Journal of Experimental Psychology:
Human Perception and Performance
1979, Vol. 5, No. 4, 579–594

# Quantification of the Hierarchy of Tonal Functions Within a Diatonic Context

Carol L. Krumhansl and Roger N. Shepard
Stanford University

Listeners rated test tones falling in the octave range from middle to high C according to how well each completed a diatonic C major scale played in an adjacent octave just before the final test tone. Ratings were well explained in terms of three factors. The factors were distance in pitch height from the context tones, octave equivalence, and the following hierarchy of tonal functions: tonic tone, other tones of the major triad chord, other tones of the diatonic scale, and the nondiatonic tones. In these ratings, pitch height was more prominent for less musical listeners or with less musical (sinusoidal) tones, whereas octave equivalence and the tonal hierarchy prevailed for musical listeners, especially with harmonically richer tones. Ratings for quarter tones interpolated halfway between the halftone steps of the standard chromatic scale were approximately the averages of ratings for adjacent chromatic tones, suggesting failure to discriminate tones at this fine level of division.

The study of perceived pitch and of the perceived relations between tones differing in pitch has generally been approached from one of two quite different traditions: the psychoacoustic and the musical. The psychoacoustic approach has typically focused on simple, physically specifiable properties of tones isolated from any musical context—properties of frequency, separation in log frequency, or simplicity of integer ratios of frequencies. The results of such studies have provided some precise information about how the ear responds to isolated tones or tones in random sequences. We believe that they have been less informative with regard to how the listener perceives tones in organized musical sequences. Music theory suggests that the perception of such sequences may rely on the listener's sensitivity to different and structurally richer principles associated with tonal and diatonic organization. Such principles are useful in explaining the cognitive phenomena of reference point, motion, tension, and resolution that underlie the dynamic force of virtually all tonal music. They have, however, been subjected to relatively little systematic laboratory investigation or quantitative formulation.

## Background

Until recently, psychoacoustic studies were dominated by the assumption that pitch is a simple psychological counterpart of the single physical dimension of frequency, and so varies along a single psychological dimension of what has come to be called pitch "height." The problem of pitch was in this way reduced to the psychophysical one of determining the functional form of the presumably monotonic relationship between physical frequency and psychological pitch (Stevens & Volkmann, 1940; Stevens, Volkmann, & Newman, 1937).

There is, however, evidence inconsistent with such a simple, unidimensional, and monotonic representation of pitch, namely, evidence that tones an octave apart (that is, tones standing in a frequency ratio of 2:1) are perceived as having something more in common than tones somewhat less than an octave apart (Allen, 1967; Blackwell & Schlosberg, 1943; Boring, 1942, pp. 376, 380; Humphreys, 1939; Licklider, 1951, pp. 1003–1004; Ruckmick, 1929). With increasing refinement, such psychoacoustic observations have led to the view that pitch is a more complex attribute having at least two components: pitch height (Balzano, 1977; Deutsch, 1972a; Dewar, 1974; Levelt, Van de Geer, & Plomp, 1966) and octave equivalence or tone "chroma" (Bachem, 1954; Balzano, 1977; Deutsch, 1973; Shepard, 1964). Certainly the precision (Ward, 1954) and cross-cultural consistency (Dowling, 1973a) with which listeners adjust a variable tone so that it is an octave away from a fixed tone appear difficult to explain in terms of any purely one-dimensional psychophysical scale of pitch. In addition, some suggestive indications have been reported for augmented similarities of tones differing by simple ratios of frequencies other than the simplest two-to-one ratio that presumably underlies octave equivalence (Balzano, 1977; Levelt et al., 1966).

There is evidence, moreover, that the effects of even these physically specified factors depend on whether the tones are interpreted by the listener as musical or nonmusical. On the basis of data from a nonmusical task of fractionation, Stevens et al. (1937) originally constructed, and Stevens and Volkmann (1940) refined, the *mel* scale of pitch, which is a markedly nonlinear, though monotonic, function of both frequency and log frequency. In a more musical task in which subjects produced transpositions of a familiar sequence of pitches, Attneave and Olson (1971) found, by contrast, that the transpositions maintained the interval relations of the original melody with respect to log frequency, rather than with respect to the psychoacoustic mel scale. The extent to which the presented tones are interpreted as musical also appears to influence the degree to which tones separated by an octave are perceived as similar. For example, Allen (1967) found that octave equivalence was strong in music students but largely absent in nonmusical listeners. Similarly, Thurlow and Erchul (1977) found that only some subjects were able to recognize musical intervals (defined, in effect, in terms of tone chroma) when the higher tone of the interval appeared randomly within any of several different octaves.

Recognizing the powerful role that musical and cognitive factors may play in the perception of tones in a musical context, a number of experimental psychologists interested in the perception of music have recently collected data on tones contained in more or less well-structured musical sequences. Such studies have emphasized the importance of larger structural, cognitive, and Gestalt-like

The work reported here was supported in part by National Science Foundation Grant BNS-75-02806 to the second author and was carried out during the first author's National Institute of Mental Health Traineeship (MH-10478-09) at Stanford University. We are greatly indebted to several research assistants and colleagues: Shelley Hurwitz, who ran the subjects in Experiment 1 and carried out much of the data analysis for both experiments; Zehra Peynircioglu, who ran the subjects in Experiment 2 (and served as our one subject with absolute pitch); Kip Sheeline, who prepared the computer-generated tapes for Experiment 2 using the facilities generously provided by the Center for Computer Research on Music and Acoustics at the Stanford University Artificial Intelligence Laboratory; and John Grey and Gerald Balzano, who provided additional help in the form of equipment and suggestions.

Experiment 1 was first reported at the symposium "Cognitive Structure of Musical Pitch" at the 1978 meeting of the Western Psychological Association in San Francisco (Krumhansl, Note 1).

Requests for reprints should be sent to either Carol L. Krumhansl, who is now at Dept. of Psychology and Social Relations, Harvard University, Cambridge, Massachusetts 02138, or to Roger N. Shepard, Department of Psychology, Stanford University, Stanford, California 94305.

properties such as melodic contour, proximity grouping, and, especially, musical scale and tonality.

Dowling and his co-workers, particularly, have established the importance of contour. Dowling and Fujitani (1971), for example, found that listeners easily distinguished transpositions that preserved melodic contour from those that did not; but, for transpositions that preserved qualitative contour, they found that listeners had difficulty in discriminating transpositions that also preserved the quantitative sizes of the pitch intervals from those that did not. Moreover, although Deutsch (1972b) found that subjects had difficulty identifying familiar melodies when the tones were dispersed in different octaves, Dowling and Hollombe (1977) subsequently showed that this disruptive effect of octave dispersal is reduced if the dispersed version is constructed so as to preserve the qualitative contour of the original melody.

The residual difficulty of octave-dispersed melodies, even after contour has been preserved, undoubtedly stems in large part from the resulting violation of the Gestalt principle, well-known in vision, of perceptual grouping and coherence on the basis of proximity. More direct evidence for the operation of such a proximity principle in the domain of auditory pitch comes from studies of auditory "stream segregation" or "melodic fission" (e.g., see Bregman & Campbell, 1971; Dowling, 1973b; Shepard, in press). Still further evidence for the importance of proximity comes from experiments by Deutsch (1978) in which listeners attempted to recognize the repetition of a given tone following the experimental interpolation of a sequence of other tones. Deutsch reported that accuracy of recognition varied inversely with the average size of the intervals in the interpolated sequence.

Finally, and most importantly from our standpoint, there are indications that tonality and the organized pattern of pitch intervals contained within a musical scale influence the perception of musical sequences. Tonality is described by music theorists (e.g., Piston, 1941; Ratner, 1962) as the single most critical factor in musical organization. Briefly, tonality is synonymous with key and designates a set of tones as the musical scale to be used in a particular context. The most common scale used in traditional Western music is the major diatonic scale, which consists of seven of the 12 musical tones contained within each octave (namely, the seven tones called *do, re, mi, fa, sol, la, ti,* generally completed by the eighth tone, *do'*, an octave above the first). These tones conform to a fixed pattern of intervals that repeats in every octave. In the key of C, they are represented by the white keys of the piano keyboard. The remaining five tones in each octave (in C major, the black keys of the piano) are nonscale or nondiatonic tones.

Recent archeological evidence suggests that the use of the diatonic scale goes back well over 3,000 yr. (Kilmer, Crocker, & Brown, 1976). And, although scales do differ considerably from one culture to another, they generally share certain basic structural features with the diatonic scale (Dowling, in press; Dowling, Note 2), which seem to reflect universal structural constraints that may have both a cognitive (Dowling, Note 2) and a group-theoretic (Balzano, Note 3) basis.

In addition to defining the musical scale, tonality designates one single tone in the scale, called the tonic, as the most structurally stable tone in the system. This tone functions as a kind of reference point to which each of the other tones has a well-defined relationship, called its tonal function. In accordance with a hierarchy described in music theory (Meyer, 1956, pp. 214–215), other tones differ with respect to how closely related they are to the tonic and, consequently, how structurally stable they are within the system. For example, in the key of C major the tonic, C, is the most stable tone, but other tones, particularly the third and fifth tones of the scale, E and G (which together with C constitute the so-called major triad chord), are also relatively stable, whereas other tones of the scale are less stable, "tending" (as music theorists say) toward more stable tones—particularly the tonic itself. The tones outside the scale, or nondiatonic tones, are still less stable and tend even more strongly toward the more stable tones within the system.

Empirical evidence for the importance of tonality and the diatonic structure comes from a number of recent studies. Cohen (Note 4) has shown that musical subjects, after hearing a short excerpt from a musical composition, are able to generate the associated diatonic scale with some accuracy. Other investigators (Dewar, 1974; Dewar, Cuddy, & Mewhort, 1977) have found that memory is better for tones contained in sequences conforming to diatonic structure than for tones in sequences not conforming to this structure. Dewar (1974) found that changes in musical sequences that took such sequences outside the diatonic scale were more easily detected than changes that left them within the diatonic scale, and Dowling's (1978) subjects found it difficult to distinguish exact transpositions (a constant shift in log frequencies, but necessarily departing somewhat from the diatonic scale) from contour-preserving shifts along the diatonic scale (changing the intervals somewhat, but keeping within the scale). Cohen (1975) and Attneave and Olson (1971) have found, further, that subjects are better able to recognize and to produce transpositions of tonal or familiar sequences than other less well-structured sequences. Finally, in a series of experiments undertaken subsequently to the experiments to be reported here, Krumhansl (1979; Krumhansl, Note 5) demonstrated that in an explicitly tonal context, the pattern of perceived similarities between tones was highly specific to the tonality of that context and recognition memory was better for diatonic than for nondiatonic tones.

These various recent findings strongly reinforce the notion that the representation of musical pitch must be more highly structured than had been supposed on the basis of earlier psychoacoustic studies in which experimenters assiduously avoided any semblance of a musical context. Apparently, in addition to the relations of the frequencies of the physically presented test tones themselves, which underlie pitch height and octave equivalence, structurally richer factors induced by a particular tonal context will have to be taken into account in attempting a complete description of the interval representation of musical pitch.

## Our Approach

The experiments that we now report were initially motivated by what, despite the recent advances just reviewed, we still perceived as a noticeable gap between the richness and power of the total structure implied by music theory, on the one hand, and detailed experimental demonstrations and quantifications of that structure, on the other. Studies of the perception of musical intervals had clearly demonstrated systematic effects of simple separation in log frequency and of octave equivalence, but the hierarchy of tonal functions or stabilities that are to be expected on the basis of the music-theoretic notions of tonality and diatonicism had curiously failed to assert itself in the laboratory.

Even the reported effects of the simplicity of the ratio of physical frequencies have not been as robust as expected. Thus, in the study by Levelt et al. (1966), such effects appeared only in the case of already harmonically rich tones as opposed to pure sinusoidal tones; and in the study by Balzano (1977), such effects emerged only for harmonic intervals (simultaneous tones), not for melodic intervals (successive tones). What effects had been reported might therefore be attributed to relatively sensory interactions between related frequencies or harmonics in the ear more than to cognitive factors having to do with the tonal functions indicated by music theory. And, even so, some of the results suggestive of structure beyond the simple acoustic one of separation in log frequency—such as the virtually identical multidimensional scaling solutions independently obtained by Levelt et al. (1966) and Wickens (reported in Shepard, 1974)—may have been partly artifactual consequences of attempts to extract more dimensions than the nonmetric analysis would support (see Shepard, 1974, pp. 387–388).

Our experiments were based on the conjecture that previous failures to bring out sharply the tonal functions that are expected within the diatonic system might stem from a failure to meet three conditions simultaneously: (a) The tones must be presented within a musical context that strongly and unambiguously establishes the tonal framework with respect to which all test tones are to be interpreted. (b) Detailed quantitative information must be obtained for each of the alternative tones of interest relative to this tonal framework. (c) Data must be analyzed separately for individual listeners who may differ widely in musical experience, training, or aptitude, and so vary markedly in their responsiveness to the experimentally introduced musical context.

In both of the following experiments, we established the tonal context by simply playing the seven tones of an ascending or descending major scale (omitting the eighth tone, an octave from the first, that normally completes the sequence). We then played a single tone from somewhere in the octave range extending from the omitted tonic tone that would normally have completed the context series to the tone just one octave beyond that note. We asked listeners, for each such final tone, to rate how well that tone fit in with or completed the preceding context sequence. We hoped that the judgments obtained for different tones within the test octave would provide a direct, quantitative indication of the stability and tonal function of each tone relative to the tonal framework induced by the preceding context.

## Experiment 1

In Experiment 1, subjects judged how well each of the 13 (equally tempered chromatic) tones within the octave range between middle C and the tone an octave above (C′) completed ascending and descending C major scales. The two alternative scale sequences and the 13 final tones used in Experiment 1 are shown in Figure 1. Since the ascending C major scale began on the C an octave below middle C and the descending C major scale began on the C two octaves above middle C, the octave range of final tones was bound by the two alternative context sequences.

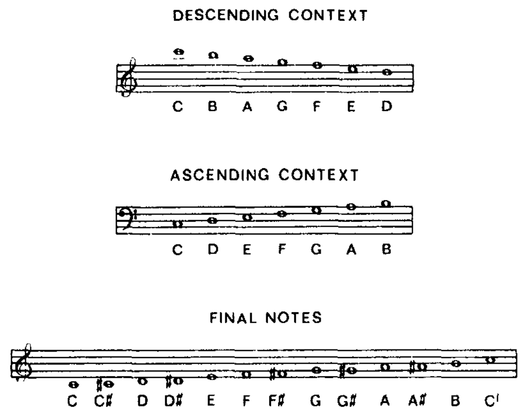


*Figure 1.* The descending and ascending context sequences of seven tones of the C major scale and the 13 chromatic alternative test tones from the intervening octave that might immediately follow the context sequence.

The choice of the key of C (and consequently of the octave range from C to C′) instead of any of the 11 other possible major keys and ranges should not, of course, restrict the generality of our results. For in accordance with music theory, previous psychological results (e.g., Attneave & Olson, 1971; Dowling, 1978; Dowling & Fujitani, 1971) confirm that the structure of the perceived relations between the various scale degrees is invariant under transposition along a log frequency scale. (We disregard here the exceptions that might conceivably arise in the case of those rare individuals possessing true absolute pitch.)

Earlier studies suggest that a number of factors may influence such judgments of scale completion. First, if pitch height is an important dimension, then tones close in frequency to the context tones should be preferred to tones farther from the context tones. Second, if the tones separated by an octave are equivalent or closely related, then the Cs at both the high and the low ends of the octave range of test tones should be judged as approximately equally good completions. Third, if tones of the diatonic scale are more closely related to the tonal center implied by the scale context, then these diatonic tones should be preferred to the nondiatonic tones. Fourth, if the more detailed music-theoretic notions concerning a hierarchy of tonal functions within the diatonic system have psychological correlates, then the most preferred tones, following the tonic, should be the other tones of the major triad chord (the third and fifth degrees of the scale), followed by the other diatonic and finally the nondiatonic tones.

## *Method*

### *Subjects*

Twenty-four Stanford University undergraduates participated. Twenty-three of the participants received credit toward an introductory course requirement. The 24th, a research assistant who was as yet naive about the purposes of the experiment, was later added because of her extensive musical training and her possession of absolute pitch. Each listener participated in a single 1-hr. experimental session, and filled in a short questionnaire describing his or her musical training and experience.

### *Apparatus*

We prepared the stimulus tapes by recording at 7.5 inches per second (ips) directly from a Farfisa (Chicago Musical Instrument Co.) electronic organ onto tape using a Revox A77 Stereo Taperecorder. We used the flute stop on the organ as the best available approximation to a pure sinusoidal tone generator. During the experimental sessions, we played the prerecorded tapes at a comfortable loundness level via the Revox Taperecorder, a Dynaco SCA-80Q amplifier, and two Bang and Olafsen Beovox S45 speakers.

### *Stimulus Sequences*

Each trial consisted of the presentation of a seven-tone sequence of either an ascending or a descending C major scale followed immediately by a final tone, which was any one of the 13 equally tempered notes within the octave range bounded by the two possible scale contexts. The duration of each tone was approximately .75 sec, and the intertrial interval during which listeners made their responses was approximately 6 sec. Each of the 13 final notes appeared with each of the context types (ascending and descending scales) once in each of eight blocks of trials. Trials were randomly ordered within blocks, and the blocks were presented to the subjects in different random orders.

### *Procedure*

Listeners were instructed to judge on each trial how well the final tone completed the presented sequence. We asked them to express their judgments by means of numerical ratings of goodness from 1 to 7, where 7 was designated "very good," 4 was designated "moderate," and 1 was designated "very bad." We encouraged them to use the full range of these ratings.

## *Results*

### *Individual Differences*

Each listener's response profile consisted of the average ratings for each of the 13 notes for the two types of context. Of the original 23 listeners, 22 of these response profiles showed one of the three general patterns in Figures 3, 4, and 5. Seeking more objective support for our informal division of the listeners into these three groups, we computed correlations between the data for each pair of listeners and then analyzed the resulting matrix of correlations by means of the hierarchical clustering method HICLUS (Johnson, 1967). The HICLUS solution using the compactness option consisted of three groups of listeners and a single unclassified listener, as shown in Figure 2. These more objective results, as well as those (not shown) obtained by the connectedness option, were in essentially complete agreement with our original subjective classification.

All eight listeners in Group 1 played an instrument or had vocal instruction (mean years of instruction, 7.4; mean years performing, 5.6). Of the eight listeners in Group 2, all but one played an instrument or sang (mean years of instruction, 5.5; mean years performing, 3.3). Of the six listeners in Group 3, only two played an instrument and one sang (mean years of instruction, .7; mean years performing, .0). Owing to rather large intragroup variation, only Groups 1 and 3 differed significantly from each other in terms of years of instruction and performing experience. The $t$ values comparing Groups 1 and 3 were $t(12) = 3.00$, $p < .05$, and $t(12) = 4.62$, $p < .01$, for years of instruction and years performing, respectively. The comparisons involving Groups 2 and 3 approached significance, $t(12) = 2.03$ and $t(12) = 2.07$, $.05 < p < .10$ for both, for years of instruction and years performing, respectively. The single unclassified listener had 1 yr. of musical instruction and 11 yr.

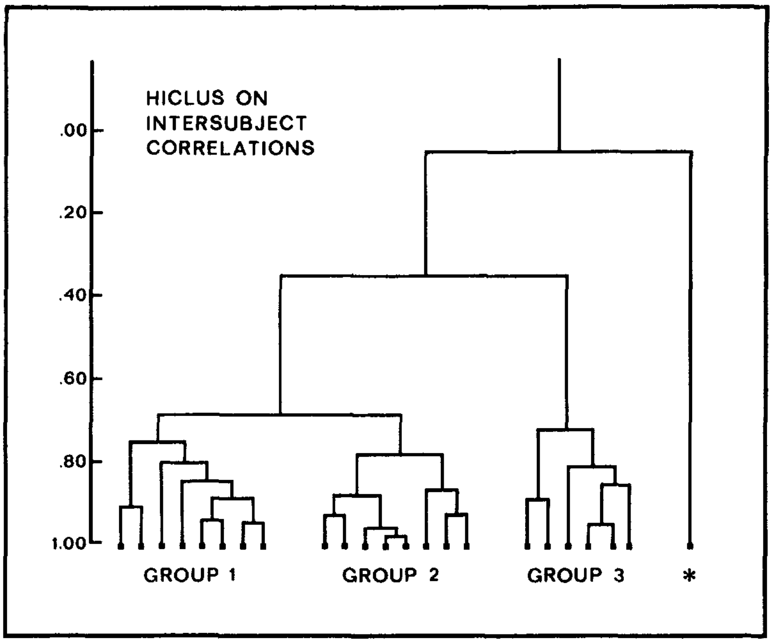


*Figure 2.* Hierarchical clustering (compactness or complete link method) of the first 23 listeners in Experiment 1, based on the matrix of correlations between rating profiles for all pairs of listeners.

of performing experience. None of the original 23 listeners responded affirmatively to a question asking whether they have absolute pitch. The subsequently added 24th listener, however, reported having absolute pitch in addition to exceptionally extensive musical training—17 yr. of formal instruction and 10 yr. of performing experience.

### *Judgments of Scale Completion*

*Group 1.* The average scale completion judgments for the 13 final tones for Group 1 listeners are plotted for both ascending and descending contexts in Figure 3. These eight listeners manifested a clear preference for the low and high Cs over all other notes $t(7) = 21.59$, $p < .001$. In addition, there was a preference for scale notes other than the tonic (D, E, F, G, A, B) over the nonscale notes (C#, D#, F#, G#, A#), $t(7) = 12.59$, $p < .001$. Among the scale tones, D and B, which are the scale tones adjacent to the tonic C, were judged to be relatively good scale completions, as was the tone G, the fifth degree of the scale. The profiles for ascending and descending scale contexts were similar, the ratings for the notes showing very little dependence on whether the scale context was ascending or descending. In particular, both high and low Cs were judged to be equally good completions of ascending and descending scales.

*Group 2.* The average responses of Group 2 listeners, presented in Figure 4, showed a strong preference for the two Cs over all other 11 notes in the octave range, $t(7) = 13.66$, $p < .001$, giving a deep U-shaped pattern to the profile. In addition, there was a small but significant difference between the scale notes other than the tonic and the nonscale notes, $t(7) = 4.63$, $p < .01$. Furthermore, the low tones were judged to be somewhat better completions for the ascending context, and the high tones were judged to be better completions for the descending context. In particular, every subject showed a preference for the low C on ascending trials and the high C on descending trials.

*Group 3.* In the ratings of Group 3 listeners, shown in Figure 5, again there was a preference for the two Cs over other tones

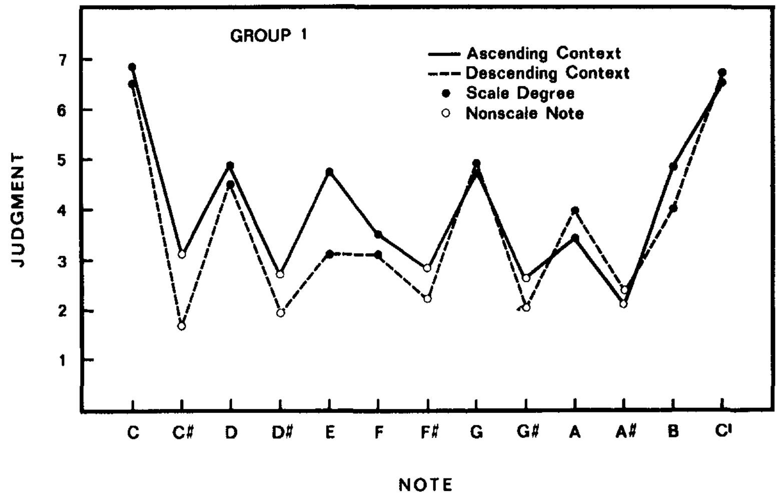


*Figure 3.* Mean rating profiles for Group 1 listeners in Experiment 1. (The mean rated goodness of each test tone as a completion of the preceding context sequence is plotted separately for ascending and descending contexts.)

in the octave range, $t(5) = 6.45$, $p < .01$, although the magnitude of the effect was not as large as in Group 1 and Group 2. Also, there was a small but significant preference for the other scale tones over nonscale tones, $t(5) = 2.78$, $p < .05$. However, distance from the tones of the preceding scale context accounted for most of the variance in ratings of these listeners. Low tones were judged to be better completions of ascending scales, and high tones were judged to be better completions of descending scales. In particular, the low C was judged to be a much better completion of ascending scales than was the high C, and the opposite held for descending contexts. While the Group 2 subjects showed this distance effect primarily at the two Cs, this effect was found throughout the octave range for Group 3. Indeed one of the listeners in Group 3 exhibited a pure distance effect with no indication of octave equivalence. For this listener, each of the two crossing curves like those plotted in Figure 5 fell off monotonically

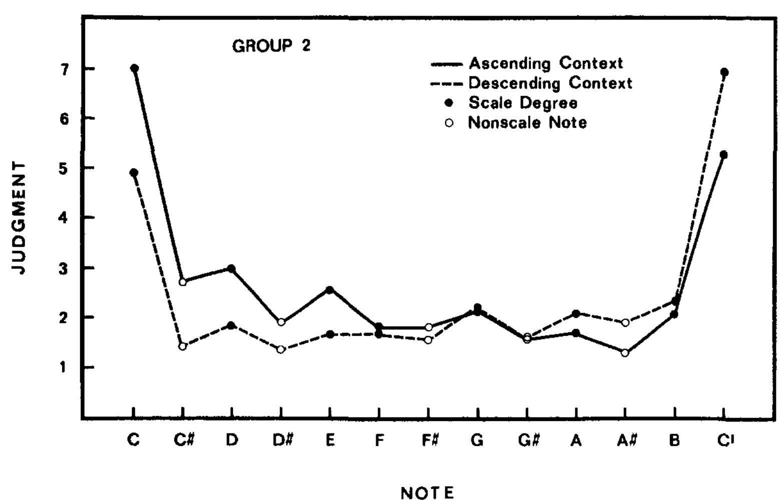


*Figure 4.* Mean rating profiles for Group 2 listeners in Experiment 1.

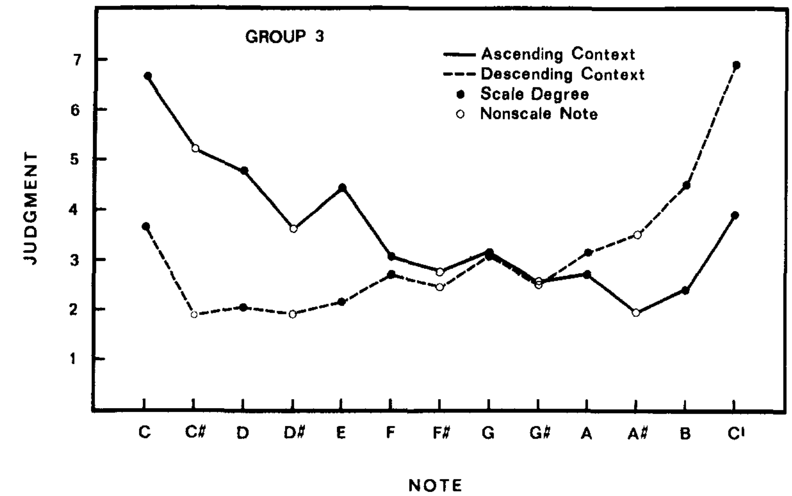


*Figure 5.* Mean rating profiles for Group 3 listeners in Experiment 1.

from the C nearest to the preceding context but without any upswing at the C at the other end of the test range.

*Unclassified subject.* The single listener in this original set of 23 who could not be classified as falling into any one of the three groups showed a pattern similar to Group 1 listeners for descending contexts, but a virtually flat response profile for ascending contexts.

*Subject 24.* Figure 6 shows the response profile for the subsequently added listener who, in addition to exceptionally extensive musical training, reported having absolute pitch. As with Group 1, there was little difference between ascending and descending contexts, and again there was a general preference for scale over nonscale tones. Within the set of scale tones there was, in addition to the high ratings for the two Cs, a very marked preference for the other members of the major triad, namely, the third and fifth degrees of the scale (E and G), over the other scale tones.

### *Discussion*

The results of Experiment 1 demonstrate the influence of a number of factors on the

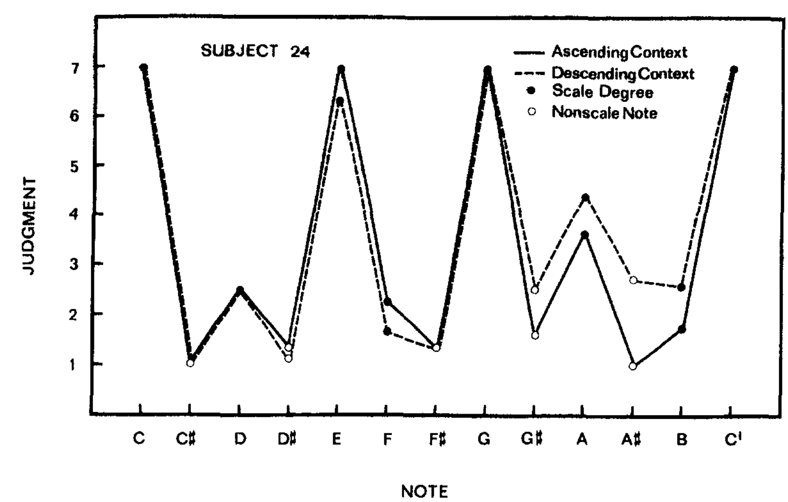


*Figure 6.* Mean rating profiles for the listener in Experiment 1 who was subsequently added because of her exceptionally extensive musical training and posssion of absolute pitch.

perceived suitability of alternative tones as completions of a scale sequence. Further, the relative contributions of these different factors evidently vary markedly with the musical backgrounds of the listeners. The previously well-established factor of pitch height emerged once again, with tones close in frequency to the context tones preferred over tones farther from the context. However, the influence of pitch height was strong only for the least musical listeners in Group 3. For these listeners, the context apparently served primarily to define a location on the continuum of pitch height, so that there were nearly opposite effects when the context tones were presented above and below the range of test tones. In marked contrast, the musical Group 1 listeners gave strikingly similar ratings following ascending and descending scale sequences. Thus, the context evidently defined for these listeners a tonal structure that, under octave equivalence, applied to tones in any octave according to their individual functions within this invariant structure regardless of their location in pitch height. In particular, the high and low Cs were judged by Group 1 to be equally good completions of both ascending and descending scales. This octave equivalence was weaker for less musical subjects, with the low C preferred in ascending scale contexts and and the high C preferred in descending scale contexts, in agreement with earlier findings (Allen, 1967).

The orderly relationship found between the results for individual listeners and their reported levels of musical interest and training is consistent with similar relationships reported in earlier studies, including those by Allen (1967), Attneave and Olson (1971), Cohen (1975), Cuddy (1970), Dewar et al. (1977), and Shepard (1964). Apparently, the kinds of structural relations that are extracted from tonal sequences vary with the extent of the listener's musical sophistication, whether innate or acquired.

Despite these individual differences, the rating profiles of all three groups of listeners revealed a general preference for scale over nonscale tones, although this preference for diatonic tones was strongest for the most musical listeners. Within the set of diatonic tones, moreover, the tonic was the most strongly preferred by all listeners. Particularly for musical listeners, the fifth and, to some extent, the third degrees of the scale were also judged to be relatively good scale completions, as were the second and seventh degrees. However, proximity to the tonic in pitch height probably contributed to the preferences for the second and seventh in many subjects; certainly the rated preferences for these two tones were lower in the listener with the most extensive musical background (Subject 24; see Figure 6).

Our quantitative results are consonant, then, with expectations based on music theory that the full musical interpretation of a set of tones entails the internal assignment of those tones to a hierarchy of tonal functions. Indeed, in the case of musical subjects, this hierarchy, together with virtually complete octave equivalence, essentially replaces the effect of distance or separation in pitch height that has been the dominant result of previous psychoacoustic studies in the absence of a tonal context. Further evidence for the reality of such an underlying hierarchy comes from a subsequent study (Krumhansl, 1979; Krumhansl, Note 5) in which musically trained listeners judged the similarity between pairs of tones in an established tonal context. Multidimensional scaling of the results of that study yielded a hierarchy in essential agreement with the hierarchy found here. A tightly organized cluster corresponding to the tones of the major triad chord was surrounded by a more widely dispersed set corresponding to the remaining diatonic tones, which set in turn was surrounded by a much more widely dispersed set corresponding to the nondiatonic tones.

## Experiment 2

Two limitations of Experiment 1 led us to undertake a second experiment. First, since in Experiment 1 we had divided the octave test range only into the standard halftone (or semitone) steps of the 13-tone chromatic scale, we precluded the possibility of discovering any still lower levels of the hierarchy of tonal functions. The possibility occurred to us (and to Balzano, Note 6) that, just as the tones directly implied by a particular diatonic context dominate all other tones from the chromatic scale of halftone steps, those other tones are themselves *indirectly* implied by such a context under transposition (modulation of key); and so they might in turn dominate other tones not even indirectly implied under transposition such as, for example, tones interpolated by quarter-tone steps, halfway between the standard tones of the chromatic scale.

Second, the flute-stop tones presented in Experiment 1 were only approximations to pure sinusoidal tones. So, even though the tones were presented sequentially rather than simultaneously, the listeners' judgments of goodness of sequence completion may have been partly influenced by some perception of overlap between the harmonics present in the test tone and the frequencies of the preceding context tones. As we have already noted, there is evidence that judgments of harmonic relations reflect a richer structure when the tones themselves are harmonically richer (e.g., Levelt et al., 1966).

Accordingly, although Experiment 2 was carried out in essentially the same way as Experiment 1, we modified the set of tones in two ways: To the 13 chromatic tones separated by halftone steps within the octave test range, we added the 12 quarter tones halfway between these to obtain a more finely graded series of 25 test tones. And, instead of generating the tones by means of an organ flute stop, we used a computer to generate all tones as pure sinusoids, free of upper harmonics.

## *Method*

### *Subjects*

We recruited seven Stanford undergraduates who served as paid listeners by posting notices in the music department. The eighth listener was the same highly musically trained research assistant who participated in Experiment 1. As before, all listeners completed a questionnaire describing their musical training and experience.

### *Apparatus*

We recorded the stimulus tapes at 7.5 ips on a Revox A77 Taperecorder from a digital signal processor controlled by the PDP-10 computer at Stanford's Center for Computer Research on Music and Acoustics. During the experimental sessions, we played these tapes by a Revox A77 Taperecorder through earphones at a comfortable loudness level.

### *Stimulus Sequences and Procedure*

The trials were similar in construction to those of Experiment 1. The context, consisting of an ascending or descending diatonic scale, was followed by a final tone, which was any one of the 25 tones in the set containing the 13 standard chromatic tones and the 12 interpolated quarter tones in the octave range falling between middle C (261.6 Hz) and the C an octave above (523.2 Hz). The tones were pure sinusoids with a duration of .5 sec each and a linear amplitude envelope having a .05-sec rise time, a .40-sec steady state, and a .05-sec fall time. Each of the 25 final tones appeared with each of the two context types once in each of the 10 blocks of trials. We randomly ordered trials within blocks and presented the blocks to the listeners in different random orders. There were five practice trials at the start of the experiment.

The procedure was identical to that of Experiment 1.

## *Results*

### *Individual Differences*

Again there was considerable variation in the rating profiles. Consequently, the intersubject correlations were analyzed by means of the program HICLUS. Using the compactness method, Subjects 3, 4, 6, and 8 were clustered together; Subjects 2, 5, and 7 formed another cluster; and Subject 1 was separated from the other listeners in the solution. Subject 1 was again the only subject claiming to have absolute pitch. Subjects 3, 4, 6, and 8 had on the average 5.2 yr. of formal instruction and 4.5 yr. of performing experience. Subjects 2, 5, and 7 had an average of 6.5 yr. of instruction and 3 yr. of performing experience. Not surprisingly, since these listeners were recruited from the music department, all subjects had at least a moderate level of instruction and performing experience and did not differ markedly from each other in this respect.

### *Judgments of Scale Completion*

*Subject 1.* The rating profiles of this highly trained musical listener, shown in Figure 7, exhibit similarities to the profiles she produced (as Subject 24) in Experiment 1. Her profiles were very similar for ascending and descending contexts. There was a strong preference for the tonic, C, over all other tones, and relatively high ratings were also given for the third and fifth degrees of the scale, E and G. In addition, there was a hint of a preference for the other scale tones over nonscale musical tones. However, there was no overall difference between the standard chromatic tones and the interpolated quarter tones. Rather, the ratings of the quarter tones seemed to approximate averages of the ratings of the neighboring chromatic tones.

*Subjects, 3, 4, 6, and 8.* These listeners also showed a general preference for the tonic, C, over the other tones. In addition, there were higher ratings for the scale tones than for nonscale tones. In particular, there was a rise in the ratings for tones near the third and fifth scale degrees, E and G. As with Subject 1, the ratings for the quarter tones approximated averages of the ratings for the neighboring chromatic tones. The profiles for this group reflected an increase in the distance effect at the expense of octave equivalence. Especially in the case of the lower tones in the range, tones nearer the context tones in pitch height were preferred to tones farther from the context tones. Thus the low C was judged as a better completion of an ascending context than a descending context.

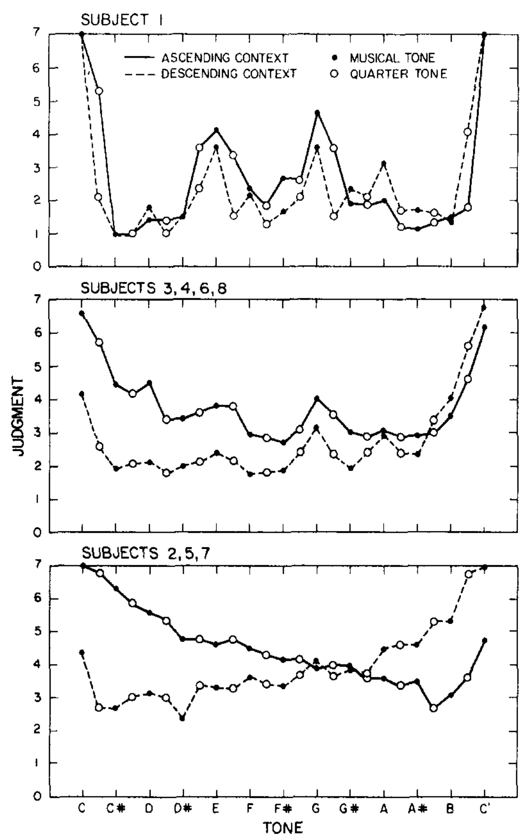


*Figure 7.* Mean rating profiles for three groups of listeners in Experiment 2, which used pure sinusoidal tones and included additional test tones interpolated halfway between the previous 13 chromatic tones.

*Subjects 2, 5, and 7.* The ratings by these subjects showed a still stronger distance effect, which held throughout the octave range of test tones, with higher ratings for low tones in ascending contexts and for high tones in descending contexts. Although there was no discernible difference between scale and nonscale tones, the upturn in the graphs at the end of the range indicated the presence of some degree of octave equivalence. Again, no difference was found between the chromatic tones and the interpolated quarter tones.

### *Discussion*

Listeners' judgments of scale completion in Experiment 2 provide no evidence of a preference for the standard musical tones over the quarter tones falling halfway between the musical tones. The ratings for the interpolated quarter tones generally were close to the average of the ratings for the two neighboring musical tones. This result suggests that the listeners were unable to distinguish clearly between tones differing by no more than a 24th part of an octave—at least under the conditions of this experiment, in which the tones themselves were free of harmonics and in which the test tone was presented after the context sequence and within a different octave. As Dowling (1978) suggests, there may be cognitive-psychological reasons why melodies in most human cultures are constructed using a restricted set of discrete pitches. For, as Dowling notes, this set typically contains from five to seven scale-step categories within each octave, as might be expected on the basis of Miller's (1956) well-known generalization concerning the limitation of the human capacity for categorical judgment to $7\pm2$ categories. Moreover, a number of investigators have found evidence for categorical perception of musical pitches that is similar to that found for speech sounds (Burns & Ward, 1978; Locke & Kellar, 1973; Siegel & Siegel, 1977a, 1977b; Blechner, Note 7; Halpern & Zatorre, Note 8).

The rating profiles showed many of the same characteristics as those of Experiment 1, with ratings of scale completion indicating again that pitch height, octave equivalence, and the tonal functions of the pitches are important factors underlying the listeners' judgments. Despite the use in Experiment 2 of the relatively less musical tones containing no overtone structure, the results of most listeners showed some degree of octave equivalence and the results of at least some of the listeners provided evidence for the hierarchy of tonal functions.

Although the subjects in Experiment 2 had on the average a moderate level of musical sophistication, the distance effect was generally stronger whereas the more structural factors of octave equivalence and tonal hierarchy were weaker than in Experiment 1. (This distance effect was particularly noticeable for the lower tones in the test range, where the absence of overtone structure may be more salient.) This finding is consistent with our general supposition that these structural factors tend to emerge only to the extent that the tones are perceived as musical. It is also possible that the inclusion of the unconventional quarter tones tended to confuse the listeners and hence weaken their maintenance of an orderly correspondence between the test tones in Experiment 2 and the listeners' discrete internal tonal system, which presumably makes no provision for such interpolated tones. We tentatively conclude that the harmonics usually contained in musical pitches may contribute to, although are probably not necessary for, the extraction by the listener of the musically important relations between tones.

Again, individual differences were found in the profiles, although these differences were not consistently related to the reported musical training of the subjects. The relatively small number of listeners who participated in Experiment 2 and the relatively gross measure of musical sophistication used, as well as the ambiguities introduced by the elimination of any overtone structure and the inclusion of quarter tones, may have obscured any relationship between the musical background of the subjects and the pattern of scale completion judgments, such as that found in Experiment 1.

## General Discussion

The results of Experiments 1 and 2, taken together with previously reported work, appear to us to be consistent with the following generalizations: To the extent that tones differing in frequency are not interpreted as musical stimuli—because they are presented in isolation from a musical context, because the tones themselves are stripped of harmonic content, or because they are played to musically unsophisticated listeners—the most potent factor governing their perceived relations is simply their separation along a one-dimensional continuum of pitch height. To the extent that tones are interpreted musically—because they are embedded in a musical context, because they are rich in overtones, and because they are played to musically sophisticated listeners—simple physical separation in log frequency gives way to structurally more complex and cognitive factors, including octave equivalence or its psychological counterpart, tone chroma, and a hierarchy of tonal functions specific to the tonality induced by the context.

That the data from some of the musically sophisticated listeners in Experiment 2 provided evidence for such an underlying hierarchy even when the tones were reduced to simple sinusoids argues against theories, widely held since Helmholtz (1863/1954), according to which the perceived relationships between tones is to be explained principally in terms of coincidences between frequencies of the harmonics (overtones or partials). Further evidence that cognitive factors must be invoked in addition to such physical factors is Krumhansl's (1979; Krumhansl, Note 5) demonstration that the perceived similarities between tones forming an interval of a fixed size depend on the relation of those tones to the contextually established tonal center.

In retrospect, we are inclined to use two related propositions to interpret the indication from Experiment 2 that the hierarchy of tonal function does not extend below the standard division of octaves into the semitone intervals of the chromatic scale. The first proposition is that there is a definite cognitive-psychological constraint on the number of categories that can be successfully distinguished by absolute judgments with respect to a single dimension (Dowling, 1978; Miller 1956)—here, presumably, the circular dimension of tone chroma (Balzano, 1977; Shepard, 1964). The second is that just as the perceptual interpretation of sounds as speech depends on the categorical assimilation of those continuously variable sounds to an underlying discrete set of phonemes (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), the perceptual interpretation of sounds as music depends upon the categorical assimilation of those continuously variable sounds to an underlying discrete set of tones (Burns & Ward, 1978; Siegel & Siegel, 1977a, 1977b; Blechner, Note 7; Halpern & Zatorre, Note 8) arranged in what we have been calling the tonal hierarchy.

Finally, our results reinforce previous indications that individual listeners differ widely in the extent to which they interpret musical tones in terms of such an underlying tonal system. From this we conclude that the indiscriminate pooling of data from such different listeners prior to analysis can obscure important patterns in those data. We suggest instead that the systematic pursuit of these individual differences by such data analytic techniques as hierarchical clustering (which we used here) or Carroll and Chang's (1970) individual differences scaling (which we have applied to this problem elsewhere; see Shepard, Note 9, Note 10) can help to separate out such underlying perceptual components as pitch height, tone chroma, and tonal function with respect to a contextually established tonality.

## Reference Notes

1. Krumhansl, C. L. Component factors in judgments of scale completion. In R. N. Shepard (Chair), *Cognitive structure of musical pitch.* Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
2. Dowling, W. J. Musical scales as cognitive structures. In R. N. Shepard (Chair), *Cognitive structure of musical pitch.* Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.

3. Balzano, G. J. The structural uniqueness of the diatonic order. In R. N. Shepard (Chair), *Cognitive structure of musical pitch.* Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
4. Cohen, A. J. Inferred sets of pitches in melodic perception. In R. N. Shepard (Chair), *Cognitive structure of musical pitch.* Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
5. Krumhansl, C. L. *The psychological representation of musical pitch in a tonal context.* Paper presented at the meeting of the Psychonomic Society, San Antonio, Texas, November 1978.
6. Balzano, G. Personal communication, 1978.
7. Blechner, M. J. *Musical skill and the categorical perception of harmonic mode.* (Status Rep. on Speech Perception SR-51/52). New Haven, Conn.: Haskins Laboratories, 1977. Pp. 139–174.
8. Halpern, A. R., & Zatorre, R. J. *Categorical perception and selective adaptation of simultaneous musical intervals.* Manuscript in preparation, 1978.
9. Shepard, R. N. The double helix of musical pitch. In R. N. Shepard (Chair), *Cognitive structure of musical pitch.* Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
10. Shepard, R. N. *Geometrical approximations to the structure of musical pitch.* Manuscript in preparation, 1979.

## References


Allen, D. Octave discriminability of musical and non-musical subjects. *Psychonomic Science,* 1967, *7,* 421–422.

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

Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America,* 1954, *26,* 751–753.

Balzano, G. J. Chronometric studies of the musical interval sense. (Doctoral dissertation, Stanford University, 1977). *Dissertation Abstracts International,* 1977, *38,* 2898B (University Microfilms No. 77-25, 643)

Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology,* 1943, *33,* 407–419.

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

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.

Burns, E. M., & Ward, W. I. Categorical perception—phenomenon or epiphenomenon? Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America,* 1978, *63,* 456–468.

Carroll, J. D., & Chang, J.-J. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika,* 1970, *35,* 283–319.

Cohen, A. *Perception of tone sequences from the Western-European chromatic scale: Tonality, transposition and the pitch set.* Unpublished doctoral dissertation, Queen's University at Kingston, Ontario, Canada, 1975.

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

Deutsch, D. Mapping of interactions in the pitch memory store. *Science,* 1972, *175,* 1020–1022. (a)

Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics,* 1972, *11,* 411–412. (b)

Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics,* 1973, *13,* 271–275.

Deutsch, D. Delayed pitch comparisons and the principle of proximity. *Perception & Psychophysics,* 1978, *23,* 227–230.

Dewar, K. M. *Context effects in recognition memory for tones.* Unpublished doctoral dissertation, Queen's University at Kingston, Ontario, Canada, 1974.

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. The 1215-cent octave: Convergence of Western and Nonwestern data on pitch-scaling. *Journal of the Acoustical Society of America,* 1973, *53,* 373A (Abstract) (a)

Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology,* 1973, *5,* 322–337. (b)

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

Dowling, W. J. Musical scales and psychophysical scales: Their psychological reality. In T. Rice & R. Falck (Eds.), *Cross-cultural approaches to music: Essays in honor of Mieczyslaw Kolinski.* Toronto, Canada: University of Toronto Press, in press.

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.

Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics,* 1977, *21,* 60–64.

Helmholtz, H. von. *On the sensations of tone as a physiological basis for the theory of music* (A. J. Ellis, Ed. and trans.). New York: Dover, 1954. (Originally published, 1863.)

Humphreys, L. F. Generalization as a function of method of reinforcement. *Journal of Experimental Psychology,* 1939, *25,* 361–372.

Johnson, S. C. Hierarchical clustering schemes. *Psychometrika,* 1967, *32,* 241–254.

Kilmer, A. D., Crocker, R. L., & Brown, R. R. *Sounds from silence: Recent discoveries in ancient near eastern music.* Berkeley, Calif.: Bīt Enki Publications, 1976.

Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology,* 1979, *11,* 346–374.

Levelt, W. J. M., Van de Geer, J. P., & Plomp, R. Triadic comparisons of musical intervals. *British Journal of Mathematical and Statistical Psychology,* 1966, *19,* 163–179.

Liberman, A. M., Cooper, F. S., Shankweiler, D., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review,* 1967, *74,* 431–461.

Licklider, J. C. R. Basic correlates of the auditory stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951.

Locke, S., & Kellar, L. Categorical perception in a nonlinguistic mode. *Cortex,* 1973, *9,* 355–369.

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

Miller, G. A. The magic number seven, plus or minus two. *Psychological Review,* 1956, *63,* 81–97.

Piston, W. *Harmony.* New York: Norton, 1941.

Ratner, L. G. *Harmony: Structure and style.* New York: McGraw-Hill, 1962.

Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review,* 1929, *36,* 172–180.

Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America,* 1964, *36,* 2346–2353.

Shepard, R. N. Representation of structure in similarity data: Problems and prospects. *Psychometrika,* 1974, *39,* 373–421.

Shepard, R. N. Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization.* Hillsdale, N.J.: Erlbaum, in press.

Siegel, J. A., & Siegel, W. Absolute identification of notes and intervals by musicians. *Perception & Psychophysics,* 1977, *21,* 143–152. (a)

Siegel, J. A., & Siegel, W. Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics,* 1977, *21,* 399–407. (b)

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

Stevens, S. S., Volkmann, J., & Newman, E. B. A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America,* 1937, *8,* 185–190.

Thurlow, W. R., & Erchul, W. P. Judged similarity in pitch of octave multiples. *Perception & Psychophysics,* 1977, *22,* 177–182.

Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America,* 1954, *26,* 369–380.

Received August 28, 1978 ■

## Correction to Kirsner and Sang

In the article "Visual Persistence and Code Selection in Short-Term Memory" by Kim Kirsner and D. L. Sang (*Journal of Experimental Psychology: Human Perception and Performance,* 1979, Vol. 5, No. 2, pp. 260–276), Figures 2 and 3 were mislabeled. The figure on page 264 should be labeled Figure 3, Experiment 2, and the figure on page 265 should be labeled Figure 2, Experiment 1.