Psychological Review
1978, Vol. 85, No. 4, 341–354

Scale and Contour: Two Components of a Theory of Memory for Melodies

W. Jay Dowling
University of Texas at Dallas

This article develops a two-component model of how melodies are stored in long- and short-term memory. The first component is the overlearned perceptual-motor schema of the musical scale. Evidence is presented supporting the lifetime stability of scales and the fact that they seem to have a basically logarithmic form cross-culturally. The second component, melodic contour, is shown to function independently of pitch interval sequence in memory. A new experiment is reported, using a recognition memory paradigm in which tonal standard stimuli are confused with same-contour comparisons, whether they are exact transpositions or tonal answers, but not with atonal comparison stimuli. This result is contrasted with earlier work using atonal melodies and shows the interdependence of the two components, scale and contour.

Remembering melodies is a basic process in the music behavior of people in all cultures. This behavior may involve production, as with the singer performing a song for an audience or the participant in a significant social event trying to remember the appropriate song. Or it may involve recognition, as with the listener whose comprehension of the later developments in a piece depends on memory for earlier parts. In this article, I concentrate on two components of memory that contribute to the reproduction and recognition of melodies, namely, melodic contour and the musical scale. I maintain that actual melodies, heard or sung, are the product of two kinds of underlying schemata. First, there is the melodic contour—the pattern of ups and downs—that characterizes a particular melody. Second, there is the overlearned musical scale to which the contour is applied and that underlies many different melodies. It is as though the scale constituted a ladder or framework on which the ups and downs of the contour were hung.

Two examples epitomize the behavior this theory addresses. First, if people in our Western European culture hear a melody from some non-Western culture using a non-Western scale, their reproductions of that melody will use their own Western scale, preserving the contour of the original melody. The scale functions as a classic example of a sensorimotor schema controlling perception and behavior. In this example, the non-Western melody is assimilated to the Western schema (Francès, 1958, p. 49). Part of the education of ethnomusicologists is directed toward freeing them from their native schemata so that they can hear accurately the pitches of non-Western music.

The second example involves the use of melodic contour in the structure of pieces of music. One way a composer can tie a piece together is to repeat the same contour at different pitch levels, at different relative place-

Earlier versions of this article were presented at the Music Educators National Conference, Western Division, San Francisco, 1975, and the meeting of New American Music, Tangents III, University of Wyoming, 1977. Support for the experiment reported here was provided in part by a Biomedical Sciences Support Grant to the University of Texas at Dallas from the National Institutes of Health.

I thank Dane Harwood, Darlene Smith, James Bartlett, Drew Hamre, Mark Goedecke, Cynthia Null, Kelyn Roberts, and Klaus Wachsmann for helpful discussions and comments. And I especially wish to thank Mantle Hood for bringing the data on Indonesian tunings to my attention.

Requests for reprints should be sent to W. Jay Dowling, Program in Psychology and Human Development, University of Texas at Dallas, Richardson, Texas 75080.

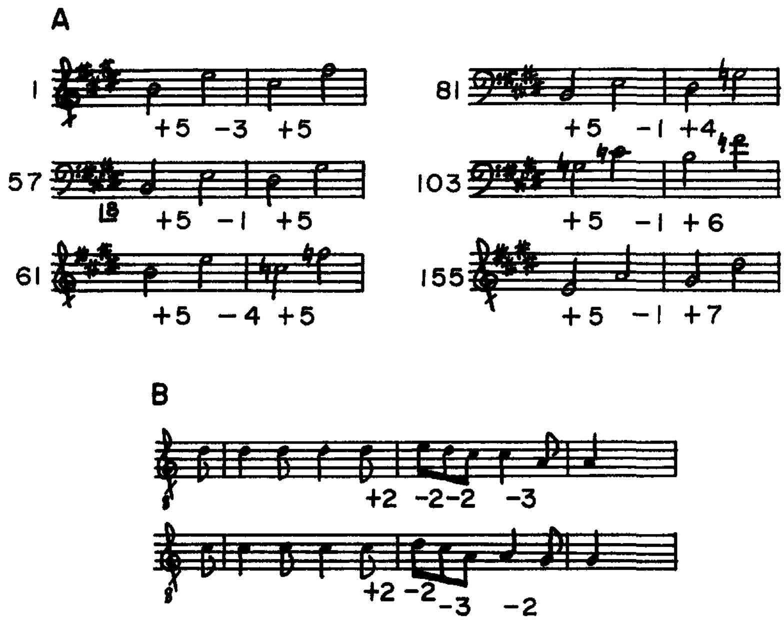

Figure 1. Section A shows examples from Beethoven's Piano Sonata opus 14 no. 1, illustrating use of the same melodic contour with different interval sizes. (Intervals between notes are shown below the staves in semitones. Excerpts are labeled with measure numbers.) Section B is an American Indian example from Kolinski (1970, p. 91).

ments on the scale, or even on different scales. The repetition provides unity without becoming boring through being too exact. Figure 1A demonstrates Beethoven's use of this device in his Piano Sonata opus 14 no. 1. Such a device relies on the listener's ability to recognize the melodic contour through transformations of pitch. That such processes are not confined to Western music is shown by Kolinski's (1970) example from the Flathead Indians (see Figure 1B). Adams (1976) has provided a guide to the uses of contour conceptualizations in ethnomusicology.

In what follows, I will discuss (a) scales of pitch and their characteristics and (b) melodic contours and their independence in memory. (*Independence* is here used in the sense of referring to cases where one remembers one thing without remembering a related thing.) I will also present a new experiment that illustrates performance in memory for melodies consistent with the present two-component theory.

Musical Scales

Several aspects of musical scales as they function in various cultures are of particular interest to the psychologist. First, scales consist of discrete steps of pitch, typically five or seven to the octave. This limitation would seem to match quite well the limitations of people as information processors. Second, musical scales of pitch are approximately logarithmic with respect to frequency, a fact that bears directly on the problem of pitch scaling in psychophysics. Third, musical scales can be slid continuously up and down in pitch without distorting their relative interval sizes. They serve as highly stable sensorimotor schemata with great flexibility of application.

Discrete Scale Steps

With few exceptions, the cultures of the world use discrete changes of pitch along a musical scale in creating their melodies. Since the singing voice and many early instruments are capable of producing a continuous variation of pitch, how is it that the music of the world moves by discrete steps? Helmholtz (1954) raised this question, and some of the answers he suggested are appealing even today. The ultimate biological functions of discrete steps are as obscure as the functions of melody or of music itself. However, given that we have melody, discrete scale steps can be very useful. Helmholtz argued that the division of the pitch continuum into discrete steps is desirable in order that the degree of melodic and rhythmic movement might be immediately apparent to the listener. If pitch moved continuously in a melody, he argued, it would be much more difficult to decide exactly how much the pitch had changed or what rhythmic, temporal units had elapsed. In fact, such a decision could be made only on the basis of a fixed set of discrete reference points anyway. That is, the listener would have to learn some form of scale as a cognitive framework through which to deal with the melodies he hears, continuous or not.

The modern cognitive psychologist can easily carry Helmholtz's argument one step further. Granting that we need discrete scale steps to comprehend melodic movement, how many should there be? Miller's (1956) classic article suggested that human beings are limited in their categorical judgment capability to using 7 ± 2 categories on any one dimension. Most cultures fall within this range, using five or seven distinct scale step categories within the octave. As will be discussed below, the system of categories repeats cyclically every octave. Thus, in Miller's terms, a pitch would be categorized in terms of two dimensions: one for pitch level within the octave (often called *chroma*) and the other for octave level. Certain cultures that would seem at first sight to provide exceptions to this pattern turn out on closer inspection not to. For example, the music of India typically uses many very narrow intervals of which 22 are available within an octave. However, melodies are basically constructed out of scales having seven or so focal pitches, and the other neighboring tones are used as auxiliary tones around these focal pitches for the purposes of ornamentation.

At this point, I need to introduce a conceptual scheme for talking about musical scales. This scheme has four levels of abstraction from the pitches of actual melodies and shares certain features with the conceptual schemes of Hood (1971) and Deutsch (1977). I have discussed it in greater detail in another paper (Dowling, Note 1). The most abstract level is that of the psychophysical scale, which is the general rule system by which pitch intervals are related to frequency intervals of tones. I will argue that the appropriate form for the psychophysical scale is approximately logarithmic, owing to the fundamental place of octave judgments in human auditory processing. The second level is that of tonal material, which includes the set of pitch intervals in use by a particular culture or within a particular genre. In Western music, the tonal material would include the set of semitone intervals of the equal tempered chromatic scale but not anchored to any particular frequency. The third level is that of the tuning system, which consists of a selection of a subset of the available pitch intervals from the tonal material that are used in actual melodies. In Western music, the tuning system might consist of the set of intervals represented by the white notes of the piano—the set of intervals with the ascending cycle (in semitones) of [2, 2, 1, 2, 2, 2, 1] repeating every octave. This set of intervals can become the basis of any of the heptatonic modes (that is, the modes with seven notes per octave): major, minor, or the medieval church modes. The set of intervals of the black notes on the piano [2, 3, 2, 2, 3] can also function as a tuning system for the Western pentatonic modes.

The most concrete level is mode. In going from tuning system to mode, two things happen. First, an anchor for the frequencies is established, and what were pitch intervals at the more abstract levels get translated into the pitches of notes. Second, a tonal focus in the tuning system is selected, and a tonal hierarchy is established on the tuning system. (The tonal hierarchy determines which notes in the mode are more important and dynamic tendencies of the notes as they function in melodies.) These two operations are usually combined in Western music under the rubric of "selecting a tonality," for example, taking the intervals of the white-note tuning system and selecting the key of C-major, A-minor, Eb-major, or C-dorian. Since *mode* in my system overlaps considerably with the term *scale* as used in the phrases *musical scale* and *diatonic scale,* I will often use scale in what follows as a synonym for mode. I want to caution the reader, however, that scale is also used in ways that are not synonymous with my use of mode, for example, in the phrase *chromatic scale,* which is more nearly synonymous with my term *tonal material.* My use of scale is informal; a strictly formal system would use only mode. I use the term *tonal* to refer to melodies using the notes of well-known modes. And I should caution the reader that what I have to say here is mostly irrelevant to the dispute among proponents of tempered, Pythagorean, and other "natural" tuning systems well discussed by Ward (1970).

Logarithmic Scale

The musical scales of most cultures repeat themselves cyclically every octave. Tones an octave apart are treated as equivalent in some sense and are typically given the same name (C, D, E; do, re, mi; or Barang, Sulu, Data [Indonesian]). Tones that are psychologically and musically an octave apart have a ratio between their fundamental frequencies of approximately 2/1. (The "approximately" will be discussed below.) Ward (1954) has shown that Westerners are quite precise in setting two pure tones (presented alternately) to the subjective interval of an octave. This is all that is required to produce a logarithmic psychophysical scale of pitch: Tones at equal distances along the subjective pitch scale (namely, an octave apart) should be related by equal frequency ratios.

Complications arise because the subjective octave usually represents a frequency ratio just slightly greater than 2/1. Ward (1954; replicated by Burns, 1974, using non-Western subjects and by others [Sundberg & Lindqvist, 1973] using Westerners) found that through the midrange of up to frequencies of about 900 Hz, the frequency ratio of a subjective, pure-tone octave was about 2.02/1. This ratio still leads to an approximately logarithmic scale. Ward's results are plotted in the curve of Figure 2, which shows the tuning of successive octaves as cumulative deviations from the 2/1 frequency ratio in logarithmic units (hundredths of a semitone, or cents). (A semitone represents a frequency ratio of $2^{1/12}/1$). Octaves having a 2/1 frequency ratio would be represented in Figure 2 as a horizontal line. Straight segments of the curve with positive slope represent logarithmic pitch scales with a greater than 2/1 frequency ratio to the octave.

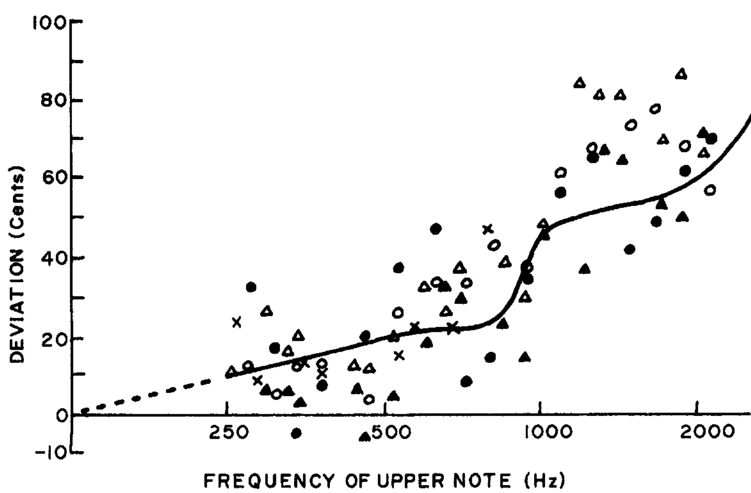

Figure 2. Cumulative tuning deviations of octaves in cents (hundredths of a semitone) as a function of the fundamental frequency of the upper note in the octave for Burmese harp (Xs; Williamson & Williamson, 1968) and for Indonesian gamelans using *sléndro* (circles) and *pélog* (triangles) tuning systems. (The Indonesian data are based on Hood [1966; filled symbols; $N = 2$] and Surjodiningrat, Sudarjana, & Susanto [1969; open symbols; $Ns = 11$ and 10 for *sléndro* and *pélog,* respectively]. The solid line is based on Ward's [1954] data for pure-tone octave judgments.)

Figure 2 also shows some measurements of non-Western instrument tunings: a set from a Burmese harp (Williamson & Williamson, 1968) and several sets from Indonesian gamelans (Hood, 1966; Surjodiningrat, Sudarjana, & Susanto, 1969). *Gamelan* refers to both the orchestra and the set of instruments, consisting chiefly of gongs and marimbas. The correspondence of the tunings and Ward's curve is very good, especially considering the sorts of systematic deviations within the octave in gamelan tunings discussed by Hood (1966). One reason that the fundamental frequencies of gamelan tunings are free to approximate the tunings of subjective octaves is that there is not the same attempt as found in the West to make fundamentals of higher tones match upper partials (overtones) of lower tones. One reason for this is that acoustically gongs have very irregular series of upper partials.

Western pianos are tuned to octaves that are slightly greater than 2/1 ratios, but the deviation is smaller than that for subjective pure-tone octaves (Martin & Ward, 1961). The reason for the deviation is that piano strings are not ideal, frictionless, vibrating bodies but have a certain amount of stiffness. This stiffness is most pronounced at the upper and lower extremes of the keyboard, where the ratio of the diameter to the length of the string is relatively large. The stiffness of the strings leads the frequencies of upper partials to be progressively sharper than true harmonics (which stand in integer ratios to the fundamental). Hence, tuning upper fundamentals to partials of lower notes will lead to stretched octaves, but they are not so severely stretched as those of the human ear. Our Western emphasis on consonance of simultaneous tones leads us to prefer this piano tuning over one in which melodic succession would be given precedence. (See Plomp & Levelt, 1965, for a discussion of the contribution of coincidence of upper harmonics to the consonance of complex tones.) These considerations also suggest that for the same reasons, a cappella choirs and string quartets will tend to use a 2/1 octave, especially during slow passages.

The precision of judgment of Ward's (1954) subjects is a very good reason for preferring a quasi-logarithmic scale to other candidates for the psychophysical scale relating pitch to frequency (Ward, 1970). Other reasons include the above cross-cultural data and the fact that one of Ward's (1954) subjects who had absolute pitch produced the same scale by note labeling as with the octave judgment method. Another strand of evidence is provided by Null (1974), who obtained a logarithmic scale using a slightly modified magnitude estimation method. One of the cleverest corroborations of the log scale is provided by an experiment of Attneave and Olson (1971). They observed that musically untrained subjects had difficulty in producing transpositions of arbitrarily selected pitch intervals. Then, they hit upon the idea that certain interval sequences might be overlearned even by the uninitiated, for example, the National Broadcasting Company (NBC) chimes. They found that the NBC chimes had always been presented at the same pitch level —on the notes G-E-C in the middle register, an acronym for the General Electric Corporation. They asked subjects to produce the pattern at various pitch levels above and below the original. The scaling question was, Would transpositions follow a log scale, or a power function, or something else? Subjects overwhelmingly responded with transpositions along a logarithmic scale. Thus, given a meaningful task, untrained subjects use a logarithmic scale for pitch.

Attneave and Olson's subjects demonstrate how stable the tonal scale system is throughout life once it is learned. So do Francès's (1958) subjects, who found it easier to notice changes on rehearing tonal melodies than atonal. Tonal scales constitute one of the most durable families of perceptual-motor schemata that have been observed in psychology, ranking perhaps only after the schemata of natural language in their stability and resistance to change in adult life. In fact, the same kinds of categorical perception phenomena found in the phonetics of a language seem to hold for musical scale pitch judgments (Burns, 1974). It would probably surprise most American psychologists how early in life these perceptual-motor schemata are acquired. Im-

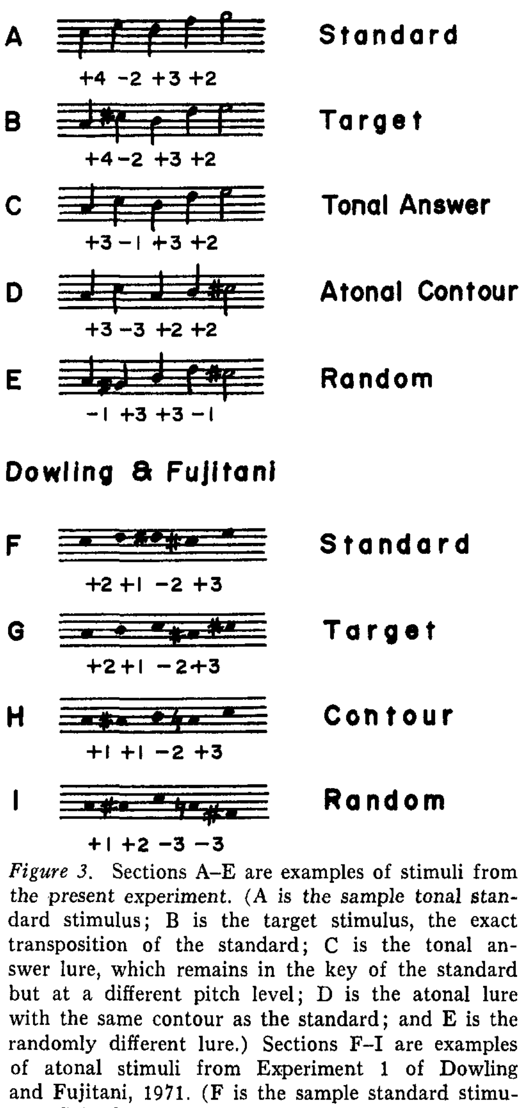

Figure 3. Sections A–E are examples of stimuli from the present experiment. (A is the sample tonal standard stimulus; B is the target stimulus, the exact transposition of the standard; C is the tonal answer lure, which remains in the key of the standard but at a different pitch level; D is the atonal lure with the same contour as the standard; and E is the randomly different lure.) Sections F–I are examples of atonal stimuli from Experiment 1 of Dowling and Fujitani, 1971. (F is the sample standard stimulus; G is the target stimulus, the exact transposition of the standard; H is the same-contour lure; and I is the randomly different lure. All staves are notated with a treble clef understood. Intervals in semitones are shown under each stave.)

berty (1969) found that 8 year olds are able to notice shifts of scales within a melody from one key to another a semitone or two higher. They found shifts of mode that left the underlying pattern of intervals unchanged (as in the tonal answer stimuli of the experiment reported below) much more difficult. Zenatti (1969) found that 8 year olds can recognize three-note tonal melodies much better than atonal. Five year olds find tonal and atonal melodies equally difficult because they have not yet thoroughly internalized the scale. Eleven year olds find both types equally easy. But with four- and six-note melodies, even adults find the atonal more difficult, in agreement with Francès (1958).

In this discussion of musical scales, I have not tried to incorporate Shepard's (1964) helical theory of pitch scales. I wish to note, however, that I believe the present theory compatible with his. It only requires that the helix be sprung somewhat to allow for Ward's enlarged subjective octaves. I believe Shepard's demonstration of the auditory barber pole could be replicated using Ward's quasilogarithmic scale. The issue of the acceptance of the helical model as a pitch scale has not been settled, however. For example, Deutsch (1972) has demonstrated that in the recognition of familiar tunes, chroma (the quality of C-ness, D-ness, or E-ness; do-ness, re-ness, or mi-ness) does not suffice as a cue to tune identity.

Memory for Melodic Contour

For everyone except the small percentage of the population having absolute pitch, well-known melodies must be stored as sequences of pitch intervals between successive notes (Deutsch, 1969). In this section, I have assembled evidence that memory for the contour (the ups and downs of the melodic intervals) can function separately from memory for exact interval sizes. That is, the contour is an abstraction from the actual melody that can be remembered independently of pitches or interval sizes. This is true for melodies in both short-term and long-term memory.

Retrieval from Short-term Memory

The contours of brief, novel atonal melodies can be retrieved from short-term memory even when the sequence of exact intervals cannot. This point is illustrated by an experiment by Dowling and Fujitani (1971, Experiment 1). They used a short-term recognition memory paradigm in which the standard stimulus on each trial was a randomly generated five-note melody having small pitch intervals between successive notes. The comparison stimulus followed the standard after a brief pause. Three types of comparison melodies were the same as the standard in both contour and pitch intervals, or had the same contour but different intervals, or were different in both contour and intervals (i.e., were novel random sequences). These are illustrated in Figure 3, Sections F–I. Different groups of subjects had the task of distinguishing among the following types of comparison melody: exactly the same versus random, exactly the same versus same contour, and same contour versus random. Half the subjects were given comparisons starting on the same pitch as the standard; the other half had comparisons whose starting note was transposed to another pitch level. The results showed that when the comparison starts on the same pitch level as the standard, exact-same comparisons are easily distinguished from either random comparisons or same-contour, different-interval comparisons. In other words, it was easy to reject any comparison stimulus that did not contain exactly the same pitches as the standard. When the comparison melodies were transposed, however, exact-same targets and same-contour comparisons were easily distinguished from random ones but almost impossible to tell apart. Listeners responded on the basis of the presence or absence of the contour and were unable to recognize the sameness of intervals when these were added to the contour.

Thus, when the listener is trying to retrieve an exact set of intervals from memory, the best he can do under the above conditions is retrieve the contour. It is critical to this result that the target melodies be atonal, that is, not using the pitches of a musical scale familiar to the subjects. Francès (1958) found alterations in tonal melodies easier to detect than alterations in atonal melodies. A major theme in the present article is that there are two components at work in the normal process of melody recognition: contour and scale. Dowling and Fujitani (1971, Experiment 1) explored the extreme case where the role of the scale had been all but eliminated. They found contour recognition to be completely dominant over pitch interval recognition. However, one would not expect the same result with tonal melodies. The overlearned scale framework should make recognition of the difference between a tonal target melody and an atonal lure having the same contour much easier. (These stimuli are illustrated in Figures 3B and 3D.)

Distinguishing between a tonal melody and a tonal lure having the same contour as the first melody but starting on a different note of the same modal scale should be very difficult. The relationship between this last pair of melodies is the same as that of a fugue subject and its *tonal answer*. Such a pair is illustrated by Figures 3A and 3C. There are two ways to think of the tonal answer in terms of the conceptual scheme of tuning system and mode.

1. We might think of the tonal answer as consisting of the same set of diatonic intervals translated into a new mode. The interval pattern of the tuning system (in Figure 3, the white notes of the piano) remains fixed at the same pitch level, while the tonality of the melody in the sense of its starting pitch level is moved to a new place in the tuning system. Thus, both the standard and the tonal answer in Figure 3 (A and C) have the diatonic intervals $[+2, -1, +2, +1]$, but the first is in C-major and the second is in A-minor.

2. The tonal answer might be thought of as remaining in the same mode as the standard, which is exactly the way a tonal answer is treated in a fugue. In that case, the pitches of the mode remain fixed while the starting pitch of the melody is shifted to a different degree of the scale. Both melodies in Figure 3 (A and C) would remain in C-major, and the difference between them would be that the standard begins on the first degree and the tonal answer on the sixth degree of the modal scale.

I prefer the second of these characterizations for describing the cognitive processing involved in the experiment presented below. This is because the time interval between the presentation of the standard and the tonal answer is short. The standard is coded as being in a particular major mode. When the listener hears the tonal answer immediately following the standard, nothing in that comparison pattern demands that he change mode. Therefore, he hears the tonal answer in the same mode as the standard. Diatonic interval pattern (contour in a restricted sense) and mode can function as features of the standard. Tonal answers would then be difficult to distinguish from standard stimuli because they share these two features. Exact transpositions force the listener to change mode in midtrial. Exact transpositions share the feature of having the same intervals as the standards when measured in semitones (that is, at the level of tonal material). The experiment thus brings these sets of feature similarities into conflict with each other.

Lest too persuasive an argument here make the results of the experiment seem a foregone conclusion, let me introduce a plausible theory that makes a different prediction from the two-component scale–contour theory. It could be, according to this theory, that contours are not stored in memory independently of interval sizes. Dowling and Fujitani got their result because they used unnatural atonal melodies. Atonal intervals are difficult to remember, both because listeners have not had much practice with them and because they do not occur as part of an overlearned scale schema. However, if we replicate Dowling and Fujitani using tonal materials, we will get a very different result. The interval sequences of tonal standard melodies will be easily learned because they fall on a well-known scale. Thus, changes in those intervals will be easily noticed whether the change is to an atonal melody or to a tonal melody in another mode. The listeners should at least do better than chance in discriminating exact transpositions from tonal answers. (To give this theory its due, I should note that pilot work with a few professional musicians convinces me that they can perform in the way just described).

The two-component theory, on the other hand, claims that even with tonal melodies, contour (in the sense of ups and downs measured in diatonic intervals) and interval sizes (measured in semitones at the level of tonal material) are stored independently, the latter simply as a mode label. Tonality can function as a cue to distinguish a tonal melody from an atonal one. But changes in interval size that leave tonality intact will be difficult to notice.

Experiment on Tonal Melodies

This experiment is based on the paradigm of Dowling and Fujitani (1971, Experiment 1). It copies the conditions of that experiment in which comparison melodies were transposed (i.e., began on a different pitch from the standard melodies) and in which the subject's task was to recognize only exact transpositions of the standards. The most important difference between the two experiments is that in Dowling and Fujitani, all melodies were atonal; while in the present experiment, all but one type of comparison melody were tonal. Other differences in method derived from differences in the available equipment, subject populations, and the desirability of equalizing expected interval size between tonal and atonal melodies used in the present experiment. Since the most comparable conditions in the two experiments—those of distinguishing targets from randomly different comparisons—gave roughly comparable results (81% and 84% vs. 89%), the effects of changes in these other variables are apparently negligible for purposes of establishing the qualitative results toward which this study is directed.

Using tonal melodies in this paradigm makes possible the introduction of one more type of comparison melody: the tonal answer. In this type, the comparison begins on a different pitch from the standard but stays within the same diatonic scale. Thus, the pitch intervals between tones will generally be changed, but the melody will still be tonal. Figure 3 shows examples of this and the other types of comparison stimuli used in the present experiment and that of Dowling and Fujitani.

Method

Subjects. Twenty-one students at the University of Texas at Dallas served in four separate group sessions, most of them receiving extra credit in upper level psychology courses for their participation. Subjects were divided into groups on the basis of a postexperiment questionnaire. Subjects in the experienced group had 2 or more years of musical training (including studying an instrument or voice or playing an instrument in an ensemble but excluding taking music courses or singing in choir). In the inexperienced group, subjects had less than 2 years training. Means for the two groups were 5.0 and .14 years training, respectively. The mean age of subjects was 30.4 years and was comparable in the two groups. There were three males and seven females in the experienced group and seven males and four females in the inexperienced group (a difference that probably reflects a tendency to give females music lessons in our culture). Performance of males and females in the two groups was roughly comparable, with a slight superiority of experienced males.

Procedure. Subjects were instructed that this was an experiment in memory for melodies. On each trial, they heard a pair of brief melodies, and their task was to say whether the two were the same or different. For this purpose, they responded using a four-category scale with responses of "sure same," "same," "different," and "sure different." There were 48 trials in the experiment and they wrote their responses on a sheet of paper provided. The experimenter explained that all comparison stimuli, even those that were the same, would start on a different note from the standard. What was important in determining sameness was that not only the ups and downs but also the distances among the notes be the same. The experimenter referred to the possibility of singing "Happy Birthday" starting on different notes as an illustration of how the same melody could occur at different pitch levels as long as the distances among the notes remained the same. The experimenter presented one example of each of the types of stimulus pairs. After responding to questions, the experimenter then presented the 48 trials.

Stimuli. Stimuli were played on a freshly tuned Steinway piano and recorded on tape. They were presented to subjects over loudspeakers at comfortable listening levels via high-quality reproduction equipment. The timing of stimuli was controlled by a Davis timer producing clicks every .67 sec, which were barely audible on the tape. In all stimuli, the rate of presentation of the quarter-note values (♩) of Figure 3 was 3 tones per sec. Thus, each stimulus was 2 sec in duration. There was a 2-sec blank interval between standard and comparison stimuli on each trial and a 5-sec response interval following the comparison stimulus. A warning, consisting of the experimenter's voice announcing the trial number, preceded the onset of each trial by 2 sec.

All standard stimuli began on middle C (fundamental frequency of 262 Hz). There are 16 possible contours for five-note sequences, providing unisons are excluded. The contours of the standard stimuli on each of the 48 trials of the experiment were selected by using three successive random permutations of the order of the 16 contours. Thus, each contour appeared as a standard exactly three times in the experiment. The interval sizes between notes of the tonal stimuli were chosen with the following probabilities of diatonic scale steps: P (±1 step) = .67; P (±2 steps) = .33.

There were four types of comparison melody. Each of these four types occurred equally often in the three blocks of 16 trials (randomly determined). All comparison melodies began on either the E above or the A below middle C (randomly determined). These transpositions were chosen as being moderately distant from C both in pitch level (+4 and −3 semitones) and in shared pitches (three and four, respectively). (Francès, 1958, has shown "remote" transpositions in the latter sense harder to recognize than "near" transpositions.) Target comparison stimuli (see Figure 3B) were exact transpositions to the standards to the key of E or A. They retained exactly the interval sizes of the standards. Tonal answer comparison stimuli (see Figure 3C) were lures that started on E or A but that remained in C, the same key as the standard, and had the same contour and diatonic intervals. Thus, the interval sizes in semitones did not remain the same as in the standard, while the tonal scale remained the same.

Atonal same-contour comparison stimuli (see Figure 3D) were lures that retained the contour of the standard but that used intervals randomly selected without regard for tonal scales. The probabilities of interval sizes of the atonal lures were P (±1 semitone) = .17, P (±2 semitones) = .33, and P (±3 semitones) = .50. (These probabilities were chosen so that the expected interval size of a tone sequence would be 2.33 semitones. This is comparable to the expected interval size of 2.27 semitones for the tonal sequences. The latter figure was calculated as a weighted average of the diatonic interval sizes, without regard for starting pitch. The fact that all standard stimuli began on C, immediately next to a 1-semitone interval, would lower this estimate slightly in the present case.) The random different stimuli (see Figure 3E) were lures having different contours from the standards (randomly selected) and different diatonic scale intervals selected just as with the standards.

Data analysis. The four-category confidence judgments were used to generate individual memory operating characteristics (MOC) for each subject for each of the three types of stimulus comparisons. Hit rates were given by responses to the target stimuli. These hit rates were plotted separately against three sets of false alarm rates given by responses to tonal answer, atonal contour, and random lures to give three areas under the MOC for each subject. *Areas under the* MOC can be interpreted as an estimate of what the proportion correct would be in the case where chance performance is .50 (Swets, 1973). Areas under the MOC were evaluated using an unweighted-means analysis of variance for unequal cell sizes due to the unequal numbers of experienced and inexperienced subjects.

Table 1
Areas Under the Memory Operating Characteristic of the Experiment Compared with Similar Conditions of Dowling and Fujitani's (1971) Experiment 1

Group	Target vs. tonal answer	Target vs. atonal contour	Target vs. random
Experienced	.48	.79	.84
Inexperienced	.49	.59	.81
Dowling and Fujitani	—	.53	.89

Results

Table 1 shows the mean areas under the MOC for the two groups and the three stimulus comparisons. The main effect of stimulus comparison type was significant, $F(2, 38) = 77.35$, $p < .001$. Distinguishing between targets and tonal answer lures was very difficult, with chance performance in both groups. Distinguishing targets from atonal same-contour lures was somewhat easier, and distinguishing targets from random lures was easiest of all. The Experience × Stimulus Type interaction was significant, $F(2, 38) = 8.01$, $p < .001$, mainly reflecting a difference of ability in distinguishing targets from atonal same-contour lures. The main effect of experience approached significance at the .05 level. This modest effect is consonant with the modest correlations found by Dowling and Fujitani between performance on their task and experience.

General Discussion of Short-term Recognition of Melodies

The present results illustrate the importance of scale as well as contour in short-term recognition memory for melodies. This point is brought out by comparison with the results of Dowling and Fujitani (see Table 1). Both studies found that the two melodies are relatively easy to distinguish in that they have different contours, with performance in the .80s. Dowling and Fujitani's subjects found it difficult to distinguish between two atonal melodies with the same contour but different interval sizes, performing at around the chance level of .50. In the present experiment, subjects found it easier to reject an atonal comparison melody when it was preceded by a tonal standard melody. This was especially true of experienced subjects who presumably have a firmly internalized modal system. What subjects in the present experiment found extremely difficult, performing at chance, was distinguishing between exact transpositions of comparison melodies to new tonal keys and shifts of the contour along the same diatonic scale as the standard. Both experienced and inexperienced subjects had trouble with this task. Phenomenologically, the comparison stimuli in Pairs A–B and A–C of Figure 3 sound "natural," while the comparison in Pair A–D sounds "strange."

This result illustrates the separateness of the functions of contour and mode. The function of mode is not to fix a set of intervals in semitones as belonging to a melody. If it were, tonal answers would not be confused with exact transpositions. The mode is simply a framework on which the contour may be hung. For brief melodies heard only once, the point on the modal scale where the melody begins is not taken into account. What subjects seem to account for is that both the mode and the contour are the same in the two melodies.

This result should not be taken to mean that diatonic scale intervals are somehow psychologically equal. The concept of subjective equality applies best to the level of the psychophysical scale. The aesthetic purpose of using differently sized pitch intervals in modal scales would be lost if all the intervals in the mode were subjectively the same. Different intervals are used because they provide melodic interest. Melodies are translated into different modes because that is more interesting than reiterating them in the same mode. Musicians find tuning systems that use only subjectively equal intervals dull because they deny a major source of melodic variety. This is the principal criticism of the "whole-tone scale" with which Debussy and others experimented around the turn of the century.

Moreover, Francès (1958, Experiment 2) found evidence that makes the subjective equality of modal intervals difficult to believe in. Francès mistuned certain notes on a piano and then placed these notes in various melodic contexts. He found that when the mistuning was in the direction of the dynamic tendency of the tone in its modal context, the alteration was much less likely to be noticed than when the mistuning was in the opposite direction. This means that the subjective interval size changes with modal context. Such changes would be consonant with the theory that diatonic intervals tend to be subjectively equal provided that the changes due to context make large intervals (in the sense of the psychophysical scale and tonal material levels) smaller and small intervals larger. However, Francès got the opposite result. In all of his cases, it was small intervals that became smaller.

The confusion of target melodies and tonal answers does not occur with well-known melodies stored in long-term memory. Attneave and Olson (1971) have shown that with familiar melodies, exact interval sizes are precisely remembered. In fact, familiar melodies can be used as mnemonics for retrieving scale intervals, as when a music student uses the song "Over There" to remember descending major sixths (-9 semitones) or the song "There's a Place for Us" for ascending minor sevenths ($+10$ semitones). But even for familiar melodies, the contour and mode can function independently, as when "Frère Jacques" or "Twinkle, Twinkle" are sung in a minor key or when their contours are recognized in spite of distorted interval sizes (Dowling & Hollombe, 1977).

That memory for exact interval sizes of familiar melodies is so good raises the question of how such melodies are stored. They cannot be stored simply as contour plus scale, since that would leave the skips unspecified. "Twinkle, Twinkle" and the "Andante" theme from Haydn's *Surprise Symphony* would be stored nearly identically: the contour [0, +, 0, +, 0, −, ±, 0, −, 0, −, 0, −] (where 0 = unison, + = up, and − = down) beginning on the tonic and in the major mode. What is needed is a way of specifying skips along the diatonic scale. What I will suggest is a three-valued dimension of quasi-linguistic marking. Intervals of one diatonic step will be unmarked; two-step intervals will be marked "s" for skip; and larger leaps will be marked "l*x*," where *x* gives the number of steps. In this system, "Twinkle, Twinkle" becomes [0, $+_{l4}$, 0, +, 0, −, −, 0, −, 0, −, 0, −], while Haydn's "Andante" becomes [0, $+_s$, 0, $+_s$, 0, $-_s$, +, 0, $-_s$, 0, $-_s$, 0, $-_s$]. This system takes advantage of the fact that relatively narrow intervals predominate in the music of the world (Dowling, 1968). Of the 26 intervals in the first two phrases of the two songs just cited, 12 are unisons (0 steps), 7 are one step (unmarked), 6 are two steps (marked "s"), and 1 is four steps (marked "l4").

In order to document this predominance of small diatonic intervals more fully and to show that using the present system of marking would put considerably less load on memory than would remembering literally every interval size, I counted all the intervals in the collection of 80 Appalachian songs by Sharp and Karpeles (1968). I chose this collection because it was compiled by careful scholars from an almost exclusively vocal tradition that does not use harmonizing accompaniments. The collection is just about the right size to produce stable data, and the songs in it were selected without any obvious biases regarding the phenomenon under investigation. I first categorized the 80 songs into those using heptatonic modes (seven notes to the octave; 46 songs) and those using pentatonic modes (five notes to the octave; 34 songs). For convenience, and because the existence of a system of hexatonic modes is not indisputable, I treated the 21 songs that used only six notes of the scale as heptatonic (an example of such a song is "Twinkle, Twinkle"). This decision would if anything bias the interval count against small diatonic intervals, since a hexatonic step across the note omitted from the heptatonic mode was counted as a skip. (It is interesting that the note omitted from the heptatonic mode was always one of the pair of notes in the underlying tuning system that stands in the unique 6-semitone interval to each other —B or F in the white-note tuning system— and that the 6-semitone interval did not occur in any of the heptatonic songs.) A subsequent

Table 2
Percentages of Diatonic Intervals of Different Sizes in the 80 Songs of Sharp and Karpeles (1968), with Unisons Omitted

Mode type	Interval size in diatonic steps 1	2	3	4	>4
Heptatonic (46 songs, 1,544 intervals)	56	30	8	3	2
Pentatonic (34 songs, 1,334 intervals)	81	15	2	1	1

check on the data showed that separating the hexatonic songs from the heptatonic has a negligible effect on the frequency distribution of interval sizes.

The songs were strophic. That is, more or less the same melody was repeated for the several verses. For present purposes, each interval in one strophe of each song was counted once. Where the rhythm differed from strophe to strophe (because of different words), I used the rhythm of the first strophe.

Unisons accounted for 23% and 25% of the intervals in the heptatonic and pentatonic modes, respectively. Taking unisons and unmarked diatonic steps together accounted for 66% and 86% of all the heptatonic and pentatonic intervals. Omitting unisons from consideration allows us to see what percentages of + and − intervals need to be marked "s" or "l." Table 2 shows these data. For both mode types, the unmarked intervals account for over half the cases. Only 13% of heptatonic and 4% of the pentatonic intervals need to be coded as leaps.

Moreover, the leaps are used in generally predictable ways in the songs. Of 202 leaps of three or more steps in the heptatonic songs, 106 (52%) occurred in one of the following two contexts: (a) pick-up notes at the start of a phrase or leaps to the start of a new phrase (as between the second and third phrases of "Twinkle, Twinkle"; 75 instances) and (b) passages in which the same interval and rhythmic pattern including the leap was repeated in the same song (as in the first and penultimate phrases of "Twinkle, Twinkle"; 31 instances). The occurrence of leaps is also redundant with mode. Among the pentatonic modes, for example, the mode whose ascending intervals in semitones are [2, 3, 2, 2, 3] had only 1% leaps in its songs versus 4% for the other modes. And each mode has its own pattern of more and less probable leaps. For example, no leaps were observed between the seventh and fourth degrees of the major mode (the 6-semitone interval).

The scarcity of leaps and their redundant usage when they do occur reduces considerably the load on memory. Therefore, it seems plausible to characterize the memory storage of the pitch material of an actual melody as a combination of mode and contour, including a specification of the starting pitch level in the mode and the marking of skips and leaps. It is important to note that this analysis is restricted to the pitch material. With real melodies, rhythm must play an important part in memory. Kolinski (1969) in his analysis of variants of the song "Barbara Allen" provides numerous instances in which the very same contour and mode, with a change in rhythm, become a different song, often in another culture. For example, a change in rhythm turns one of the variants into the sextet from *Lucia di Lammermoor*. In fact, it is artificial even to think of rhythmless melodies. In the state of our knowledge, this is a necessary artificiality, since the problem becomes enormously complex when rhythm is added.

Thinking

Rather than just holding a melodic phrase in short-term memory, subjects in Dowling's (1972) study on melodic transformations had to turn it upside down, backwards, or both. Recognition follows that order of increasing difficulty. Of interest here is the fact that subjects were able only to recognize the melodic contour when so transformed and not the exact interval sizes.

Long-term Memory

In addition to being preserved in short-term memory, melodic contours also seem to be retrievable from long-term memory independently of interval sizes. White (1960) distorted familiar melodies such as "Yankee Doodle" and "On Top of Old Smokey" by changing the pitch intervals between notes by doubling them, adding a semitone, and so on. While the undistorted melodies were recognized 94% of the time, melodies subjected to these contour-preserving transformations were recognized at about the 80% level. Other contour-preserving transformations such as setting all intervals to 1 semitone produced a greater decrement in performance (to around the 65% level). However, contour-destroying transformations produced even greater decrements. For example, randomizing the intervals or subtracting 2 semitones from them even when this changed their sign produced performance in the 45–50% range. (It is difficult because of sampling problems to say what the chance level in White's task was, though it was undoubtedly in the 10–20% range. The rhythmic patterns of the tunes presented alone on one pitch were recognized with 33% accuracy.)

Dowling and Fujitani (1971, Experiment 2) refined somewhat White's approach using five familiar tunes whose first two phrases could be stylized into the same rhythmic pattern. They used three kinds of distortion: preservation of contour, preservation of contour plus relative interval size (i.e., the greater-than/less-than relationships of successive intervals were preserved), and preservation of the underlying harmonic structure of the melody while destroying contour. The contour-preserving distortion produced a decrement in performance from 99% (undistorted) down to 59%. Preserving the relative interval sizes only boosted this a little (to 66%). Destroying contour led to a performance of 28%. (Chance level was probably between 20% and 30%).

Melodic contours of familiar tunes can be recognized even though the distortion of interval sizes is very great. Deutsch (1972) showed that distorting a tune by adding or subtracting one or two octaves from each pitch made it very difficult to recognize. But if the octave distortion is carried out with a preservation of contour, performance is remarkably improved (Dowling & Hollombe, 1977).

One final consideration is that long-term memory has the function in musically nonliterate societies of preserving the musical culture, including the melodies. What we would expect from the present theory is that variants of a tune would share certain similarities, namely, those that arise from very similar contours being hung on the underlying scale framework in various ways. This variation might occur for three principal reasons: forgetting of the original intervals, the desire to create interesting innovations by manipulating interval relationships (as shown in Figure 1), and changes of instruments and scale systems that necessitate transformations of melodies (as in the case of adapting a non-Western melody to a Western scale). All of these processes are probably operating, for example, in the variations of the tune "Barbara Allen" studied by Seeger (1966) and Kolinski (1969). The most striking result of Seeger's study of 76 variants of "Barbara Allen" in the U.S. Library of Congress is that they fall into families based on the sharing of closely related contours. These contours fall onto their scales in various ways, producing a profusion of similar but interestingly varied melodies.

Reference Note

1. 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.* Manuscript in preparation.

References

Adams, C. R. Melodic contour typology. *Ethnomusicology*, 1976, *20*, 179–215.

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

Burns, E. M. Octave adjustment by non-Western musicians. *Journal of the Acoustical Society of America*, 1974, *56*, S25. (Abstract)

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

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

Deutsch, D. Memory and attention in music. In M. Critchley & R. A. Henson (Eds.), *Music and the brain*. London: Heinemann, 1977.

Dowling, W. J. *Rhythmic fission and the perceptual organization of tone sequences*. Unpublished doctoral dissertation, Harvard University, 1968.

Dowling, W. J. Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Perception & Psychophysics*, 1972, *12*, 417–421.

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.

Francès, R. *La perception de la musique*. Paris: Vrin, 1958.

Helmholtz, H. von *On the sensations of tone*. New York: Dover, 1954.

Hood, M. *Sléndro* and *pélog* redefined. *Selected Reports, UCLA Institute of Ethnomusicology*, 1966, *1*(1), 28–48.

Hood, M. *The ethnomusicologist*. New York: McGraw-Hill, 1971.

Imberty, M. *L'acquisition des structures tonales chez l'enfant*. Paris: Klincksieck, 1969.

Kolinski, M. "Barbara Allen": Tonal versus melodic structure, part II. *Ethnomusicology*, 1969, *13*, 1–73.

Kolinski, M. Review of *Ethnomusicology of the Flathead Indians* by Alan P. Merriam. *Ethnomusicology*, 1970, *14*, 77–99.

Martin, D. W., & Ward, W. D. Subjective evaluation of musical scale temperament in pianos. *Journal of the Acoustical Society of America*, 1961, *33*, 582–585.

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

Null, C. H. *Symmetry in judgments of musical pitch*. Unpublished doctoral dissertation, Michigan State University, 1974.

Plomp, R., & Levelt, W. J. M. Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 1965, *38*, 548–560.

Seeger, C. Versions and variants of the tunes of "Barbara Allen." *Selected Reports, UCLA Institute of Ethnomusicology*, 1966, *1*(1), 120–167.

Sharp, C., & Karpeles, M. *Eighty English folk songs from the southern Appalachians*. Cambridge, Mass.: MIT Press, 1968.

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

Sundberg, J. E. F., & Lindqvist, J. Musical octaves and pitch. *Journal of the Acoustical Society of America*, 1973, *54*, 922–929.

Surjodiningrat, W., Sudarjana, P. J., & Susanto, A. *Penjelidikan dalam pengukuran nada gamelan gamelan djawa terkemuka di Jogjakarta dan Surakarta*. Jogjakarta, Indonesia: Universitas Gadjag Mada, 1969.

Swets, J. A. The relative operating characteristic in psychology. *Science*, 1973, *182*, 990–1000.

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

Ward, W. D. Musical perception. In J. V. Tobias (Ed.), *Foundations of modern auditory theory* (Vol. 1). New York: Academic Press, 1970.

White, B. W. Recognition of distorted melodies. *American Journal of Psychology*, 1960, *73*, 100–107.

Williamson, M. C., & Williamson, R. C. The construction and decoration of one Burmese harp. *Selected Reports, UCLA Institute of Ethnomusicology*, 1968, *1*(2), 45–76.

Zenatti, A. Le développement génétique de la perception musicale. *Monographies Francaises de Psychology*, 1969, No. 17.

Received September 19, 1977 ■