Perception of Structure in Short Melodic Sequences

Article *in* Journal of Experimental Psychology Human Perception & Performance · September 1981
DOI: 10.1037/0096-1523.7.4.869 · Source: PubMed

3 authors:

Lola L. Cuddy
Queen's University

Annabel Cohen
University of Prince Edward Island

Douglas J Mewhort
Queen's University

Journal of Experimental Psychology:
Human Perception and Performance
1981, Vol. 7, No. 4, 869-883

0096-1523/81/0704-0869$00.75

Perception of Structure in Short Melodic Sequences

Lola L. Cuddy, Annabel J. Cohen, and D. J. K. Mewhort
Queen's University, Kingston, Ontario, Canada

Three experiments studied the perception of tone sequences having various degrees of musical structure. Ratings of perceived structure and ease of recognition in transposition were both influenced by harmonic progression (as defined by music theory), the contour (directional changes in pitch), and the excursion or repetition pattern within the sequence. The relation between the original and transposed sequence also affected ease of recognition in accordance with the number of tones shared between the two sequences. The results are described in terms of the abstraction and analysis of levels of pitch relations, an analysis conducted even by musically untrained listeners. The conceptual framework emphasizes the application of musical rules as an illustration of rules governing auditory sequences in general.

Musical context influences the detection of an absolute or relative pitch alteration, usually in the direction of enhancement. We have previously argued that the degree of improvement in judgment is related to the success with which the listener detects and applies musical rules (Cuddy, Cohen, & Miller, 1979). Included under the rubric of musical rules are the codified structures of classical musical analysis—the notions of scale and tonal relationship (cf. Deutsch, 1977; Dowling, 1978). When the properties of an auditory pattern are varied according to musical rules, musically trained listeners will show corresponding variations in accuracy of note naming in an absolute judgment task (Cuddy, 1971), judgment of intervals and of triads (Cuddy & Cohen, 1976; Cuddy, Cohen, & Dewar, Note 1), and recognition of untransposed (Dewar, Cuddy, & Mewhort, 1977) and transposed (Cuddy et al., 1979) melodies. Dowling (1978) and Francès (1972) have also shown in several series of experiments that listeners differentiate melodic fragments that retain the rules of musical tonality from those that do not.

Not all of the above findings are restricted to musically trained listeners. Listeners with minimal formal training in music find the tasks more difficult than do their trained counterparts. Their performance indicates, however, that they do respond to the relational information provided by a musical tone sequence (Attneave & Olson, 1971; Cuddy et al., 1979; Dewar et al., 1977).

The present paper extends an earlier report (Cuddy et al., 1979) in which broad categories of sequences were defined according to musical rules. Now a coding system for individual sequences is developed, taking into account both the formal rules of musical analysis and also the patterns of melodic contour (Dowling & Fujitani, 1971). We again ask to what extent formal musical analysis is representative of perceptual structure and whether the representations hold across levels of musical training. As a further extension of our work, we will consider the relevance of the results to general issues of serial processing.

The classification of sequences as "musical" (i.e., those drawn from a musical text) or "random" (i.e., those in which rules of musical structure are violated and in which musical pattern is difficult, if not impossible, to detect) yields highly reliable differences in performance when overall ease of recog-

This research was supported by a grant to L. L. Cuddy from the Natural Science and Engineering Research Council of Canada and the Advisory Research Committee of Queen's University, Kingston, Ontario, Canada.

We thank J. Pinn, G. Millington, M. Beesley MacHattie, and P. Dobbins for technical assistance, and M. G. Wiebe for editorial comments.

A. J. Cohen is now at the Department of Psychology, Scarborough College, University of Toronto.

Requests for reprints should be sent to L. L. Cuddy, Department of Psychology, Queen's University, Kingston, Ontario K7L 3N6, Canada.

nition is compared between the two categories. This classification is not sufficiently precise to predict performance differences arising from the individual characteristics of sequences within a category. For example, the prediction that recognition of a musical sequence transposed to a musically near key (the dominant) should be more accurate than recognition of the sequence transposed to a musically distant key (the tritone) was confirmed in the earlier report. However, the effect did not hold for all musical sequences (cf. also Francès, 1972). A more specific characterization of individual musical sequences—one corresponding as nearly as possible to auditory perceptual organization—should facilitate greater experimental control of recognition effects.

Our contention, therefore, is that musical rules describe relations that contribute to the perceived structure of melody. The next section will briefly outline a conventional coding system applied by music analysts to describe properties of musical sequences. The code describes the unfolding in time of frequency relations based upon three-note or four-note patterns contained within the major scale. Following the outline of the code, we will explore the notion that the uncertainty specified by the code as a measure of stimulus uncertainty is a correlate of perceptual functioning.

Symbols of Musical Analysis

The purpose of musical analysis is to elucidate music: to provide, for a piece of music, the "best fitting" logic in terms of musical rules. The choice of the best fit depends upon the analyst's knowledge of the redundancies (or grammar) of music (Green, 1965) and, as we shall see, to some extent upon individual persuasion. Nevertheless, all conventional analysis starts from notions of scales and triads.

Figure 1 (first stave) represents a major scale in the key of C. The notes are C_4–D_4–E_4–F_4–G_4–A_4–B_4–C_5. The interval between adjacent tones is indicated in semits. All major scales contain the same ordered set of intervals of whole tones (2 semits) and semitones (1 semit). Below each note a Roman numeral indicates the serial position of the tone in the scale. Roman numerals, therefore, refer not to specific notes or pitches, but to scalar functions. Associated with each tone of the scale are musical functions: I is the tonic, the keynote, the most important note in the scale to which all other notes are assumed to refer. V is the dominant, the second most important note and, as the central component in the I–V–I progression described below, contributes to the sense of completed motion. VII is the leading tone (LT) and it is assumed to create a sense of leading to the tonic, one semit above.

For musical analysis, Roman numerals are also used as symbols for three-note chords formed at each position of the scale. These chords, called triads, are formed by choosing every other successive note in the scale. Here, the Roman numeral is used to indicate the scalar position of the lowest note of the triad.

Triads for the C major scale are shown in Figure 1 (second stave). The interval between tones of the triad is indicated in semits to the right of the musical notation. For every major scale, there are three types of triads: major (+), minor (−), and diminished (o). Major triads, found on positions I, IV, and V, consist of an interval of four semits (the major third) beneath an interval of three semits (the minor third). For minor triads, found on II, III, and VI, the intervals are reversed; the minor third occurs beneath the major third. The single diminished triad on VII contains two minor thirds.

Intervals of a major or minor third may be added to the basic triadic patterns of Figure 1 (second stave), creating four- (or more) note patterns. Symbolic notation here is necessarily more detailed, but only one point is immediately relevant to the present studies. If an interval of a minor third is added to the triad at position V, a chord known as the "dominant seventh" is formed. It is so called because the bottom and top notes of the chord form a musical seventh (10 semits). The notation is V_7. In Figure 1 (second stave), the note completing the dominant seventh is shown in brackets at position V. Observe the similarities between V_7 and VII. They share three tone frequencies—in our example, the notes B–D–F. Confronted with these three notes in the

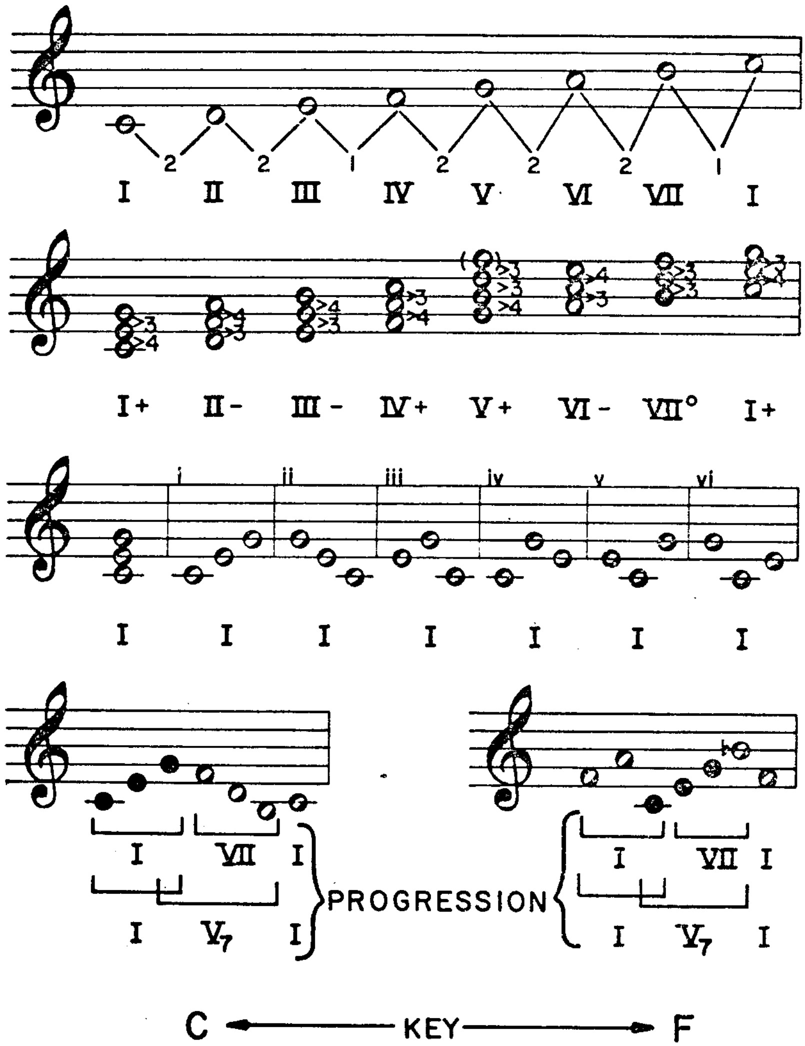

Figure 1. First stave: Diatonic major scale with C as tonic (I). (Size [in semits] of intervals between adjacent tones is shown. Roman numerals signify the ordinal position of each tone in the scale.) Second stave: Triads formed on each note of the diatonic major scale. (Size [in semits] of intervals between adjacent tones is shown next to the triad. Triad type—major [+], minor [−], or diminished [o]—appears below notation. Bracketed note over the triad on V is the added seventh tone resulting in the V_7 chord.) Third stave: A major triad chord followed by six melodic exemplifications. Fouth stave: Two different melodic sequences of seven notes each, with two possible chord progression analyses. (The solid notes form a major triad [CEG] in each sequence.)

scale of C, the analyst has a choice of notations, depending on context. The choice V_7 may be made if the analyst wishes to emphasize a harmony at the dominant.

Figure 1 (third stave) illustrates the six possible sequential configurations of a triad and demonstrates the point that musical analysis does not differentiate among various sequential orders or contours. In the key of C, all orders of the notes C–E–G represent an underlying harmony at the tonic (I). That is, sequential relations within a group of notes are not as important as sequential relations between groups of notes. One of the problems of analysis is to identify the groups of notes to which sequential rules are applicable. As described below, the solution involves two stages.

Samples of musical analyses for sequences of the type used in our experiments are shown in Figure 1 (fourth stave). The samples illustrate two steps of analysis: first, the identification of the key, tonality, or scale of the piece; and second, the identification of the sets of tone relations—in particular, chord patterns—within the assumed key. The notes of a melody are sequentially grouped according to the triadic pattern to which they belong. The sequence of patterns is called a chord or harmonic progression. For example, the progression I–V–I represents a melody in which the notes of the tonic triad are followed by the notes of the dominant triad resolving back to the tonic. The process of analysis may appear to be one of increasing differentiation (i.e., from the identification of the tonality to the identification of structures within the tonality), but it also involves the application of musical rules on many levels simultaneously. Assignment of a key is not always straightforward and, as a result, the analyst may work backward from the possible chord relations to a tonality. In such cases the decision is based on knowledge of likely chord progressions.

In the present example, the analyst has chosen the key of C for the first sequence and the key of F for the second sequence. This assignment of key is supported by the following musical reasons: (a) All notes of the sequence belong to the key; (b) the final note is the tonic; and (c) a progression known as a cadential ending resolves to the tonic of the sequence. All three points are satisfied only for the assigned key. However, two alternative (and, as described above, closely related) analyses are provided for each sequence—the progression I–V_7–I or the progression I–VII–I. The sample analyses also illustrate that analysis is not concerned with the order of tones within triadic patterns and is not concerned with the identification of absolute frequencies. Note, for example, that different serial orders of the opening triad are both symbolized by the notation I, that different frequencies of the opening triad are symbolized by I, and that identical frequencies C–E–G (blackened notes) are given different symbols depending upon the tonal context. Musical analysis therefore deemphasizes or overlooks sequential relations within a set of notes and emphasizes sequential relations between sets of notes.

Musical Analysis and Perceptual Structure

Musical analysis can be employed within an experimental framework to refine and to clarify our understanding of auditory perceptual structure. Quite simply, the analysis may be regarded as a set of hypotheses about what is heard in music. For example, the identification of the tonic note of a melody by musical analysis implies that the note fulfills a central or focal role in the perception of the structure of the melody (Helmholtz, 1885/1954). The description of chord progressions implies the detection of sequential relations among tone sets. Further, the task of musical analysis is an instance of musical problem solving (Laske, 1977). Where the solution to the problem offers alternative descriptions, it may be stated, in information-theoretic terms, that the stimulus uncertainty has increased. Our prediction is that the amount of stimulus uncertainty will be reflected in the relative difficulty with which the listener apprehends the auditory structure (cf. Garner, 1970).

The relevance of classical music analysis to perception may, according to our current argument, be much more direct than that of the mathematical abstractions that generate serial music. As for the latter, it has been hotly disputed whether its formal representations have any contact whatsoever with the auditory process (e.g., Bernstein, 1966). The composer and musicologist Graham George has pointed out that, even within traditional frameworks, "the history of musical analysis is littered with discussion of pseudo-structures existing only in the minds of their theorist-inventors" (George, 1970, p. 30). George argues that the appropriate test of the validity of tonal analysis is the correspondence of the analysis with the "consciously apprehensible elements of the musical structure" (p. 30). Our experiments are designed to provide tests of this correspondence, following the directions indicated by our earlier work.

The musical analyses of the sequences for the experiments were obtained from experienced musicians—teachers and composers of music in a music faculty. The measures of perceptual organization were twofold—accuracy of recognition of a sequence in transposition and ratings of perceived structure. Various levels of musical experience within a university population were sampled across experiments. While to some extent the choice of levels tested for a particular experiment was constrained by the availability of listeners, the levels across experiments ranged from little or no training to that of the professional musician.

Experiment 1

The use of structure ratings as an experimental technique for demonstrating what perceivers "know" about simple patterns has the precedent of Garner and Clement (1963), Garner (1974), and Vitz and Todd (1969). In all cases, in accordance with information-theoretic notions, structure ratings were correlated with stimulus uncertainty (defined in terms of an objective stimulus code). For example, patterns whose code indicated a high level of uncertainty received correspondingly low structure ratings. The first experiment examines a similar relation for musical analysis and rating of seven-note sequences.

Thirty-two seven-note sequences were constructed according to the following procedure. First, a prototypic sequence was composed based upon a priori principles of tonality that should lead to the highest rating of perceived structure. The prototypic sequence was constructed to contain the strongest musical cues to tonality—these include diatonicism, the I–V–I progression, the arrangement of the notes of the triad at position V so that the leading note of the scale (note VII) moves directly to the tonic (I), and the beginning and closing of the sequence on the tonic note. This last point, which we refer to as the excursion of the sequence, was discussed by Helmholtz (1885/1954) with reference to the development of tonal music.

> The tonic note, as the connecting core of all the tones in a regularly constructed melody, must be heard on the first accented part of a bar, and also at the close, so that the melody starts from it and returns to it; the same is true for the tonic chord in a succession of chords. (p. 296)

In addition, two psychological considerations were taken into account. Pitch contour alone may influence ratings of "good form" (Meyer, 1956). Within the constraints of the tonal rules, the notes of the prototypic sequence were arranged in the simplest possible contour. Finally, since repetition affects memory for the pitch of tones (Deutsch, 1972), its presence should be carefully controlled. In the present case, repetition of only the tonic was allowed. The sequence shown on the left-hand side in Figure 1 (fourth stave) fulfills all of the above requirements and was chosen as the prototypic sequence for the experiment.

The remaining 31 sequences were variations and alterations of the prototypic sequence, intended to degrade the structure of the prototypic sequence by degrees. The alterations included changes of the harmonic progression, violations of the rules of diatonicism, variations of contour, and combinations of the manipulations.

Method

Sequences. The prototypic sequence is shown in musical notation as Sequence 1 in Figure 2, and the derived sequences are given in order of the final mean ratings of perceived structure for the most highly trained group of listeners. The prototypic sequence is notated in the key of C and, for purposes of comparison, all derived sequences are shown as transformations from that key. The exact alteration of the prototypic sequence for each of the derived sequences is provided in Cohen (1977, Experiment 2), but it is not necessary to pursue these details to understand the main results of the experiment. A few descriptive examples will suffice: Sequence 2 preserves diatonicism, but alters the V–I and the leading-note-to-tonic progression. In addition, the contour is slightly smoothed. Sequences 3 and 4 preserve the diatonicism and progression of Sequence 1, but alter the contour and the excursion of the tonic (Sequence 4 only). Sequences 10, 13, 15, 21, and 27 each involve one violation of the rule of diatonicism. Sequences 17, 22, 24, 26, 28, 30, 31, and 32 each involve several such violations.

Musical analysis and analysts. Each of the 32 sequences was transposed to a frequency location randomly selected in the treble clef and copied onto musical manuscript paper with accidentals replacing (where relevant) the key signature. Each sequence was copied on the top of a separate page with space below for the musical analysis. The analysts—two professors of music

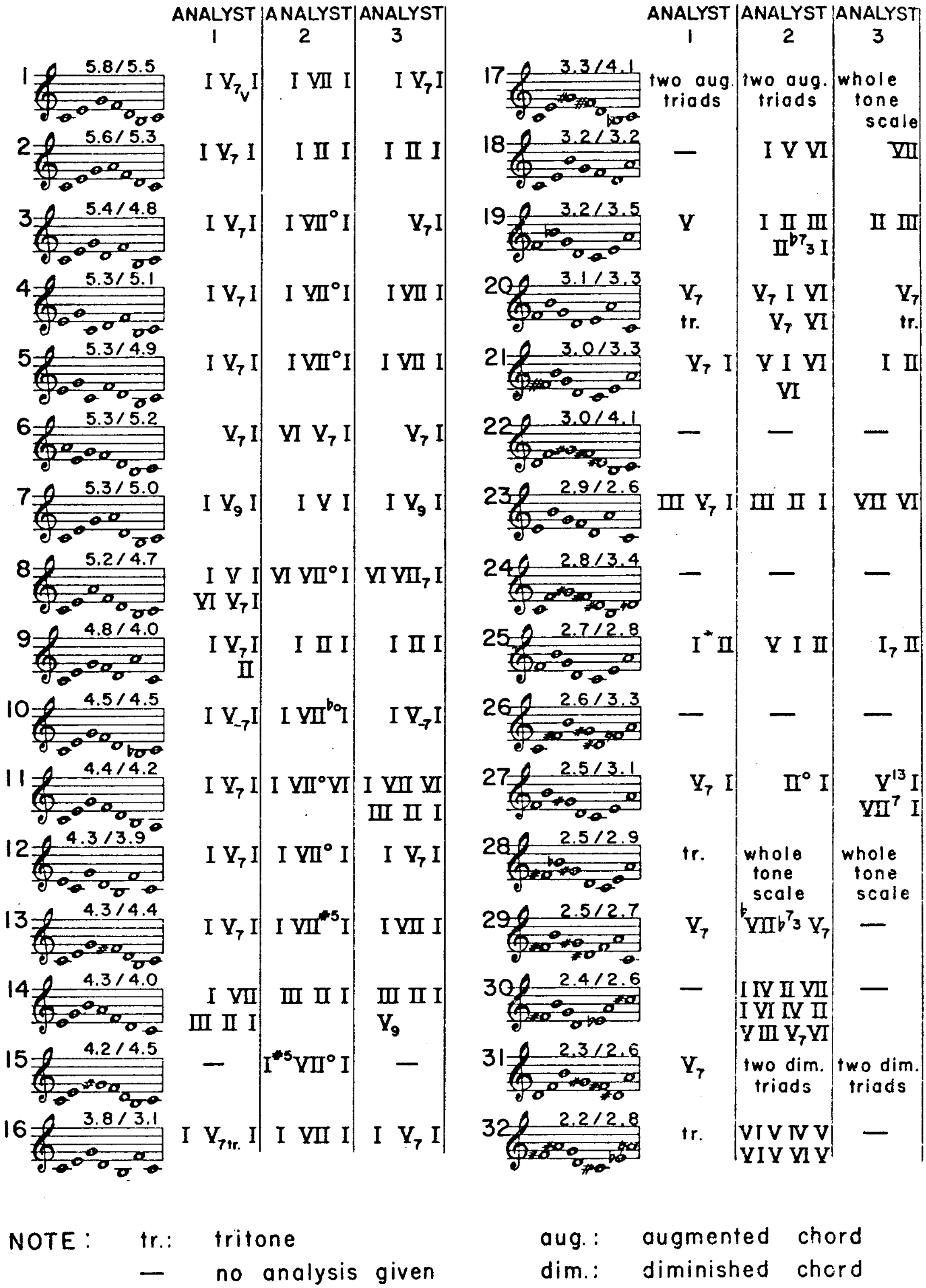

Figure 2. Sequences, mean ratings, and musical analyses for Experiment 1.

theory and a graduate student in musical composition—were asked to identify both the keys and the harmonic progression(s) implied by each sequence. No specific time limit was imposed. The analyses were collected at the end of one week.

Structure ratings and listeners. The sequences were prepared for auditory presentation by recording on tape (Scotch 202, using a Crown PRO-800 recorder at 7.5 ips). The notes were sine tones from the tempered system over the range 246.9 Hz to 987.8 Hz. The tones were generated by a coherent-decade frequency synthesizer (General Radio 1161-P1) controlled by a minicomputer (Digital Equipment Corporation PDP8/I). To minimize transient clicks, the tones were low-pass filtered with cutoff set at 1,250 Hz. Each tone was presented for 500 msec, and the intersequence interval was 6.5 sec. The stimulus sequences were transposed to randomly selected frequency locations within the overall range and were recorded in two different random orders. In addition to the 32 sequences, 10 practice trials were recorded. Practice trials did not duplicate any of the main test sequences.

Listeners were 120 undergraduate volunteers. All indicated an interest in music and would be considered musically trained according to our earlier classification system (Cuddy & Cohen, 1976; Cuddy et al., Note 1). Members of the highly trained (H) group ($n = 60$) had achieved, on the average, Grade X Conservatory performance level[1] and played at least two instruments regularly. The listeners of the trained (T) group (also $n = 60$) had some informal training on one musical instrument and, typically, sang in a choir or were faithful attenders of concert series. None of the latter group had formal training in musical analysis.

The listeners were asked to rate each sequence in terms of its tonality or tone structure by placing a check on a 6-point scale. They were asked to reserve the highest ratings for sequences with "musical keyness" or "completeness" and to assign lower scale values to sequences that contained "unexpected" or "jarring" notes. Half of the listeners within each group received the trials in one order; the rest received the other order. The sequences were presented binaurally on headphones (MB Electronic Model MBK 68) in sound-attenuated rooms.

Results

The mean rating for each sequence for each group of listeners, and the musical analysis provided for each sequence by each analyst, are shown in Figure 2.[2] The sequences are ordered in descending order of mean rating for group H. Above the musical notation for each sequence are two values separated by a slash—the first value is the mean rating for group H, the second, the mean rating for group T. Although the range of mean ratings was slightly wider for group H than for group T (2.2 to 5.8 for group H; 2.6 to 5.5 for group T), the correspondence between groups was very high, $r(30) = .91$, $p < .001$. The prototypic sequence was rated highest in tonal structure by both groups; 52 of the 60 listeners in group H rated the sequence with "full marks" of 6.

The Roman numeral symbols for each sequence reflect each analyst's attempt to encode the harmonic progression of the sequence (cf. Figure 1, fourth stave, and accompanying text). Each analytic solution is placed on a single line, and the figure shows that the code for most sequences involved two or three successive chord progressions. Where an analyst provided alternative solutions, the solutions are recorded on successive lines. A blank means that no solution was found possible by that analyst. Occasionally (particularly as structural ratings decreased), analysts used variants of the basic system described above. The point of major relevance in Figure 2, however, is simply this: As the perceptual ratings decrease, there is an increasing diversity of musical analyses. From Sequence 15 onward, there is an increasing tendency to report that no solution is possible. As the rules of diatonicism are violated, there is a breakdown both of successful attempts at musical analysis and of rated structure.

Three classes of analyst agreement were defined as follows. (1) Agreement among all three analysts with one equivalence, that of V_7 and VII, allowed. (In fact, Analyst 2 commented that, where VII was noted, he usually also meant to imply harmony at the dominant.) (2) Agreement between two analysts, allowing the same, but no other, equivalence. (3) Disagreement among all analysts, or the failure of at least one analyst to provide a solution. The mean listener ratings for agreement classes 1, 2, and 3 were 5.0, 4.5, and 2.8, respectively, for group H, and 4.6, 4.4, and 3.2 for group T. The probability of obtaining such an order for both groups by chance is less than .028. In gen-

[1] Royal Conservatory levels are designated by Roman numerals to distinguish the levels from school grades. There are 11 Conservatory levels. Both Grade X and Grade XI requirements must be met to obtain the diploma of Associate of the Royal Conservatory of Toronto (ARCT).

[2] An analysis based upon medians yielded essentially the same results, as did an analysis where responses were normalized in terms of the endpoints for each individual listener.

eral, key assignments reflected the same pattern as the musical analyses.

The agreement of analysts does not, however, present the whole story. First, even within cases of close agreement, the highest ratings tend to be given to sequences incorporating the I–V–I progression, or its allowed alternative. Second, a sequence with the I–V–I progression that also has the leading-note-to-tonic ending (Sequence 1) was given a higher rating than a sequence with the same progression but with a more complex contour (Sequence 5), or a sequence with the same progression but without the leading-note-to-tonic ending (Sequence 16). Third, within Class 3 disagreement, sequences that maintain a smooth contour, or maintain a tonic-to-tonic excursion, were generally rated higher than sequences that do not maintain such constraints. Notice, for example, the moderately high ratings given Sequences 15 and 17 (4.0 on the average), despite the violation of diatonicism.

The design of the present experiment does not allow the independent evaluation of harmonic or chord progression, contour, and excursion. Experiments 2 and 3 examined the effects of these variables on rating and recognition performance where the variables were entered into a cross-classified design.

Experiment 2

For the second experiment, 20 sequences were composed to represent the factorial combination of harmonic structure, contour, and excursion. Recognition of each sequence was tested in a two-alternative forced-choice paradigm. The recognition test involved the transposition of each sequence either to the dominant or to the tritone of the original tonic (cf. Cuddy et al., 1979), and involved a discrimination between the original sequence and the sequence with one note altered by one semit. Recognition of a sequence in transposition requires the abstraction of tone relations; to the extent that the factors reflect a perceptual representation of melody, performance in the transposition task should parallel the structure imposed on the stimuli.

Method

Sequences. The 20 experimental sequences are shown in Figure 3. All sequences are shown in musical notation with reference to the key of C. The rows of Figure 3 represent five levels or classifications of harmonic structure (S). Levels S_1, S_2, and S_3 obey the rule of diatonicism. The sequences for the highest level, S_1, contain the notes of the major scale, the I–V–I progression, and the leading-note-to-tonic ending. At the second level, S_2, the leading note is altered by one whole tone (two semits) and the I–V–I progression is weakened. For level S_3, the opening triad on I is altered along with the V–I progression. Levels S_4 and S_5 are nondiatonic in that the sequences involve one (S_4) or three (S_5) violations of the diatonic scale upon which the sequences of the first three levels are based.

In its simplest form for each level of harmonic structure, each sequence had only two directional changes in frequency, and each began and ended on the same note. These sequences are represented in the first column of Figure 3 under the headings "Simple Contour" and "Zero Excursion" (E_0) (the subscript following E refers to the number of semits between the first and last note). Each sequence was subjected to further transformations by rearranging the order of the notes to provide four directional changes in frequency, called "Complex Contour." Within each level of contour, there were two levels of excursion—an excursion from beginning and ending note of either zero (E_0) or five semits (E_5). The only repetition allowed in any sequence was that of the tonic, including the "tonic" from which levels S_4 and S_5 were derived. The repetition was in serial positions 1 and 7 for E_0, and in serial positions 2 and 7 for E_5.

Procedure. The recognition test used a two-alternative forced-choice procedure. On each trial a standard test sequence was presented, followed by a correct and an incorrect transposition of the test sequence. For the incorrect transposition, one tone, at either serial position 3 or serial position 5, was altered by plus or minus one semit. The direction of the alteration was randomly determined with the constraint that the altered tone was always a tone outside the key or tonality of the diatonic sequences.[3] (In two instances, to avoid repetition of the tone at serial position 2, it was necessary to test serial position 4 rather than serial position 3. The two instances were level S_4, Complex Contour, E_0; and level S_5, Complex Contour, E_5.) The transposed sequences were placed at a frequency location either six or seven semits above, or five or six semits below, the original test sequence. For all sequences, the frequency location seven semits above or five semits below corresponds to the musical location of the dominant of the final note of the test sequence, and the frequency location six semits above or below corresponds to the musical location of the tritone.

The entire recognition test consisted of 80 trials—

[3] Cuddy et al. (1979) indicated that for diatonic sequences, an error within the key is more difficult to detect than an error outside the key. Key of error was not included in the present experiment simply to avoid an inordinate number of factors in the design; it is, however, again being studied in current research.

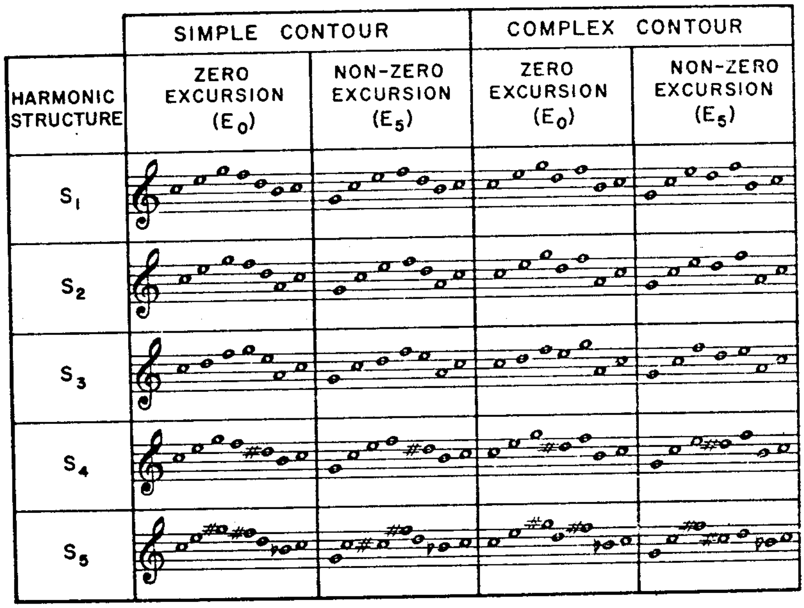

Figure 3. Sequences used in Experiments 2 and 3.

each of the 20 test sequences was tested at each of two serial positions and each of two transpositions. The direction of the transposition (above or below the test sequence) and the order of the transposition (correct versus incorrect) were determined at random. In addition, eight practice trials preceded the recognition test, and eight "dummy" trials were inserted at random locations within the recognition test. For the dummy trials, the altered note in the incorrect transposition was at serial positions 2 or 6. The purpose of dummy trials was to encourage listeners to attend to the entire test sequence. The construction of practice trials and dummy trials was similar, but not identical, to that of the main test trials.

The apparatus for the presentation of recognition trials was identical to that of the first experiment. The duration of each tone was .67 sec, the intersequence interval was 2 sec, and the intertrial interval was 4 sec. Each trial was assigned to a randomly selected frequency location within the range 174.6 Hz to 1480.0 Hz (musical notes F_3 to $F\#_6$). The cutoff frequency of the low-pass filter was set at 2.5 kHz.

Listeners and experimental design. Listeners were 60 undergraduate volunteers. The 30 listeners in the highly trained (H) group had, as in Experiment 1, an average performance level of Grade X Conservatory, and each regularly practiced several instruments. The 30 trained listeners had achieved, on the average, a higher performance level than that of the T group of the first experiment—by contrast, a mean performance level of Grade VIII Conservatory and regular practice on at least one instrument. This group was coded T_1 to indicate that the higher performance level represents a subset of our T classification. Each listener was tested under all possible combinations of the factors of structure, contour, excursion of the test sequence, key of transposition of the test sequence, and serial position of the note altered. They were asked to respond to each experimental trial by recording on a test sheet the temporal position (first or second) of the correct transposition of the test sequence.

Results

The proportion of correct responses averaged across serial position is presented in Table 1.[4] The table shows the proportion of correct responses for both groups of listeners as a function of structure, contour, excursion, and transposition. The top entry in each cell represents the data for Group H, and the bottom entry that for Group T_1.

Overall performance was slightly more accurate for the more highly trained group; the mean proportion correct was .82 for Group H and .74 for Group T_1, $F(1, 58) = 13.61$, $p < .001$. However, the same pattern of results was obtained for both groups. With one exception described below, no interactions with training level were statistically significant.

The extreme right-hand column of Table

[4] A preliminary analysis indicated no significant effects attributable to the serial position of the note altered in transposition.

Table 1
Mean Proportion Correct Recognition Responses for Highly Trained Listeners (Upper Entry) and Trained Listeners (Lower Entry)

Level of harmonic structure	Simple contour: Zero excursion (E_0) D	T	Nonzero excursion (E_5) D	T	Complex contour: Zero excursion (E_0) D	T	Nonzero excursion (E_5) D	T	M
S_1	.98 .97	1.00 .87	.98 .83	.92 .80	.93 .90	.83 .85	.97 1.00	.97 .98	.92
S_2	.93 .80	.88 .85	.93 .87	.85 .77	.90 .82	.90 .73	.95 .77	.90 .82	.85
S_3	.88 .77	.87 .68	.88 .82	.77 .77	.73 .73	.77 .63	.87 .73	.72 .78	.78
S_4	.88 .67	.82 .77	.87 .55	.73 .68	.75 .72	.70 .72	.80 .77	.80 .65	.74
S_5	.65 .62	.78 .77	.68 .57	.77 .53	.63 .55	.70 .68	.35 .38	.47 .43	.60
M	.82		.78		.76		.75		

Note. In the column headings, D signifies dominant and T signifies tritone.

1 shows a regular decrease in accuracy from .92 to .60 as cues to structure are removed or contraindicated, $F(4, 232) = 85.45$, $p < .001$. The main effects of contour and of excursion were significant; simple contour led to greater accuracy than did the complex contour, $F(1, 58) = 20.40$, $p < .001$, and zero excursion produced greater accuracy than did nonzero excursion, $F(1, 58) = 4.16$, $p < .05$. A slightly greater difference in accuracy attributable to contour was obtained for Group H than for Group T_1 (a difference of .06 versus .02), yielding a significant interaction between contour and training level, $F(1, 58) = 9.15$, $p < .01$.

The factors of structure, contour, and excursion also entered into triple interaction, $F(4, 232) = 5.95$, $p < .001$. With reference to Table 1, it may be seen that when both the cues provided by simple contour and zero excursion are removed, by far the greatest difficulty occurs at the lowest level of structure (S_5). The component of the triple interaction that reflects the relatively greater drop in performance at level S_5 than for the remaining four levels was significant, $F(1, 58) = 19.50$, $p < .001$, and accounts for 81% of the variability attributable to the triple interaction.

The effect of key distance in transposition is shown in Figure 4. Structural levels are represented along the abscissa, with musical location (as defined above) as the parameter. Key of transposition was itself not a significant factor ($p > .30$), but it entered into interaction with levels of structure. As shown in Figure 4 for both training groups, transposition to the dominant was superior to transposition to the tritone for the highest levels of structure, but the effect reversed at the lowest level of structure, $F(4, 232) = 6.23$, $p < .001$. The component of interaction reflecting the reversal at level S_5 is significant, $F(1, 58) = 23.48$, $p < .001$, and accounts for 94% of the variability attributable to the double interaction of transposition and structure. No interaction with training level occurred and the two panels of Figure 4 may be treated as essentially replicates of one another.

Experiment 3

The second experiment demonstrated that the proposed factors of harmonic progression, contour, and excursion influenced accuracy of recognition in a systematic fashion. As a further test of the reliability of the

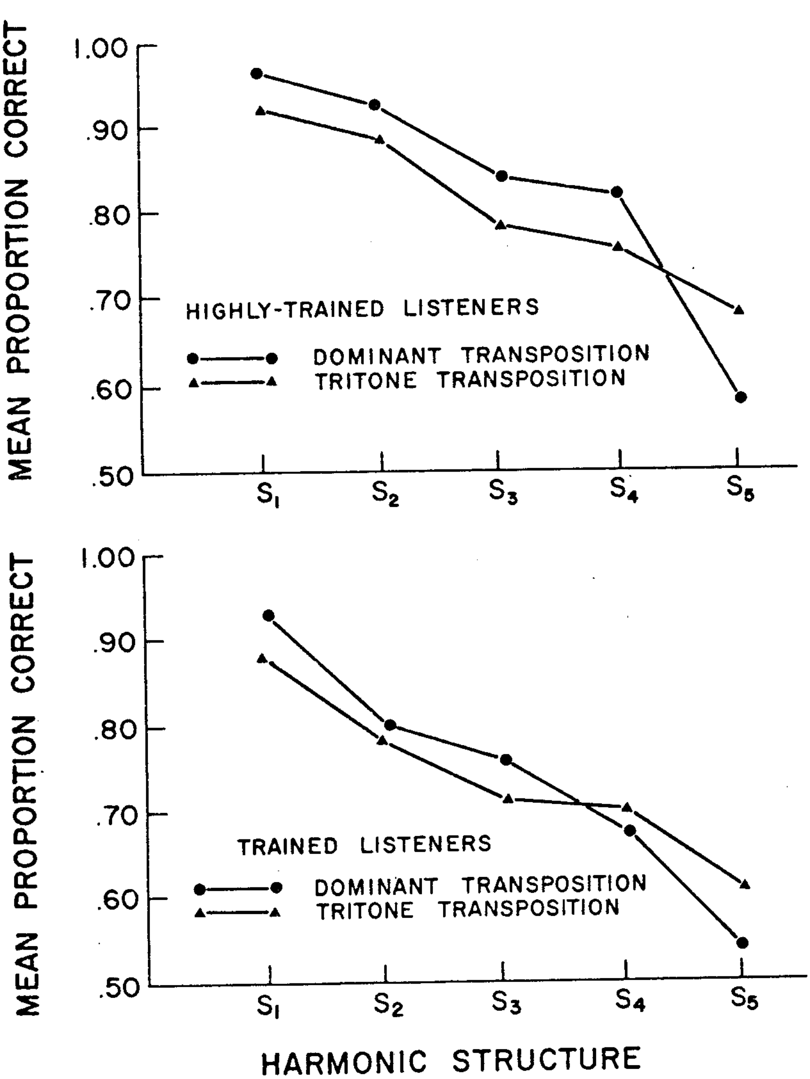

Figure 4. Mean proportion correct for five levels of decreasing harmonic structure. (Levels are plotted on a linear scale but they represent an ordinal scale.)

manipulations, the third experiment studied the sequences of the second experiment with a different response task—the rating task of the first experiment.

Method

The procedures were identical, with one main exception, to those described above for the rating method. Sequences were generated on-line under control of the computer. This modification provided 10 different randomizations of sequence orders and a reduction in the signal-to-noise ratio of about 20 dB. In addition, a .5-sec warning light immedately preceded each sequence presentation.

Three groups of 30 listeners each were tested. The criteria for inclusion in Groups H and T_1 were identical to those for Experiment 2. A third relatively untrained group (U) contained members with no more than 2 years of formal musical training.

Results

The mean rating for each sequence for each group of listeners is given in Table 2, in a format similar to that of Table 1. First, correlations between training levels were obtained across sequences for the rating data of the third experiment. Second, correlations between rating and recognition data were calculated for each possible pairwise comparison of training levels in the second and third experiments. The matrix of product-moment correlations is given in Tables 3 and 4.

Table 3 shows the correlations for the rating data of the different training levels; Table 4 shows the correlations for the correspondence between rating and recognition scores; for all correlations, $p < .01$, $df = 18$. The relation between rating and recognition scores for the 20 sequences closely approximated a linear trend (slope = .13, range, .10 to .14, average error of estimate = .06). In other words, a shift of one unit along the rating scale corresponded to just slightly more than a 10% shift in recognition accuracy.

Discussion

Psychology has periodically looked to musical intuition, knowledge, and theory to shed light upon problems of auditory perception. Helmholtz (1885/1954) in *Sensations of Tone* drew repeatedly upon musical examples, and the controversies between Wundt and Stumpf surrounding musical issues (Boring, 1942) foreshadowed later diverging traditions of empirical and introspectionist methods in psychology. Contemporary empirical psychologists (e.g., Krumhansl & Shepard, 1979) have turned to musical considerations to study the discrepancy between psychoacoustical and musical representations of pitch—why invariant logarithmic frequency relations are revealed in response to musical patterns (Attneave & Olson, 1971) but are not detected by classical psychophysical scaling paradigms (e.g., Stevens & Volkmann, 1940). The notion that we have explored in this paper is that the perception of invariant frequency relations depends upon the temporal structure of the tone sequence. To characterize temporal structure we have relied heavily, though not exclusively, upon musical description.

The results of the experiments support the earlier conclusion that "the rules of musical tonality apply to psychoacoustic tasks of tone recognition and that the rules are

Table 2
Mean Rating Scores for Highly Trained Listeners (Upper Entry), Trained Listeners (Middle Entry), and Untrained Listeners (Lower Entry)

Level of harmonic structure	Simple contour: Zero excursion (E_0)	Simple contour: Nonzero excursion (E_5)	Complex contour: Zero excursion (E_0)	Complex contour: Nonzero excursion (E_5)	*M*
S_1	5.30	4.83	4.77	5.13	
	5.43	5.27	4.83	5.10	5.01
	5.23	4.53	4.67	4.97	
S_2	4.27	4.30	3.93	4.07	
	3.93	3.73	3.10	3.50	3.78
	3.67	3.70	3.30	3.87	
S_3	3.40	3.93	3.30	3.53	
	3.40	3.00	2.97	3.37	3.33
	3.40	3.67	2.73	3.30	
S_4	3.83	3.73	3.53	3.37	
	3.27	2.93	2.57	2.83	3.21
	3.23	3.27	2.97	2.93	
S_5	3.03	2.70	2.63	2.50	
	2.57	2.50	2.07	1.93	2.53
	3.20	2.80	2.57	1.90	
M	3.81	3.67	3.33	3.49	

Table 3
Correlations Between Mean Ratings Obtained for the Three Training Levels of Experiment 3

Training level	Training level U	T_1	H
U	-	.95	.95
T_1	-	-	.94

Note. H = highly trained, T = trained, and U = untrained. T_1 is a subgroup of the trained listeners.

amenable to experimental manipulations" (Cuddy et al., 1979, p. 154). A greater depth of analysis of the rules of harmonic progression, plus the addition of contour and excursion as factors in a descriptive system, has led to increased differentiation among sequences in terms of performance characteristics. Further evidence of the application of a rule system is found in the key-distance effect in transposition (Experiment 2): Accuracy of recognition for transposed sequences is greater at the dominant than at the tritone. The scale at the dominant shares all but one member of the scalar alphabet of the original tonic; the tritone, only one member. The sharing of tones between the tone sets underlying two melodies appears to facilitate recognition of melodic similarities and differences. Facilitation through sharing may also apply at the lowest level of structure; at level S_5 there is greater overlap of tones between the original sequence and the transposition at plus or minus six semits (the tritone) than between the original and the alternative (dominant) transposition.

Listeners across a variety of levels of formal musical training were able to respond to the characteristics of the sequences. This conclusion is borne out by the high degree of correlation between training levels in the first and third experiments, and by the absence of qualitatively different patterns of interaction and experimental conditions in the second. The *overall* mean difference in recognition accuracy for the training levels of the second experiment should be treated with caution; it could reflect nonacoustical factors such as general interest in or attentiveness to the experiment. On the other hand, musical training may serve to provide additional (correlated) sources of information to apply to the task of encoding stimulus relations. While the untrained listener relies solely upon previously acquired implicit rules (Bamberger, 1978), the highly trained listener has access to a formal system that increases the probability of detecting pitch relations.

Table 4
Correlations Between Mean Recognition (Experiment 2) and Mean Rating (Experiment 3) Scores

Training level (recognition)	Training level (rating) U	T_1	H
T_1	.90	.85	.90
H	.85	.81	.85

Note. H = highly trained, T = trained, and U = untrained. T_1 is a subgroup of the trained listeners.

In the earlier report (Cuddy et al., 1979) we acknowledged a debt to Gestalt psychology for pointing out the psychological relevance of melodic transposition. However, the treatment of melody as an emergent whole accommodates neither the differential ease with which melodies are recognized and transposed nor questions concerning when and how we are able to extract logarithmic frequency relations. In order to handle these and other related problems in auditory pattern recognition, techniques for defining the meaningfulness of auditory pattern are sorely needed (Jones, 1978). Musical analysis, we suggest, provides an instance of a useful description of frequency patterns. Musical analysis provides descriptions of levels of different relations—for example, the sharing of the scalar alphabet between two melodies and the unfolding of harmonic progressions within a melody—while analysis of contour describes the order or configuration of the notes themselves. The data suggest that the complexity of pitch contour alone may influence the detection of harmonic relations, but until the role of contour is more fully explored with longer sequences and a greater variety of contours, the suggestions must remain tentative.

Our approach to auditory sequence recognition is compatible with models described by Jones (1978) as the "structure-detecting listener"—that is, models such as her own (Jones, 1974) and Restle's (1970, 1973) that propose the notion of an active search by the listener for invariant properties in a milieu of psychological and physical variance. Earlier, with respect to the learning and labeling of members of tone sets, Cuddy (1970) proposed that learning occurs most rapidly and effectively when three conditions are met: There must be (a) a system with the capacity to perceive structure, (b) a teaching method that initiates the adoption and development of structural rules, and (c) a tone set in which structure is emphasized through the physical presence of musical structure or frequency classification.

Now we add a further notion, that the listener detecting musical structure appears to operate in a fashion analogous to that which we described for the music analyst—that of working at several levels of musical application. Tone relations are encoded forward through the application of knowledge of likely progressions, but, in addition, the progressions are defined (within the limitations of memory space) from the coding of the final cadence and excursion. Thus, an entire sequence may be recognized or rated not merely in terms of the detectability of structure of the ongoing sequence but also in terms of whether it has a "strong" ending.[5]

An alternative account might propose that a pattern is the result of concatenated or clustered outputs of a series of frequency-ratio detectors. A neurological mechanism for interval recognition has been proposed (Deutsch, 1969). Following Deutsch's statement that a tune consists of successive intervals, a further step could view melody perception as the result of prior association or rehearsal of output clusters. Such a model would predict greater difficulty with unfamiliar (level S_5) as opposed to familiar (level S_1) clusters. The model would have difficulty, however, with key-distance effects such as those reported here where performance varies while the interval sequence remains invariant. Further, it would have difficulty explaining instances of tonal lures reported by Dowling (Bartlett & Dowling, 1980; Dowling, 1978, 1979) where listeners do *not* detect interval shifts in transposition if the scalar alphabet is maintained. Finally, we suggest that such a model would have difficulty incorporating recent evidence on the development of song learning (Gardner, Davidson, & McKernan, in press). Adult singers make interval and even contour errors but still obey the structural harmonic rules of progression.

However, both accounts have in common the notion that Gestalt principles must be superseded by an analysis of the detection of pattern. Our account focuses upon the role of structural rules and the expectations that the rules convey. Musical "form" is perceived when expectation is confirmed and a problem of analysis is solved. "Good form" is thus related to the ease and the certainty with which the solution is achieved.

---

[5] We have called the task in the second experiment a recognition task despite the fact that the paradigm is modeled after the two-alternative forced-choice discrimination procedure. The theoretical notion behind the choice of term is that the paradigm captures the listener's attempt to recognize scalar and serial rules. However, in order to bring the implications more in line with those of recognition tests in general, it would be useful in the future to apportion variance both to the nature of the standard sequence and to the perceived distance between the structure of the standard and comparison sequences.

Reference Note

1. Cuddy, L. L., Cohen, A. J., & Dewar, K. D. *Judgment of musical intervals: Relations between patterns of accuracy and musical training.* Presented at the 39th meeting of the Canadian Psychological Association, Ottawa, Canada, June 1978.

References

Attneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, *84*, 147–166.

Bamberger, J. Intuitive and formal musical knowing: Parables of cognitive dissonance. In S. S. Madeja (Ed.), *The arts, cognition and basic skills.* St. Louis, Mo.: CEMREL, 1978.

Bartlett, J. C., & Dowling, W. J. The recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 1980, *6*, 501–515.

Bernstein, L. *The infinite variety of music.* London: Weidenfeld & Nicolson, 1966.

Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century, 1942.

Cohen, A. J. *Perception of tone sequences from the Western European chromatic scale: Tonality, transposition and the pitch set.* (Doctoral dissertation, Queen's University, Ontario, Canada, 1975). *Dissertation Abstracts International,* 1977, *37,* 4179B. (National Library in Ottawa Order No. PC26238)

Cuddy, L. L. Training the absolute identification of pitch. *Perception & Psychophysics,* 1970, *8,* 265–269.

Cuddy, L. L. Absolute judgment of musically-related pure tones. *Canadian Journal of Psychology,* 1971, *25,* 42–55.

Cuddy, L. L., & Cohen, A. J. Recognition of transposed melodic sequences. *Quarterly Journal of Experimental Psychology,* 1976, 28, 255–270.

Cuddy, L. L., Cohen, A. J., & Miller, J. Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology,* 1979, *33,* 148–157.

Deutsch, D. Music recognition. *Psychological Review,* 1969, *76,* 300–307.

Deutsch, D. Effect of repetition of standard and comparison tones on recognition memory for pitch. *Journal of Experimental Psychology,* 1972, *93,* 156–162.

Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.), *Music and the brain.* Springfield, Ill.: Charles C Thomas, 1977.

Dewar, K. M., Cuddy, L. L., & Mewhort, D. J. K. Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, *3,* 60–67.

Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review,* 1978, *85,* 341–354.

Dowling, W. J. The cognitive psychology of music. *Humanities Association Review,* 1979, *30,* 58–68.

Dowling, W. J., & Fujitani, D. S. Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America,* 1971, *49,* 524–531.

Francès, R. *La perception de la musique.* Paris: Librairie Philosophique J. Vrin, 1972.

Gardner, H., Davidson, L., & McKernan, P. The acquisition of song. *Proceedings of National Symposium on the Application of Psychology to the Teaching and Learning of Music.* Reston, Va.: Music Educators National Conference, in press.

Garner, W. R. Good patterns have few alternatives. *American Scientist,* 1970, *58,* 34–42.

Garner, W. R. *The processing of information and structure.* Toronto, Canada: Wiley, 1974.

Garner, W. R., & Clement, D. E. Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior,* 1963, *2,* 446–452.

George, G. *Tonality and musical structure.* London: Faber & Faber, 1970.

Green, D. M. *Form in tonal music.* Toronto, Canada: Holt, Rinehart & Winston, 1965.

Helmholtz, H. L. F. *On the sensations of tone as a physiological basis for the theory of music.* A. Ellis (Ed., and trans.). (4th ed.). New York: Dover, 1954. (Originally published, 1885.)

Jones, M. R. Cognitive representations of serial patterns. In B. H. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition.* Hillsdale, N.J.: Erlbaum, 1974.

Jones, M. R. Auditory patterns: Studies in the perception of structure. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception VIII: Perceptual coding.* New York: Academic Press, 1978.

Krumhansl, C. L., & Shepard, R. N. Quantification of the hierarchy of tonal functions. *Journal of Experimental Psychology: Human Perception and Performance,* 1979, *5,* 579–594.

Laske, O. *Music, memory and thought: Explorations in Cognitive Musicology.* Ann Arbor, Mich.: University Microfilms International, 1977.

Meyer, L. B. *Emotion and meaning in music.* Chicago: University of Chicago Press, 1956.

Restle, F. Theory of serial pattern learning: Structural trees. *Psychological Review,* 1970, *77,* 481–495.

Restle, F. Coding of nonsense vs. the detection of patterns. *Memory & Cognition,* 1973, *1,* 499–502.

Stevens, S. S., & Volkmann, J. The relation of pitch to frequency: A revised scale. *American Journal of Psychology,* 1940, *53,* 329–353.

Vitz, P. C., & Todd, T. C. A coded element model of the perceptual processing of sequential stimuli. *Psychological Review,* 1969, *76,* 433–449.

Received January 28, 1980 ■