

The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour

W. J. DOWLING

Program in Psychology and Human Development, University of Texas at Dallas, Richardson, Texas 75080

and

AVA W. HOLLOMBE

Valley Psychiatric Medical Clinic, Studio City, California 91604

Splitting up the familiar tune "Yankee Doodle" by placing successive notes in different octaves makes it very difficult to recognize. The present experiments generalize this finding to more than one familiar tune and to presentation by actual musical instruments. Tunes differ in the degree to which their recognizability is affected by scrambling into different octaves, but recognition of all 10 tunes investigated was substantially diminished by the distortion. Increasing pitch proximity of temporally adjacent notes by leaving pairs or triples of successive notes intact within the same octave increases recognizability of at least some scrambled tunes. Leaving the melodic contour (the pattern of ups and downs) intact while splitting up the melody increases recognizability for subjects informed of the preservation of contour, though recognition is still worse than for undistorted versions.

Deutsch (1972) found that virtually every one of her subjects could recognize the tune "Yankee Doodle" as usually presented. However, these same subjects had great difficulty recognizing the tune when its notes were divided up randomly into three octaves so that no two successive notes fell within the same octave. Figure 1A shows "Yankee Doodle" in its usual form. Figure 1B shows the same tune with successive notes in different octaves substituted for the original notes. The substituted notes are said to be equivalent in terms of their "chroma" determined by their position in the musical scale (designated by their note names "G, G, A, B, G, etc." or "Do, Do, Re, Mi, Do, etc."). This notion of "chromatic equivalence" of tones an octave apart has, for several reasons, proved a useful construct in describing musical behavior. People are generally very precise and consistent in making octave judgments (Ward, 1954). Tones an octave apart have very similar disruptive effects on memory for pitch when used as interference tones in recognition memory experiments (Deutsch, 1973). Comparison melodies in short-term recognition memory experiments are more easily recognized when transposed exactly an octave from the standard than when transposed some other amount (Francès, 1958). Persons having

absolute pitch, or the ability to name arbitrarily selected notes, often make errors not of note name but of octave level (Ward, 1963). Tones an octave apart are functionally equivalent in Western harmony. And finally, even white rats generalize more readily to the octave than to intervening pitches (Blackwell & Schlosberg, 1943). In spite of the importance of chromatic equivalence in certain types of behavior, subjects do not seem to be able to take advantage of it in recognizing tunes whose notes have been scrambled into several octaves. This may be because even at relatively slow presentation rates listeners are not able to follow the unusually large leaps in pitch of the scrambled version. Practically all the cultures of the world construct melodies out of pitch intervals of the order of a few semitones (Dowling & Fujitani, 1971). (A semitone is 1/12th of an octave in binary log units of frequency.) And people are generally very precise in reproducing well-known musical scale intervals within the octave (Attneave & Olson, 1971). On the other hand, Bregman and Campbell (1971) showed that recognition of identity between three-note melodies with very wide pitch skips in them is worse than for melodies with relatively narrow skips. The processing of melodic information seems particularly difficult when wide changes of pitch are involved.

The present experiments are aimed at extending Deutsch's (1972) finding to other tunes besides "Yankee Doodle" and other timbres besides the sine-

We thank Julie Thompson for carrying out Experiment 1, and Ted Allen, Roberta Greer, and Kelyn Roberts for helpful suggestions and extremely helpful encouragement.

wave oscillator. It is possible that the presence of harmonics at the octave and double-octave of the fundamental frequencies of the lower tones might make the task considerably easier, since all the necessary stimulus information to solve the task would be present in the highest octave. That this is a real possibility is suggested by van Noorden's (1975) demonstration in which he rapidly alternates a pure tone and a complex tone, with the pure tone near to the fundamental, and second and third harmonics of the complex tone in frequency. The perceptual effect is that of a warbling pure tone superimposed on a beeping complex tone. The upper harmonics are heard as part of a separate perceptual figure when they are near in pitch to the alternating pure tone. If the auditory system automatically separates out upper harmonics to create a melodic figure in van Noorden's demonstration, it would seem plausible that it might be able to do a similar kind of thing in deciphering octave-scrambled melodies.

The present experiments also explore the degree to which leaving certain features of the stimuli intact while scrambling them into separate octaves can serve to make them more recognizable. Experiment 1 determines the effect of increasing pitch proximity of temporally adjacent notes. Three successive pairs or triplets of notes remain in the same octave in the midst of the tune (Figure 1C and D). If splitting up a tune so that all pitch intervals are wide has a disruptive effect on recognition, providing fragments of the tune with its usual narrow intervals in the midst of the wide intervals should provide additional cues to the tune's identity. "Yankee Doodle" should be more recognizable in the versions of Figure 1C and D than in the version of Figure 1B. Experiment 1 tests this notion with three tunes: "Yankee Doodle," "Twinkle, Twinkle, Little Star," and "Oh, Susanna." The tunes were selected for their familiarity and because, with very slight changes, their rhythms could be stylized so as to make them very similar to each other rhythmically (Dowling & Fujitani, 1971). Both Deutsch (1972) and White (1960) found that subjects recognize familiar melodies such as these at ostensibly better than chance levels simply on the basis of their rhythmic patterns alone. Rhythmic similarity was introduced here to lower that baseline rate of recognition on the basis of rhythm.

Melodic contour in the sense of the successive ups and downs of the melodic line as been shown to be very important in the recognition of melodies. The musical styles of quite diverse cultures utilize the device of changing the pitch relationships of a melody while leaving its contour intact (see, for example, Kolinski, 1970, on the American Indians). Dowling and Fujitani (1971) found that listeners were able to identify distorted versions of familiar

tunes with much greater than chance accuracy provided those distortions preserved the melodic contours of the original tunes intact. Dowling and Fujitani's distortions, however, utilized the small melodic interval relationships characteristic of these melodies. Experiment 2 tests whether subjects can use the melodic contour patterns of familiar tunes which have been distorted so that no two successive notes fall in the same octave—so that all the intervals are very wide. The subjects in the condition of Experiment 2 in which melodic contour was preserved as a cue were thus informed of that fact. Figure 1F shows "Yankee Doodle" as presented in the contour-preserving condition of Experiment 2. Experiment 2 includes 10 melodies, selected for their presumed familiarity, in order to extend even further the generalizability of Deutsch's (1972) effect.

EXPERIMENT 1

Method

Subjects. Thirty-seven UCLA undergraduates served in group sessions of four or five subjects each for course credit in Introductory Psychology. There were 8 to 10 subjects in each of the four conditions of Experiment 1. Subjects in Experiments 1 and 2 were found to be representative of a subject population that in a large number of experiments has been found to have a mean of approximately 3 years of musical training (including playing in an ensemble and lessons on an instrument or voice, but not including singing in choirs or music appreciation classes) with a median of approximately 1 year. The subjects were assigned to conditions blindly.

Stimuli. The stimuli were produced on soprano and tenor recorders having a range of fundamental frequencies from 262 Hz (middle C) up 3 octaves to 2,096 Hz. The stimuli were tape recorded and played to subjects over high-quality reproduction equipment through loudspeakers. Timing was controlled during recording by a Davis timer which was barely audible on the recording. The stimuli consisted of the first two phrases (the first 16 quarter-note beats) of the melodies "Yankee Doodle," "Oh, Susanna," and "Twinkle, Twinkle, Little Star," and were presented at a rate of 2 beats/sec. Thus each stimulus was 8 sec in duration. The tunes began on G, E, and C, respectively. The rhythmic patterns of the tunes were slightly stylized (as in Figure 1A) to increase their similarity to each other. All the tunes had exactly the same rhythmic pattern, except for the last measure (the last four beats), which contained two notes, one note, or three notes, respectively.

There were four conditions in Experiment 1. In Condition A, the subjects heard the tunes each presented within 1 octave, as in Figure 1A. The octave used for each tune was randomly selected. In Condition B, the subjects heard the tunes split up into 3 octaves so that no two successive notes fell within any 1 octave, as in Figure 1B. The distribution of notes in the 3 octaves was randomly determined and was the same for all subjects in each of Conditions B, C, and D. In Condition C, the tunes were split up as in Condition B, but with the constraint that three randomly selected pairs of successive notes in each tune fall within the same octave (Figure 1C). Condition D was the same as Condition C except that, instead of three pairs, two groups of three notes were randomly selected to fall within the same octave (Figure 1D).

Procedure. In each condition, the subjects heard the three tunes in the same random order. The experimenter instructed the

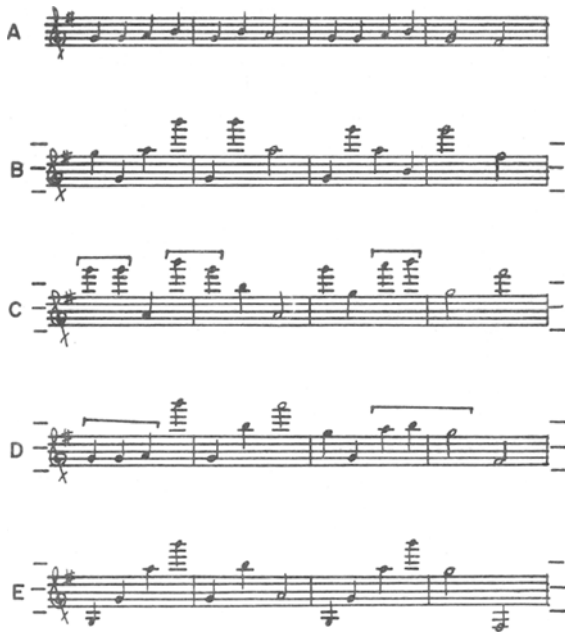


Figure 1. "Yankee Doodle" as presented in Experiment 1: Condition A, undistorted; Condition B, in randomized octaves; Condition C, with pairs of notes preserved in the same octaves (brackets); Condition D, with triples of notes preserved in the same octaves (brackets). Part of "Yankee Doodle," as presented in Experiment 2, is shown in E, which preserves the contour of the original melody. The marks at both ends of the staves indicate the boundaries of the octaves to which notes were assigned, and fall at $F_0 = 262, 524, \text{ and } 1,048 \text{ Hz}$.

subjects to respond after each tune by writing down the name of the tune, or some of the words, or a characterization of the tune if they could not remember the name or words. (In fact, all subjects responded with the name, or words, or "Don't know.") The subjects were given as much time as they needed to respond. After responding to the three stimuli for their particular condition, subjects in Conditions B, C, and D responded to all of the other conditions, ending with the undistorted tunes of Condition A. Data from these subsequent responses to distorted melodies were not used in the analysis, however, as it is reasonable to suppose that they would be heavily influenced by previous responses. Responses of subjects in Conditions B, C, and D to the undistorted tunes were virtually the same as those of subjects in Condition A, and served as a check on whether the subjects could recognize undistorted versions of tunes they could not recognize when distorted. At the end of the session, the subjects completed a brief musical autobiography.

Results

The results of Experiment 1 are shown in Table 1. The data for "Yankee Doodle" and "Oh, Susanna" in Condition B are comparable to the 12% correct obtained by Deutsch (1972) for "Yankee Doodle." "Twinkle, Twinkle," however, would seem to be easier than the other two. To test for the dependence of responses on conditions, separate chi squares were run on the rows of Table 1. For the tunes "Yankee Doodle" and "Oh, Susanna," the conditions differed significantly from each other ($p < .01$). Conditions C and D, in which the separation of the tunes into different octaves was weakened by the

introduction of groups of successive notes in the same octave, are easier than Condition B for those tunes.

Experiment 1 demonstrates the difference in results obtained when different tunes are presented in octave-scrambled form. Experiment 2 carries this further by making more precise observations with 10 different tunes.

EXPERIMENT 2

Method

Subjects. Thirty-four subjects, as described under Experiment 1, served in group sessions of three to five subjects each with four sessions in each condition. The subjects were assigned blindly to one of the two conditions of Experiment 2, with 17 subjects in each group.

Stimuli. The stimuli were produced on the piano using a 4-octave range from a fundamental frequency of 131 Hz (C below middle C) to 2,096 Hz. Four octaves were necessary to preserve the contour relationships of Condition E. The stimuli were timed, recorded, and presented as in Experiment 1. The stimuli consisted of the first 32 quarter-note beats of the 10 melodies listed in Table 2, presented at a rate of 1.33 beats/sec, and so each stimulus was 24 sec long. Familiar rhythms of the of the tunes were retained since it would have been impossible to stylize all 10 melodies to produce similarity. Condition B was comparable to Condition B of Experiment 1 and to the manipulation of Deutsch's (1972) experiment. The notes of each melody were randomly assigned to the 4 octaves so that no two successive notes fell in the same octave. In Condition E, notes were similarly assigned to octaves but with the constraint that the melodic contour, the ups and downs from note to note of the melody, would be preserved. That is, where the original melody went up by a small melodic interval, the distorted version would go up by that same interval plus 1, 2, or 3 octaves. Where the original melody went down, the distorted version went down, again by more than an octave. Wherever possible, the size of the multiple octave increment was randomly determined. If a note was repeated in the original, the distorted version went up or down by exactly 1 octave, the direction being randomly determined within the constraints of the pitch range. Series of repeated notes in the original were replaced by series of 1-octave skips in the same direction, up or down. Where this procedure took the distorted version beyond the pitch range used, as in "Jingle Bells," a skip was made to the other end of the pitch range and the series continued in the same direction in which it had started. Figure 1E shows the first part of "Yankee Doodle" distorted according to the constraints of Condition E. A different random order of presentation of the stimuli was used for each of the four groups in each condition.

Procedure. The subjects responded as in Experiment 1. The subjects in Condition E were informed that the distorted melodies preserved the contours, the patterns of ups and downs, of the original melodies. After responding to the distorted versions

Table 1
Percent Correct Identifications of Tunes in Experiment 1

Tune	Condition			
	A	B	C	D
		Random	Pairs	Triples
Yankee Doodle	88	10	44	70
Oh Susanna	88	10	33	70
Twinkle, Twinkle	100	50	56	50
N	8	10	9	10

Table 2
Percent Correct Identification of Tunes in Experiment 2
Including Subsequent Identification of Undistorted Versions

Tune	Conditions		Mean of B & E	Undis- torted (both groups)
	B	E		
When the Saints Go Marching In	0	0	0	65
Battle Hymn of the Republic	0	24	12	79
Jingle Bells	12	88	50	100
Yankee Doodle	29	71	50	97
Old MacDonald	41	59	50	94
Oh Susanna	47	53	50	97
I've Been Working on the Railroad	24	94	59	97
Mary Had a Little Lamb	53	82	68	94
Twinkle, Twinkle, Little Star	53	94	74	97
Three Blind Mice	88	94	91	97
Mean	35	66	51	92

of the melodies, all subjects responded to the original melodies to make sure that the melodies were as familiar as the experimenters supposed. At the end of the session, the subjects completed a brief musical autobiography.

Results

Table 2 summarizes the results of Experiment 2. The conditions differed from each other significantly [$F(1,32) = 24.38, p < .01$]. Condition E, in which contour information from the tunes were available, was easier than Condition B, in which it was not. Performance on Condition E was still substantially worse than with the undistorted versions for all but three of the tunes, confirming Deutsch's (1972) result. The tunes differed significantly [$F(9,288) = 19.15, p < .01$], confirming the suggestions found in the data of Experiment 1. Table 2 shows the tunes ordered according to overall difficulty. The Tunes by Conditions interaction was significant [$F(9,288) = 4.62, p < .01$], but due to the nature of the data cannot be interpreted beyond the suggestion that different tunes are affected differently by these distortions. Table 2 also gives the rates of correct response to the subsequent presentation of the undistorted melodies for the two conditions combined. Both conditions were very comparable in this, with somewhat more errors from subjects in Condition E. Since this difference runs counter to the main effect observed, it cannot be used to explain the results of the experiment.

Correlations between years of musical training and recognition performance were .38 for Condition B and $-.19$ for Condition E. This is consistent with the results of Dowling and Fujitani (1971), who found that short-term recognition memory for melodic contours had low negative correlations with musical training.

DISCUSSION

Breaking up a tune so that successive notes fall in different octaves does, in fact, diminish the recognizability of all the familiar tunes investigated here, establishing the generality of Deutsch's (1972) effect. Experiment 2 showed that this lowering of recognition rates varied from the extremes of "When the Saints Go Marching In," which went from 65% in the normal version to completely unrecognizable when distorted, to "Three Blind Mice," which only went from 97% to 88% when scrambled (Table 2). The recognition rates for "Yankee Doodle" in the randomly scrambled conditions of Experiments 1 and 2 were 10% and 29%, which, given all the differences in method, seem comparable to the 12% obtained by Deutsch (1972). The effect does not depend on presentation of the tunes in pure tones as in Deutsch's study. Just as with most effects involving patterns of pitches [as, for example, Miller and Heise's (1950) "trill threshold," see Dowling (1973)], presentation via complex tones does not alter the effect appreciably. The representation of pitch in the nervous system must be abstracted from the auditory stimulus to such a degree that harmonic content of the tones has little effect on the processing of pitch information (cf. Terhardt, 1974).

We view the difficulty subjects have in recognizing melodies split into several octaves as one more example of the difficulty listeners have in achieving coherent perceptual organization when the stimulus contains wide leaps of pitch. We see the present phenomena as lying on a continuum with those involving perceptual splitting of rapid pitch sequences alternating between widely separated pitch regions. (Bregman & Campbell, 1971; Dowling, 1973). However, this is not the only possible way of viewing these phenomena. Deutsch (1973), for example, sees the difficulty of recognizing octave-scrambled tunes as requiring a different theoretical explanation. As she points out, the scrambled-melody effect can occur at very slow rates of presentation—much slower than those used in melodic-fission experiments with interleaved melodies. Further experiments are needed to settle between these points of view. The results obtained may very well depend on the subject's set. Van Noorden (1975) found that at slow presentation rates the perceptual continuity of trills was markedly affected by set. Subjects with a set for hearing a continuous pattern heard that with much wider intervals between the notes of the trill than did subjects with a set for separation. At slow presentation rates, the listener can more easily impose a perceptual organization of melodic continuity on a stimulus containing wide skips. For example, the opening Kyrie of Bach's *Mass in B*

Minor moves slowly (ca 2 notes/sec) and presents a series of pairs of tones progressively diverging in pitch. One usually listens to it with a set for hearing it as a continuous melody. It is possible, however, to change set and hear it as an alternation between diverging melodic lines.

The distortions of melodies in the present experiments violate the gestalt principle of proximity in figural organization (Hochberg, 1971). In Condition B of Experiment 1, the individual notes of the melodies lack pitch proximity to temporally adjacent notes. Conditions C and D provide better proximity by preserving pairs and triples of notes out of the original melodic lines, increasing recognition rates in two out of three cases. This lends partial support to the notion that preserving some of the familiar proximity of notes in a melody can aid recognition.

Since the octave scrambling of melodies produces such a dramatic perceptual effect, it is not surprising that composers have used similar techniques when they wanted to break down the listener's reliance on continuous melodic lines. Webern (1883-1945), in particular, developed a style of dividing up pitch material so that successive notes would not only fall into different octaves but would be played on instruments with different timbre. His *Symphonie, Op. 21* (1929) is a good example of this. In the opening measures, single tones and pairs of tones are presented at a rate of about 1 tone/sec, with intervals between successive tone of the order of an octave. As Grout (1960) says, "The result is a texture made up of sparks and flashes of sound blending in a unique balance of color" (p. 656).

REFERENCES

- ATTNEAVE, F., & OLSON, R. K. Pitch as medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, 84, 147-166.
- BLACKWELL, H. R., & SCHLOSBERG, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, 33, 407-419.
- 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.
- DEUTSCH, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, 11, 411-412.
- DEUTSCH, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, 1973, 13, 273-275.
- DOWLING, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, 5, 322-337.
- DOWLING, W. J., & FUJITANI, D. S. Contour, interval, and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America*, 1971, 49, 524-531.
- FRANCÈS, R. *La perception de la musique*. Paris: Vrin, 1958.
- GROUT, D. J. *A history of western music*. New York: Norton, 1960.
- HOCHBERG, J. Perception: I. Color and shape. In J. W. Kling & L. A. Riggs (Eds.), *Experimental psychology* (3rd ed.). New York: Holt, Rinehart, & Winston, 1971. Pp. 395-474.
- KOLINSKI, M. Review of A. P. Merriam *Ethnomusicology of the Flathead Indians*. *Ethnomusicology*, 1970, 14, 77-99.
- MILLER, G. A., & HEISE, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, 22, 637-638.
- NOORDEN, L. VAN. Rhythmic fission as a function of tone rate. Institute for Perception Research, *Annual Progress Report*, 1971, 6, 9-12.
- NOORDEN, L. VAN. *Temporal coherence in the perception of tone sequences*. Eindhoven, The Netherlands: Institute for Perception Research, 1975.
- TERHARDT, E. Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, 1974, 55, 1061-1069.
- WARD, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, 26, 369-380.
- WARD, W. D. Absolute pitch. *Sound*, 1963, 2(3), 14-21, & 2(4), 33-41.
- WEBERN, A. *Symphonie, Op. 21*. Vienna: Universal Edition, 1929.
- WHITE, B. Recognition of distorted melodies. *American Journal of Psychology*, 1960, 73, 100-107.

(Received for publication April 5, 1976;
revision received November 5, 1976.)