**

**Method**

**Materials.** Two patterns of contiguous tones (Table 1) were combined with two constant presentation rates (5 tones/sec and 10 tones/sec) based, respectively, on tone durations of 200 msec and 100 msec. Tone frequencies (hertz) are listed in Table 1 along with corresponding musical notation for each pattern.

A NOVA 1200 digital computer from the Data General Corporation controlled pattern timing and event order via a set of analog switches connected in series with outputs of six RCA oscillators. Tone rise times were less than 5 msec. Outputs

**Table 1**  
**Distance-Nested and Non-Distance-Nested Patterns of Study 1**

Pattern Formula		Patterns*							
Nested T(Rℓ(N(c <sup>1</sup> )))	Musical Notation	c <sup>1</sup>	e <sup>1</sup>	g# <sup>1</sup>	e <sup>1</sup>	e <sup>2</sup>	g# <sup>2</sup>	c <sup>3</sup>	g# <sup>2</sup>
	Frequency (Hz)	262	330	415	330	659	831	1046	831
Nonnested Rℓ(N(T(c <sup>1</sup> )))	Musical Notation	c <sup>1</sup>	e <sup>2</sup>	e <sup>1</sup>	g# <sup>2</sup>	g# <sup>1</sup>	c <sup>3</sup>	e <sup>1</sup>	g# <sup>2</sup>
	Frequency (Hz)	262	659	330	831	415	1046	330	831
Common Tone Set, S	Musical Notation	c <sup>1</sup>	e <sup>1</sup>	g# <sup>1</sup>	e <sup>2</sup>	g# <sup>2</sup>	c <sup>3</sup>		
	Frequency (Hz)	262	330	415	659	831	1046		

\*Both patterns are generated by the application of rules applied to the common tone set according to formulas listed in Column 1.

were fed through a Channel Master mixer to a Sony TC-140 cassette recorder. The subjective loudness of the pitches was adjusted so that all tones were equal according to reports of four judges. Subjects listened to the patterns through a set of headphones with a matched impedance level (.08  $\Omega$ ); the average sound level was 71 dB (re .0002 microbar). A display of the tones to be used, noted in musical notation and numbered, was available to the subjects in the form of a drawn keyboard.

**Subjects.** Thirty-two high school students attending marching band camp at Camp Crescendo in Rio Grande, Ohio, during the summer of 1976, participated in the study. Schools represented were Carlisle High School, Carlisle, Ohio, Chillicothe High School, Chillicothe, Ohio, and Walnut Hills High School, Cincinnati, Ohio. Each subject was required to be familiar with the piano keyboard. Each subject also underwent a screening procedure to insure he/she could hear and identify all individual tones as well as basic sequences (e.g., ascending and descending melodic lines). Each qualified subject was randomly assigned to one of four conditions ( $n = 8$ ).

**Design.** The design was a 2 by 2 by 2 by 8 factorial. Rate (5 tones/sec, 10 tones/sec), and counterbalance order (1, 2) were between-subjects variables, and pattern structure (nested, non-nested) was a within-subjects variable. Each subject received one nested and one nonnested pattern at the same rate. Half received the nested sequence first and half the nonnested first. Eight subjects were randomly assigned to each counterbalanced condition.

**Procedure.** Subjects were familiarized with the recall sheet and the appropriate keyboard display. A recall sheet was composed of four rows and eight columns. Subjects were instructed to write responses which signified one repetition of the melodic line across a row of eight blocks. If they heard sounds cooccurring as chords or intervals, they were to indicate that these sounds occurred together by writing them in the same column. Thus, vertical representation using the columns signified chordal or intervallic structure, while the row designations signified melodic lines. Subjects were told to respond in musical notation.

All subjects received one practice pattern followed by two test patterns. Elapsed time between first and second test pattern presentations was approximately 2 min, which subjects occupied by filling out a questionnaire. Subjects received three continuous repetitions each of practice and test patterns.

**Scoring method.** Data were scored in two ways: (1) a total score based on the number correct per scoring position and (2) an octave score (1, 2) which reflected a subject's use of tones from one (1) or both (2) octaves. The first technique took into account a subject's serial organization of the pattern and was the best method with which to characterize accurately both nested and nonnested pattern recall. The octave score indexed streaming, since the splitting of a pattern was usually in terms of tones in one of the two octaves.

## Results

**Total correct score.** Analysis of variance showed that main effects of both structure and rate were significant, confirming predictions of the space-time expectancy theory. Nested patterns were significantly easier than nonnested ones [ $F(1,28) = 12.41$ ,  $MSe = 3.15$ ,  $p < .005$ ]. Slow patterns (5 tones/sec) were easier than fast (10 tones/sec) [ $F(1,28) = 5.45$ ,  $MSe = 7.17$ ,  $p < .05$ ]. An interaction of rate with structure approached significance [ $F(1,28) = 3.36$ ,  $MSe = 3.15$ ,  $p < .08$ ]. Table 2 shows the mean correct total scores for the four pattern conditions.

**Octave scores.** Analysis of variance showed the same effects to be significant as for the total scores. Nested

Table 2  
Mean Scores as a Function of Rate and  
Structure for Patterns in Study 1

	Fast		Slow	
	Total Correct	Octave	Total Correct	Octave
Nested	2.63	1.94	5.00	2.00
Nonnested	1.88	1.44	2.63	1.81

pattern responses showed more use of two octaves than did nonnested pattern responses [ $F(1,28) = 16.61$ ,  $MSe = .11$ ,  $p < .001$ ]. Subjects receiving slow patterns made more use of two octaves than did subjects receiving the fast sequences [ $F(1,28) = 5.44$ ,  $MSe = .14$ ,  $p < .05$ ].

The Rate by Pattern Structure interaction again approached significance [ $F(1,28) = 3.43$ ,  $MSe = .11$ ,  $p < .08$ ]. The mean number of octaves used for the two structures and rates are also shown in Table 2. Subjects receiving nested patterns used both octaves regardless of presentation rate, while the fast rate inhibited use of both octaves with nonnested patterns.

## Discussion

Musically sophisticated listeners displayed greater difficulty recollecting the order of tones presented in the non-distance-nested pattern context than in the nested context. Indeed, if octave scores can be taken as an indication that loss of temporal order in the former case resulted from listeners attending to one pitch-related subgroup of tones, then the nonnested patterns appeared more susceptible to streaming. Indeed, many listeners spontaneously reported experiencing streaming of the nonnested patterns, particularly at the rapid rate. When individuals heard the same notes rearranged into a distance-nested sequence, their recall of the ordering of notes from both octaves was markedly superior, and few remarked upon streaming.

In general, the findings are consistent with the idea that listeners must prepare themselves to detect forthcoming event relations. Neither pattern conformed precisely to a simple schema upon which a space-time expectancy prime might be generated, but the distance-nested sequence more closely approximated such a scheme. Furthermore, if unexpected events within either pattern happen too quickly and/or require too great a cognitive reorientation (relative to a simple scheme), listeners fail to detect relations between events involved. In this respect, faster patterns were more difficult. Even with the nested pattern, as rate increased, the two directional changes specific to the reflection rule caused listeners to trip up. More detailed analyses of serial location errors indicated that errors tended to pile up especially at these two discrepant positions with the nested pattern. By far, the most errors occurred at

**Table 3**  
The Eight Pitch Patterns of Study 2

Condition	C <sub>p</sub> Value	Notation		Serial Position							
				1	2	3	4	5	6	7	8
Nested	2	N1	Hertz	258	290	324	364	410	364	324	290
			Music	C <sub>5</sub>	D <sub>5</sub>	E <sub>5</sub>	F <sub>5</sub> <sup>#</sup>	G <sub>5</sub> <sup>#</sup>	F <sub>5</sub> <sup>#</sup>	E <sub>5</sub>	D <sub>5</sub>
Nonnested	1/2	U1	Hertz	258	410	324	517	290	182	230	144
			Music	C <sub>5</sub>	G <sub>5</sub> <sup>#</sup>	E <sub>5</sub>	C <sub>6</sub>	D <sub>5</sub>	F <sub>4</sub> <sup>#</sup>	A <sub>4</sub> <sup>#</sup>	D <sub>4</sub>
Nested	3	N2	Hertz	258	290	364	410	730	650	517	460
			Music	C <sub>5</sub>	D <sub>5</sub>	F <sub>5</sub> <sup>#</sup>	G <sub>5</sub> <sup>#</sup>	F <sub>6</sub> <sup>#</sup>	E <sub>6</sub>	C <sub>6</sub>	A <sub>5</sub> <sup>#</sup>
Nonnested	1/3	U2	Hertz	258	730	364	1035	290	103	410	144
			Music	C <sub>5</sub>	F <sub>6</sub> <sup>#</sup>	F <sub>5</sub> <sup>#</sup>	C <sub>7</sub>	D <sub>5</sub>	G <sub>3</sub> <sup>#</sup>	G <sub>5</sub> <sup>#</sup>	D <sub>4</sub>
Nested	4	N3	Hertz	258	290	410	460	1642	1463	1035	922
			Music	C <sub>5</sub>	D <sub>5</sub>	G <sub>5</sub> <sup>#</sup>	A <sub>5</sub> <sup>#</sup>	G <sub>7</sub> <sup>#</sup>	F <sub>7</sub> <sup>#</sup>	C <sub>7</sub>	A <sub>6</sub> <sup>#</sup>
Nonnested	1/4	U3	Hertz	258	1642	410	2607	290	46	182	29
			Music	C <sub>5</sub>	G <sub>7</sub> <sup>#</sup>	G <sub>5</sub> <sup>#</sup>	E <sub>8</sub>	D <sub>5</sub>	F <sub>2</sub> <sup>#</sup>	F <sub>4</sub> <sup>#</sup>	A <sub>1</sub> <sup>#</sup>
Nested	5	N4	Hertz	517	581	922	1035	9290	8277	5214	4645
			Music	C <sub>6</sub>	D <sub>6</sub>	A <sub>6</sub> <sup>#</sup>	C <sub>7</sub>	D <sub>10</sub>	C <sub>10</sub>	E <sub>9</sub>	D <sub>9</sub>
Nonnested	1/5	U4	Hertz	517	8277	922	9290	581	36	326	32
			Music	C <sub>6</sub>	C <sub>10</sub>	A <sub>6</sub> <sup>#</sup>	D <sub>10</sub>	D <sub>6</sub>	D <sub>2</sub>	E <sub>5</sub>	C <sub>2</sub>

Note—Subscripts refer to the number of the octave in a 10-octave display (numbered 1-10). Each octave began on C.

Serial Positions 4 and 8, where the pattern's contour deviates from expectancy.

The importance of this study lies in the fact that nested and nonnested patterns were generated from the same tone set and by the same transformational rules hierarchically combined. Only rule arrangement differs. The results suggest that relative locations of rules within the hierarchy are crucial. With nonnested patterns, the transpose rule, which spanned 16 ST, brought relatively large pitch distances into small time intervals. This, when combined with the next rule, at a higher level, created a sequence that contained some large space-time ratios (k<sub>0</sub>), as well as many directional changes.

However, a difficulty with Study 1 is that we cannot determine whether both pitch distance and contour contribute to differences between distance-nested and non-distance-nested sequences or if only one of these variables is the significant factor. Our theory suggests that both should play a role, but these variables could not be decoupled in a study designed to construct both nested and nonnested patterns as rearrangements of tones from a common tone set. Study 2 was designed to clarify these questions.

**STUDY 2**

Eight auditory patterns, four nested (N) and four nonnested (U) were generated from sets of eight pure tones. Patterns are presented in Table 3.<sup>4</sup> The four patterns within each nesting condition all followed a common contour; for N patterns the contour was “++++- - -,” while for U patterns it was “+ - + - - + -” (“+” indicates a pitch increase and “-” indicates a decrease). Within each nesting condition,

the four patterns were differentiated by the amount of pitch distance spanned by generating rules. This structural property is indexed by a pitch expansion ratio,  $C_{pn} = \Delta p_n / \Delta p_{n-1}$ , involving pitch distances at successive hierarchical levels (i.e., n - 1 and n). In this study, C<sub>p</sub> was constant over hierarchical levels, taking on either integer values (m) for N patterns (m = 2, 3, 4, 5) or fractional values (1/m) for U patterns (1/m = 1/2, 1/3, 1/4, 1/5). Larger m values mean that larger values of pitch intervals (relative to some value at a given level) are incorporated at particular levels of a pattern's hierarchical structure. For example, if C<sub>p</sub> = m, then the pattern will be a nested one; if m = 2, then for a given pitch difference at the lowest level (i.e., for  $\Delta p_0 = 2$  ST), successively higher levels (n - 1, 2) will have pitch differences that are multiples of  $\Delta p_0$  and powers of m. Thus,  $\Delta p_1 = 2^1 \Delta p_0 = 4$  ST and  $\Delta p_2 = 2^2 \Delta p_0 = 8$  ST, if  $\Delta p_0$  is 2 ST. Table 4 shows two nested patterns with m = 2 and 3, respectively (N1 and

**Table 4**  
Hierarchical Pitch Intervals Determined by  $\Delta p_n = C_p^n \Delta p_0$  for Some Typical Nested (N) and Nonnested (U) Patterns

Pattern Type	Pitch Expansion Ratio	Pitch Difference*		
	C <sub>p</sub>	0	1	2
N1	2	2	4	8
N2	3	2	6	18
U2	1/2	8	4	2

\*Pitch difference,  $\Delta p_n$ , at each hierarchical level, n, in semitones. Note— $\Delta p_0 = 2$  ST for N1 and N2 and  $\Delta p_0 = 8$  for U1;  $\Delta p_n$  refers to the difference in pitch between two sounds that correspond to a given hierarchical level in a tree-structure representation. (See, for example, Figures 1a and 1b.)

$\Delta p_n$  = Distance in pitch space (in semitone units) at level  $n$ .  
 $\Delta t_n$  = Distance in Time (ms)

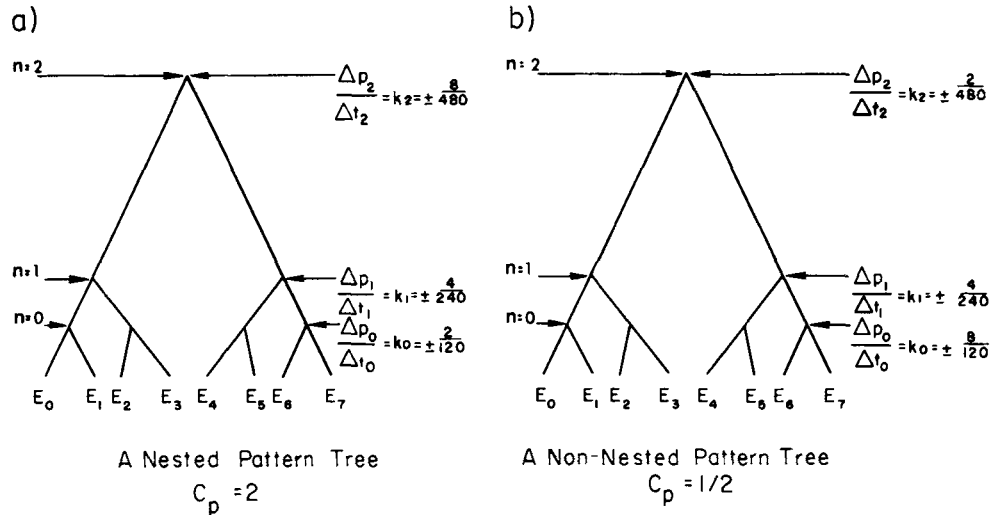


Figure 3. (a) Hierarchical tree structure for the nested pattern with  $C_p = 2$ , used in Study 2. (b) Hierarchical tree structure for the nonnested pattern with  $C_p = 1/2$ , also used in Study 2. Note that in representing hierarchical structure as pitch velocity,  $k_n$ , abstract rules, such as  $N^i$ , that relate only to pitch changes are replaced by space-time ratios.

N2). In general,  $\Delta p_n = C_p^n \Delta p_0$ , so that as  $C_p$  takes on larger integer values, the resulting changes in pattern structure reflect expanding pitch differences. With nonnested patterns, increments in  $m$  also incorporate correspondingly large pitch intervals. Here, however, the smallest pitch interval is at the highest tree level (i.e.,  $\Delta p_2 = 2$  ST); increasingly larger pitch distances occur at successively lower tree levels.

Figure 3 (and Table 3) illustrates the differences between nested ( $C_p = 2$ ) and nonnested ( $C_p = 1/2$ ) pattern types based on a common  $m$  value. In these sequences, abstract distance rules such as  $N$  and  $R\ell$  have been replaced with pitch velocity values,  $k_n$ , at each level. A space-time analysis emphasizes the ratio of the pitch distance spanned at each level (i.e.,  $\Delta p_n$ ) to the time interval at that level (i.e.,  $\Delta t_n$ ).

The  $C_p$  parameter determines contour differences that accompany nesting (i.e.,  $C_p = m, 1/m$ ), while the exact value of  $C_p$  (i.e.,  $m$ ) determines pitch distance. In the present study,  $m$  ranged from 2 to 5 across patterns. Thus, pitch distance, and so the pitch/time ratio, varies according to pattern type within each nesting condition. Of course, this ratio also varies with rate (fast, slow). Faster patterns should be more difficult than slow ones, especially if they contain large pitch distances ( $m$  values). Furthermore, in an expectancy view, these effects should be most apparent on a listener's first encounters with a sequence. At this stage, the listener's use of initial space-time relations to forecast upcoming events is maximal. Accordingly,

in this study, listeners recalled a pattern after the first in a series of presentations. To realistically approach the idea that people generate expectancies about pitch space, these subjects received no advance knowledge of the tone set; they knew only the first tone and the pattern's contour, not the actual tones. In this respect, the second study differed from the first, where listeners were aware of the tones, but not the contour. In Study 1, errors could be justifiably termed "order" errors, whereas Study 2 effectively required tone identification as well, via report of pitch intervals.

After initial pattern presentation (Trial 1), listeners received two more recall tests (Trials 2 and 3, respectively), each based on four pattern repetitions. Listeners were required to follow presented contours in reconstructing sequences on Trials 1 and 2, but not on Trial 3, which preceded judgments about pattern coherence.

To summarize, contour ( $C_p = m$  vs.  $C_p = 1/m$ ) varied with nesting, while pitch distance ( $m = 2, 3, 4, 5$ ) identified pattern type. Pattern rate was either fast or slow across nine presentations of a pattern. We predicted that (1) nested patterns would be easier to recall than nonnested ones; (2) as  $m$  increased within each nesting condition, recall performance would deteriorate; (3) faster patterns would be more difficult than slower ones, especially if large pitch distances (i.e., if  $m$  is large) were involved; (4) all effects would be most pronounced in listener's initial encounters with a pattern.

A final prediction relates to auditory streaming.

Streams form perceptually from the remnants of a broken pattern. Thus, streaming with nonnested patterns should occur only when these sequences have been presented fast enough that several relations exceed the critical threshold. While nonnested patterns are better candidates for streaming than nested ones, at slow rates only nonnested patterns with fairly large pitch distances should stream. Thus, the slow U1 pattern, while difficult to recall, should be least likely to stream; the fast U4 pattern should be most likely to stream.  $\Delta p_0$  is measured in semitones and  $\Delta t_0$  in milliseconds. The  $\Delta p_0/\Delta t_0$  values would, thus, be much smaller for U1 patterns than for U4 patterns.

## Method

**Pattern construction.** Patterns of eight tones were constructed in a manner similar to that of Study 1. A NOVA 1200 Data General Corporation digital computer precisely controlled event timing through a set of analog switches connected in series with RCA oscillators. Outputs were fed through a Studio Mix Master (Switchcraft, Mark VI convertible) to a Sony TC-143SD cassette recorder. All tones within each tone set, presented in ascending order, were judged equivalent in loudness (average loudness = 63 dB).

A pattern's pitch structure was fixed by  $C_p$ ,  $\Delta p_0$ , and an initial note. For nested patterns,  $\Delta p_0$  was 2 ST; this determined  $\Delta p_2$  for a corresponding nonnested pattern. The initial tone for all patterns except N4 and U4 was middle C. The great pitch range engaged by the  $C_p = 1/5$  patterns required some modifications to keep within normal bounds of hearing. The two extreme tones (i.e.,  $D_{10}$  and  $C_2$ ) were adjusted inward by 8 ST from high and low pitch levels, respectively. Table 3 presents the tone patterns, determined by the eight  $C_p$  values, in frequency (hertz) and in musical notation.

Patterns occurred either at a fast (8.30 tones/sec) or a slow (4.16 tones/sec) tempo, based on tone durations of 120 and 240 msec, respectively. Tones were contiguous, with individual rise times less than 5 msec throughout.

**Procedure.** All subjects were screened for ability to judge musical intervals; each subject received three different musical intervals and was required to estimate their size on an octave display drawn to resemble a piano keyboard. Subjects who were unable to accurately estimate two of the three intervals were eliminated.

Recorded instructions described pattern presentation and recall procedure to each subject. A subject heard and recalled one practice pattern consisting of eight tones and one experimental pattern. An experimental pattern was presented nine times, with the listeners being required to recall the notes, in writing, after the first, fifth, and ninth presentations. During pattern presentation and recall, the subject was shown a visual diagram of the pattern's actual contour. The diagram consisted of a series of eight "arrows" of equivalent lengths; an arrow pointed either upward or downward (45 deg) relative to the horizontal to indicate, respectively, an increment or decrement in pitch for the relevant tone in the sequence.

During recall, subjects were required to indicate both the notes and the succession of notes by assigning a number reflecting the note's serial location in the pattern to a note on the octave array. Each response sheet thus consisted of labeled musical notes from the lowest pitch levels to the highest; the initial note was always labeled with the number 1. The subject's task was to assign numbers 2 through 8 to appropriate notes in the octave array. While this may appear somewhat complicated to the musically untrained, the majority of subjects found it a routine and straightforward task. On recall Trials 1 and 2, all listeners were required to follow the presented contour for each pattern; on the last recall trial, contour information was

removed and they were not so constrained. One minute was allowed for recall on each trial. To minimize carry-over effects from the practice pattern, immediately after its presentation, subjects listened to and evaluated (without recall) the musical merit of 45 sec of rock and roll music (taken from "Country Joe and the Fish"). In each condition, practice patterns were haphazard combinations of some tones from the forthcoming experimental pattern; while practice patterns did not contain all eight tones to be used in the forthcoming experimental pattern, they always contained the most extreme sounds.

Finally, subjects were asked to judge the coherence of patterns they heard, using one of two 7-point scales. Depending on their subjective impressions, listeners indicated their confidence that either (1) tones formed a temporal sequence or (2) the pattern streamed.

**Design.** The design was a 2 by 2 by 2 by 4 by 7 factorial involving two levels of contour ( $C_p = m, 1/m$ ), two levels of rate (fast, slow), two levels of trials (first and second recall), four levels of ratio pitch distance ( $m = 2, 3, 4, 5$ ), and seven subjects. Trials formed the only repeated-measures factor.

**Subjects.** The subjects were 112 men and women with at least 2 years of musical training within the past 5 years. Each received \$3 for her/his participation. Subjects were randomly assigned to each of the 16 conditions ( $n = 7$ ); due to an error in the experimental procedure, not all subjects received the third recall trial and questionnaire.

## Results

Performance on each of the first two trial blocks was measured by a deviation score,  $D$ , which was an overall index of the difference between a listener's recall protocol and the actual pattern. Specifically,  $D$  is a cumulative error deviation score:

$$D = \sum_{i=1}^7 |IO_i - IP_i|,$$

where  $IO_i$  is the pitch interval observed in the listener's protocol at a given serial location  $i$  and  $IP_i$  is the corresponding interval in the presented pattern. Table 5 gives the mean  $D$  values for each condition ( $n = 7$ ) on each of the first two trial blocks. On the third recall trial, the  $D$  measure was unsatisfactory because many missing responses occurred in protocols with nonnested sequences. When these listeners were not required to follow the presented pattern's contour during recall, they tended to report only pattern fragments. Accordingly, data from Trial Block 3, which is most

Table 5  
Average  $D$  Scores for  $n = 7$  in Each Rate by  
Pattern by Nesting Condition

	$C_p$ Level	Fast		Slow	
		Trial 1	Trial 2	Trial 1	Trial 2
Nested Patterns	2	7.00	3.71	1.71	2.00
	3	13.14	13.28	21.71	15.57
	4	34.86	29.57	30.86	26.43
	5	44.29	38.71	28.43	31.14
Nonnested Patterns	1/2	26.57	28.57	27.00	26.86
	1/3	89.86	85.43	84.43	80.00
	1/4	155.57	164.00	134.39	121.00
	1/5	269.43	228.57	222.29	231.29



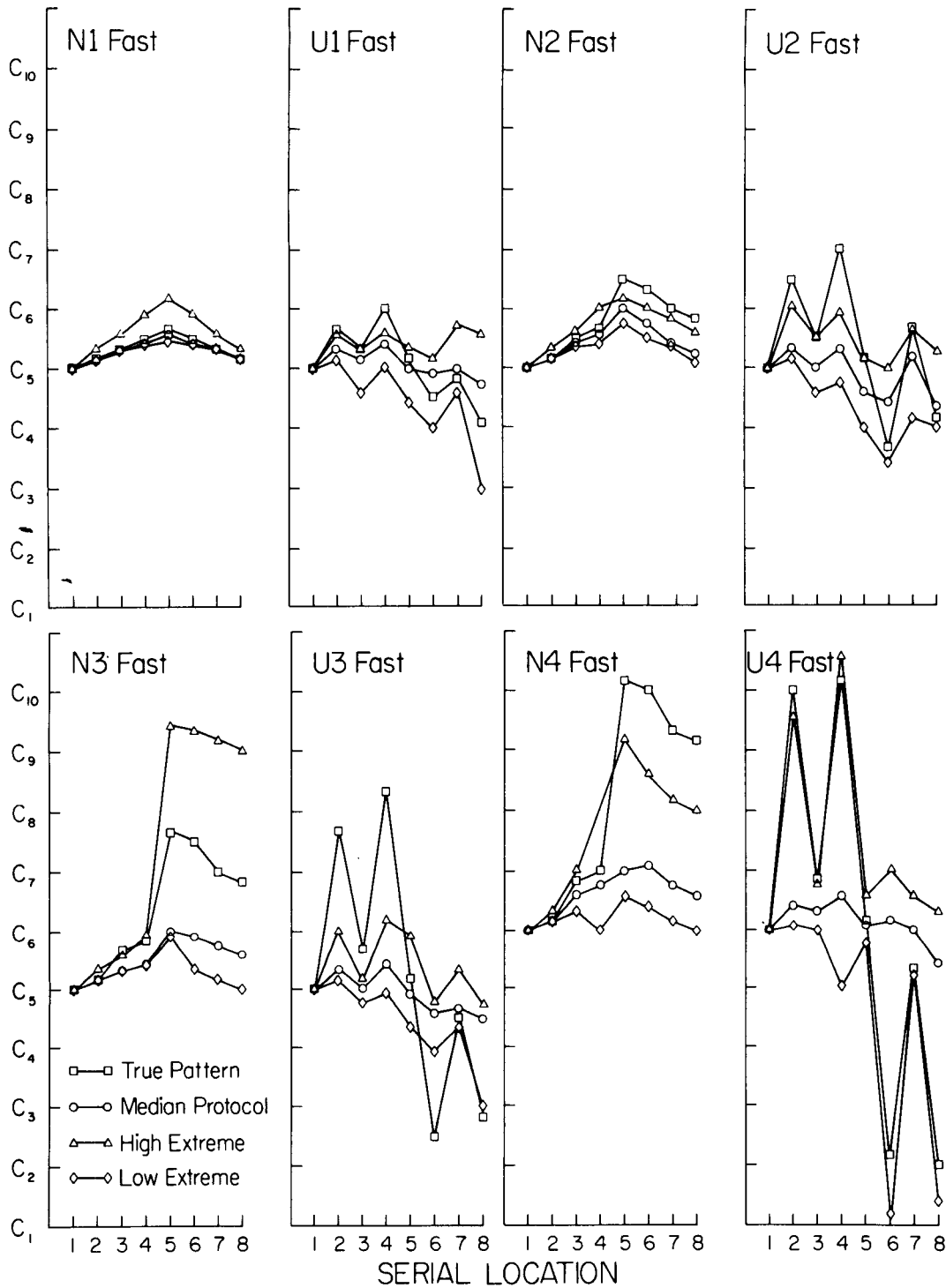


Figure 4. Reported pitch intervals as a function of serial location for the four nested (N1, N2, N3, N4) and four nonnested (U1, U2, U3, U4) patterns presented at the fast rate on the second recall trial. The C<sub>1</sub> through C<sub>10</sub> identify the octave used. Protocols shown for high and low extremes as well as the median protocol were based on the data from seven subjects per condition treated collectively, so that none represents data from a single subject across all serial positions.

relevant to a streaming analysis, are presented separately.

**Trial Blocks 1 and 2.** The nesting variable ( $C_p = m$  or  $1/m$ ) had the greatest impact upon listeners' ability to reconstruct pitch intervals from the patterns they heard. Nonnested patterns that contained five direc-

tional changes were much more difficult than nested ones, which possessed only one directional change [ $F(1,96) = 585.8$ ,  $MSe = 995.4$ ,  $p < .001$ ]. This difference existed across all four matched pairs of pattern types, although as pitch distances increased

(i.e., larger *m* values), the nonnested patterns became disproportionately difficult, leading to a significant Pattern Type by Nesting interaction [ $F(3,96) = 81.6$ ,  $MSe = 995.4$ ,  $p < .001$ ].

As predicted, pattern type had a large effect [ $F(3,96) = 151.6$ ,  $MSe = 995.4$ ,  $p < .001$ ]. Patterns with greater pitch intervals (higher *m* values) were increasingly harder for listeners to recall. Not only did contour differences, reflected by the nesting variable, enhance this effect, but so also did rate and trials. Detrimental effects of larger pitch intervals were most apparent in listeners' initial recall of fast patterns; the three-way interaction of rate, trials, and pattern type was statistically significant [ $F(3,96) = 3.98$ ,  $MSe = 263.2$ ,  $p < .05$ ].

As main effects, rate significantly affected performance [ $F(1,96) = 4.8$ ,  $MSe = 995.4$ ,  $p < .05$ ], but trials, although in the right direction, did not attain significance [ $F(1,96) = 3.5$ ,  $MSe = 263.2$ ,  $p < .10$ ]. Faster patterns were more difficult and performance was poorer after a single repetition than after five pattern presentations.

In Figure 4, protocols are graphically represented as a function of serial location for each of the eight fast conditions of Trial 2. Slow conditions were generally similar, but with somewhat less dramatic telescoping of N3 and N4 patterns. The median pitch intervals of listeners in a condition plus the upper and lower extremes at each location are shown relative to the actual interval sequence for each presented pattern. The most striking aspect of these protocols is the "telescoping" reflected by the median protocol with many patterns based on larger *m* values. In the nested conditions, a common error involved listeners' underestimation of the larger pitch intervals, especially with patterns of *m* = 4 and 5. Increments in rate enhanced this telescoping effect, a finding consistent with the space-time analysis of structure. Subjects reconstructing nonnested patterns, on the average, were even less likely to report the larger intervals, although some listeners focused only upon either the very high or very low tones, thus accounting for relatively accurate extreme responses with some of these patterns.

**Trial Block 3.** On the third trial block, listeners were not constrained to follow the presented pattern contour. This primarily changed behavior of listeners hearing nonnested patterns who were more likely to

**Table 6**  
Proportion of Listeners Correctly Ordering Either Four Middle-Range Tones or High and Low Tones on Trial 3

	Fast		Slow	
		n		n
U1	.43	7	.40	5
U2	.50	6	.00	4
U3	.50	6	.50	4
U4	1.00	5	.75	7

report that they heard a sequence composed of the four middle-range pitches. Protocols of subjects reconstructing nonnested patterns on this trial were classified into two categories: (1) those who reported the ordering only of tones within some subset (i.e., the four middle-range pitches, or two high and two low tones) and (2) those who attempted to follow the previously presented contour. The proportion of subjects listening to each pattern type who fell into the first category are presented in Table 6 for fast and slow rates. Subjects were more likely to focus upon pattern subparts as pitch intervals and rate increased. An exception to this trend is the U2 slow condition, where data from only four subjects were considered.

**Questionnaire data.** The two 7-point scales on pattern coherence (one measuring the degree of confidence in a streaming percept, and the other confidence in a sequential percept) were combined to form one 14-point scale, ranging from very sure of a sequential percept to very sure of streaming. Mean ratings for each condition are presented in Table 7. Rate, contour, and pitch distance all had expected effects on these scores. Stream judgments increased with rate [ $F(1,67) = 7.03$ ,  $MSe = 6.88$ ,  $p < .02$ ] and were far more likely with nonnested patterns than with nested patterns [ $F(3,67) = 18.01$ ,  $MSe = 6.88$ ,  $p < .001$ ] and with larger *m* values [ $F(1,67) = 163.84$ ,  $MSe = 6.88$ ,  $p < .001$ ]. There was also a significant  $C_p$  Level (*m* value) by Nesting interaction [ $F(3,67) = 7.90$ ,  $MSe = 6.88$ ,  $p < .001$ ], indicating that the effect of pattern type on streaming judgments was greater for nonnested than for nested patterns.

Finally, the data of Table 7 indicate that it is not the nesting variable alone that determines whether people judge a pattern to stream. With slow patterns, people found the nonnested U1 pattern as coherent as the nested N1 sequence. It is only as rate (fast U1) or pitch

**Table 7**  
Mean Confidence Ratings of Pattern Coherence

$C_p$ Level	Fast				Slow			
	Nested	n	Nonnested	n	Nested	n	Nonnested	n
2	2.0	5	8.7	7	1.2	5	1.2	5
3	2.8	6	12.5	6	1.5	4	11.8	5
4	2.0	5	12.6	6	2.4	5	12.2	4
5	7.7	4	11.6	5	3.8	5	13.0	7

Note—A score of 1 indicates "very sure pattern was temporally coherent," and a score of 14 indicates "very sure pattern streamed."

distance (slow U2) increases the pitch velocity value that people are more likely to judge a nonnested sequence as less coherent.

Taken together, these findings are consistent with theoretical expectation. However, alternative interpretations of the effects of pattern type are possible. Systematic increments in D scores could have resulted from listeners' tendencies to rely upon equal arrow lengths of contour displays during recall. To rule out this possibility, Study 3 was performed.

### STUDY 3

In this study, contour information was indicated by a series of "plus" and "minus" signs.

#### Method

**Material.** Four fast eight-tone nested sequences of Study 2 ( $C_p = m = 2, 3, 4, 5$ ) were used in Study 3.

**Design.** Four levels of pattern type ( $m$ ) were combined in a one-way completely randomized, analysis-of-variance design.

**Procedure.** Procedures were identical to those of Study 2, except that in this study: (1) listeners recalled note names (e.g., A<sub>6</sub>) of a pattern's last six notes, ordering them appropriately in a row of boxes (instead of numerically identifying them on a keyboard display). Labeled visual keyboard displays (as in Study 2), on which the first two notes of the presented sequence were also shown, were available for reference during recall. (2) Response sheets indicated directions of all pitch differences between to-be-recalled notes as "+" or "-" signs; the initial two notes and their 2-ST intertone pitch interval were shown also on the response sheets. (3) Subjects were carefully instructed about contour information and were told to fill in note names (interval information) only when they had some idea about the pitch distance of the notes involved. Otherwise, they were to leave the relevant box blank. (4) Listeners received one presentation each of a practice and an experimental pattern.

**Subjects.** Twenty Ohio State University students, each with at least 2 years of recent piano training, were randomly assigned to the four conditions ( $n = 5$ ). The median number of years of instrument training was 8.5. Each subject received \$5 for participation.

#### Results

D scores, based on six cumulative pitch intervals for each pattern, are presented in Table 8. Clearly, performance deteriorated systematically with increments in  $m$  for pattern type [ $F(3,16) = 11.99$ ,  $MSe = 427.37$ ,  $p < .0002$ ]. On the average, for a given interval in the  $C_p = 2$  condition, listeners were off the mark by 1.01 ST, while corresponding deviations were 2.03, 5.50, and 12.83 ST (i.e., approximately one octave) for  $C_p = 3, 4, \text{ and } 5$ , respectively. Telescoping again accounted for much of the systematic variation. Especially in recalling the first several notes, many listeners tended to underestimate presented pitch intervals. Using a criterion that defined a telescoped protocol as one in which at least half the reported pitch intervals were underestimated by at least 2 ST, Table 8 indicates that a majority of subjects telescoped in conditions with  $C_p$  greater than 2.

In spite of the fact that there were fewer intervals

**Table 8**  
Average D Scores and Percent Telescoped Protocols ( $n = 5$ )  
in Each Pattern Type ( $C_p$ ) Condition in Study 3

Pattern Type	D Score	Percent Telescoped
2	6.40	40
3	12.20	60
4	33.00	60
5	77.00	100

*Note*—D scores are cumulated over six pitch intervals in Study 3 and over seven in Study 2.

to recall in this study than in Study 2, effects of pattern type were actually greater. Perhaps recall requirements of Study 2 were more natural for listeners. In any case, Study 3 confirms that musically sophisticated listeners are less accurate at reproducing pitch intervals as patterns deviate systematically from simple schemes.

### GENERAL DISCUSSION

These studies indicate that the manner in which pitch distances between tones are arranged is important in describing the difficulty people have remembering an auditory sequence. It is not simply the rule arrangement nor the particular set of tones that determines recall difficulty, as Study 1 illustrated. It is the way in which rules engage small or large pitch differences and place those distances in relation to other pitch differences and in relation to temporal distances that matters. Thus, theories that emphasize formalisms based upon abstract rules must be eventually replaced by ones which emphasize motions in space and time.

In the second study the importance of both pattern contour and pitch distance were established. In this study, the simplest pattern, namely N1, in which  $C_p$  was 2, most closely approximated an ideal schema of the strictly ascending pattern. In order to keep contour constant within nesting conditions and to avoid this exceptionally simple case of straight ascension, the N1 and other nested patterns all contained a single directional change (after Serial Location 5). Nonetheless, to subjects hearing the N1 pattern, this presented little difficulty; indeed by the second recall trial, differences in recall of this pattern as a function of rate were slight, with an average D score at the slow rate of 2.00 and at the fast, of 3.71. Other nested patterns with equivalent contours became more difficult for listeners to reconstruct as the pitch intervals at levels  $n = 1$  and  $n = 2$  expanded with the  $m$  ratio of 5 typically yielding the most difficult sequence. Largest differences in difficulty as a function of  $m$  came between the N1 pattern and others.

The contour of nonnested patterns clearly caused major problems for listeners. Even with the simplest pattern, U1, average D scores were between 4 and 13 times those of the N1 pattern. Contour differences that

separate nested from nonnested patterns relate to the relative placement of pitch distances at different hierarchical levels. Thus, in the nonnested sequences, the large pitch differences are placed close together in time, while the reverse is true of nested patterns. Therefore, not only do nonnested sequences contain many more changes in direction than nested patterns, but these directional changes engage large pitch intervals. This undoubtedly also contributes to the finding that the nesting variable produced a larger effect in this study than in Study 1. Other contributing factors are probably task and differences in tone set. However, it is unlikely that nesting differences can be attributed entirely to the fact that nonnested patterns possessed greater pitch ranges than nested counterparts. The nested pattern with a pitch range of over five octaves ( $C_p = 5$ ) was easier than the nonnested sequence with a two-octave range ( $C_p = 1/3$ ). Results of Study 1 also argue against this interpretation. Nonetheless, differences in contour dictated by nesting do effectively increase critical pitch distances if we assume that listeners actively anticipate "where" forthcoming notes will occur in pitch space using simple schemes. With nonnested patterns, simple schemes are drastically inappropriate and require great amounts of cognitive "backtracking" in pitch space. That distance in pitch space is important, apart from contour, is apparent in the large effects of pattern type ( $m$  value). The argument that directional changes magnify the effect of the pitch distance a listener must cope with receives support from the finding that increments in  $m$  were more detrimental to pattern coherence in nonnested patterns than in nested ones, even with contour held constant. In fact, the U3 and U4 patterns, where  $m$  was 4 and 5, respectively, were exceedingly frustrating for listeners. Many greeted with relief final recall where contour could be ignored.

The argument here is that a revised expectancy scheme finally governs perception of serial relations and pattern recall. In this way, both the level of recall difficulty over a range of patterns and qualitative phenomena such as streaming can be addressed. One criticism of this view holds that nesting differences result from reconstructive differences during recall, not from schemes refined during listening. While this interpretation is also possible, it merely shifts the explanatory burden to an undefined "output mechanism," which in turn must be systematically affected by pattern contour, pitch distance, and rate. It is more parsimonious to argue that a pattern's space-time structure affects whether a particular relation is detected in the first place. Crucial to the space-time view is the proposition that event-event relations, namely pitch velocities (i.e.,  $k_n$ ), can be directly perceived. To the extent that deviant relations are not accurately perceived, listeners apparently "simplify" their memory for pattern relations to conform to the

uniform velocity case. This proposal is at variance with the idea that individual sounds are first perceived as separate events and subsequently organized into a pattern.

As for the difficulty musically sophisticated listeners exhibited with nonnested patterns, both Studies 1 and 2 suggest that the difficult sequences were also more likely to stream at fast rates. It is not simply pattern contour that determines streaming, for the nonnested U1 pattern did not stream. Indeed, listeners recalling this nonnested pattern were more accurate than those recalling the N3 and N4 nested patterns. It was the U3 and U4 patterns that were most likely to stream and, with the exception of U4 (Trial 2), the faster rate enhanced stream formation.

In summary, the way in which time and pitch space combine to produce simple or difficult serial patterns can be meaningfully described in terms of pitch/time ratios. Patterns that possess several extreme pitch velocities are difficult to track cognitively and so to recall; at rapid rates these patterns are more susceptible to streaming.

#### REFERENCES

- BREGMAN, A. S., & CAMPBELL, J. Primary auditory stream segregation and perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, **89**, 244-249.
- DANNENBRING, G. L., & BREGMAN, A. S. Stream segregation and the illusion of overlap. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 544-555.
- DOWLING, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, **5**, 322-337.
- GREENO, J. G., & SIMON, H. A. Processes for sequence production. *Psychological Review*, 1974, **81**, 187-197.
- HEISE, G. A., & MILLER, G. A. An experimental study of auditory patterns. *American Journal of Psychology*, 1951, **64**, 68-77.
- JONES, M. R. Cognitive representations of serial patterns. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition*. Potomac, Md: Lawrence Erlbaum, 1974.
- JONES, M. R. Levels of structure in the reconstruction of temporal and spatial serial patterns. *Journal of Experimental Psychology: Human Learning and Memory*, 1976, **2**, 475-488. (a)
- JONES, M. R. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 1976, **83**, 323-345. (b)
- JONES, M. R. Auditory patterns: The perceiving organism. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 8). New York: Academic Press, in press.
- KOTOVSKY, K., & SIMON, H. A. Empirical tests of a theory of human acquisition of concepts for sequential patterns. *Cognitive Psychology*, 1973, **4**, 399-424.
- MILLER, G. A., & HEISE, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, **22**, 637-638.
- RESTLE, F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 1970, **77**, 481-495.
- RESTLE, F. Structural ambiguity in serial pattern learning. *Cognitive Psychology*, 1976, **8**, 357-381.
- SIMON, H. A. Complexity and the representation of patterned sequences of symbols. *Psychology Review*, 1972, **79**, 369-382.

VAN NOORDEN, L. P. A. S. Temporal coherence in the perception of tone sequences. Unpublished doctoral thesis, Eindhoven University of Technology, 1975.

#### NOTES

1. Although symbols have served as alphabets for pattern generation whenever the members possess ordinal properties, the idea of associating these alphabets with cognitive dimensions is new. In the study by Jones (1976b), the cognitive dimensions involved were those of pitch and loudness, although the general idea is applicable to any physical (e.g., intensity) or conventionally derived (e.g., letters, digits) alphabet.

2. A semitone is the smallest interval in traditional Western music. If  $F$  is the frequency of a given tone, then a semitone distance from  $F$  is given by that frequency increment,  $\Delta F$ , which conforms to  $1 \text{ ST} = \Delta F/F \cong .0595$ .

3. Identification of the T rule with a distance of 16 ST

instead of 12 (with a serial difference of 4) modifies the rule definition of Figure 1 slightly. This modification was required to avoid the complication of involving spurious harmonic intervals in one of the patterns. As they are constructed, patterns of Study 1 represent virtually the only possible pair of patterns based upon a common tone set of Size 6 that are roughly equivalent in harmonic level and satisfy other constraints related to rule arrangement and event redundancy.

4. The N1 pattern, while generated from a tone set of eight, possessed only five tones. This resulted from the constraint that all nested patterns have equivalent contour. In order to avoid the trivial case of strictly ascending patterns, in which N1 would have eight tones, the single change in contour necessitated event repetition.

(Received for publication August 9, 1977;  
accepted January 6, 1978.)