Perception & Psychophysics
1978, 24 (6), 551-565

A bidimensional model of pitch in the recognition of melodies

WENDY L. IDSON
University of Texas at Austin, Austin, Texas 78712

and

DOMINIC W. MASSARO
University of Wisconsin at Madison, Madison, Wisconsin 53706

Pitch can be conceptualized as a bidimensional quantity, reflecting both the overall pitch level of a tone (tone height) and its position in the octave (tone chroma). Though such a conceptualization has been well supported for perception of a single tone, it has been argued that the dimension of tone chroma is irrelevant in melodic perception. In the current study, melodies were subjected to structural transformations designed to evaluate the effects of interval magnitude, contour, tone height, and tone chroma. In two transformations, the component tones of a melody were displaced by octave intervals, either preserving or violating the pattern of changes in pitch direction (melodic contour). Replicating previous work, when contour was violated perception of the melody was severely disrupted. In contrast, when contour was preserved the melodies were identified as accurately as the untransformed melodies. In other transformations, a variety of forms of contour information were preserved, while eliminating information for absolute pitch and interval magnitude. The level of performance on all such transformations fell between the levels observed in the other two conditions. These results suggest that the bidimensional model of pitch is applicable to recognition of melodies as well as single tones. Moreover, the results argue that contour, as well as interval magnitude, is providing essential information for melodic perception.

Psychophysical investigations have focused upon developing unidimensional psychological scales of pitch as a monotonically increasing function of the unidimensional physical scale of frequency. Though such scales have been developed with reasonable success (e.g., Stevens & Volkmann, 1940; Stevens, Volkmann, & Newmann, 1937), there is a body of research to suggest that pitch perception may be more complex than a simple monotonic scale would indicate (Attneave & Olson, 1971; Bachem, 1954; Baird, 1971; Deutsch, 1972b, 1973a, 1974; Shepard, 1964). Rather, it has been proposed that pitch can be more appropriately conceptualized as a bidimensional quantity representing both the overall pitch level of the tone (tone height) and its position in the octave (tone chroma) (Bachem, 1948; Deutsch, 1969; Meyer, 1904; Révész, 1954; Ruckmick, 1929; Shepard, 1964; Drobish, cited in Ruckmick, 1929). Most bidimensional descriptions share the assumptions embodied in the original model proposed by Drobisch (1846, cited in Ruckmick, 1929). Figure 1 shows a graphical representation of this model, in which pitch is conceptualized as a helix, with tone height represented by the vertical axis of the helix and tone chroma by the circular scale at its base. Tones separated by an integer number of octaves (a 2:1 frequency ratio) are located at corresponding points on successive turns of the helix.

A bidimensional model is implicit in the traditional musical scale, which is based on the underlying similarity of tones which stand in an octave relationship. Such tones share the same name (e.g., $A_4 = 435$ Hz, $A_5 = 870$ Hz). and in some cases are treated as harmonically equivalent (Tovey, 1956). For example, if the component tones of a chord are displaced into different octaves, this new chord will be treated as harmonically equivalent to the original chord. These similarities are not limited to Western music, the

This research was supported by U.S. Public Health Service Grant MH-19399 to D. W. Massaro, and was conducted while W. L. Idson was supported first by a University of Wisconsin Graduate Traineeship and later by a University of Wisconsin Graduate Fellowship. The work represents a collaborative effort; order of authorship is arbitrary. We would like to thank David E. Klitzke for his assistance with the computer programming. Portions of this work were discussed at the Meeting of the Acoustical Society of America, San Diego, 1976, and at the Meeting of the Midwestern Psychological Association, Chicago, 1977. Requests for reprints may be sent to either author: Wendy L. Idson at the Department of Psychology, University of Texas at Austin, Austin, Texas 78712; or Dominic W. Massaro at the Department of Psychology, University of Wisconsin, Madison, Wisconsin 53706.

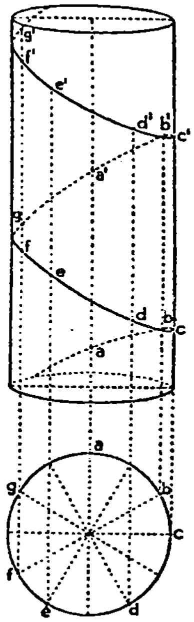

Figure 1. Graphic representation of a bidimensional model (taken from Révész, 1954).

octave relationship being fundamental in music from a variety of cultures (Nettle, 1956). This musical equivalence appears to be reflected at the perceptual level, as well. Both Bachem (1954) and Baird (1971) have found that while subjects with absolute pitch can almost invariably name single tones correctly, they are often unable to place them in the appropriate octave. For example, a subject might be aware that a particular tone was C, but would be unable to distinguish whether it was C_4, C_5, or C_6.

If a bidimensional model is appropriate for pitch perception, then two tones which differ by an exact octave should be perceived as more similar than two tones which differ by less than an octave, despite the fact that the frequency separation is greater in the former case. Allen (1967) reported this result for the subjective ratings of tone pairs. Moreover, early conditioning work demonstrated enhanced generalization effects, both in rats (Blackwell & Schlosberg, 1943) and humans (Humphreys, 1939), for tones at octave intervals. The most interesting evidence comes from a study by Shepard (1964). He reasoned that if one of the components of pitch was circular, then it should be possible to eliminate transitivity in pitch judgments by suppressing the other component. To eliminate the dimension of tone height, Shepard generated a set of 12 complex tones, each composed of pure tones separated by octave intervals. When the complex tones were sounded up (or down) the scale in half-note steps, at a rate of 1 tone/sec, the tones were heard as constantly ascending (or descending) in pitch. A series of complex tones, whose components were not at octave intervals, was heard appropriately as a sequence of chords.

Deutsch (1969) incorporated a bidimensional model of pitch into a more general model designed to account for the recognition of music. In Deutsch's model, the pitch of a tone is first stored in a unidimensional array of tone height. Processing then continues along one of two channels. Single tones will be analyzed for tone chroma, along a channel in which a convergence of pitch information occurs from all tones at octave intervals. The tones of a melody will be analyzed along the second channel, in which octave convergence is assumed not to occur. Instead, the pitch intervals between successive tones will be abstracted and synthesized, in accord with a hierarchical neural-network model analogous to that of Hubel and Wiesel (1962).

There are two principal consequences of Deutsch's (1969) model. The first, which would also be true for any bidimensional model, is that perception of single tones should be affected by tone chroma as well as tone height. The second maintains that perception of the tones of a melody should not be influenced by tone chroma. It should be stressed that this latter result is not a necessary consequence of a bidimensional model of pitch, but is unique to Deutsch's (1969) model.

In order to evaluate the first of these consequences, Deutsch (1973a, 1974) investigated the effects of tone chroma on the recognition of single tones. She employed a task in which subjects were to determine whether a comparison tone was the same as or different in pitch from a previously presented standard tone. Earlier work had demonstrated three specific effects in this task, when a series of interference tones—all drawn from a single octave—are interposed between the standard and comparison tones. First, if the sequence of interference tones is chosen randomly from within the octave, performance is decreased relative to a blank interval (Deutsch, 1970). Second, when the standard and comparison tones differ in frequency, including an interference tone which is identical to the comparison will disrupt performance more greatly than a randomly selected tone (Deutsch, 1970, 1972a). Third, an even larger performance decrement is found when the standard and comparison tones are identical in frequency and the sequence of interference tones includes one a semitone higher and another a semitone lower than the standard/comparison (Deutsch, 1973b). In a series of studies in which the critical tone(s) were displaced by octave intervals, Deutsch (1973a, 1974) was able to replicate all three of these effects, though their magnitude was somewhat smaller. These results provide strong support for the bidimensional model. The occurrence of interference from tones drawn from the higher and lower octaves argues for the significance of tone chroma, while the decrement in the magnitude of the interference, relative to interference tones in the same octave, supports the contribution of tone height.

Deutsch (1972b) also attempted to provide support for the second prediction of her model, by demonstrating simultaneously that octave convergence will not occur for the component tones of melodies and that successive interval abstraction provides the basis for melodic perception. If octave convergence occurs for melodies, then displacing the component tones of a melody by octave intervals should not disrupt recognition; the pitch information provided by tone chroma should be sufficient. If, however, melodic perception is based upon successive intervals *and* octave convergence does not occur, perception should be disrupted, as displacing the tones by octave intervals will eliminate tone height and distort successive intervals. To evaluate this hypothesis, Deutsch presented subjects with the first half of the tune *Yankee Doodle*, played either entirely in one of three adjacent octaves or with the component tones displaced by octave intervals, randomly throughout the three adjacent octaves. Although the melody was easily identified when it was played entirely in any one of the three octaves, subjects were completely unable to recognize the melody when the tones were randomly displaced among the three octaves. On the basis of these results, Deutsch (1972b) argued that octave convergence does not occur for melodies and that successive intervals provide the critical information for melodic perception. This latter conclusion is based upon the fact that the octave displacement manipulation primarily distorts information for tone height and successive intervals. Yet tone height does not appear to be critical, as the melody could be recognized when it was played entirely in any of three octaves. Deutsch concluded that the critical information which is eliminated by octave displacement comes from the magnitude of successive pitch intervals.

The intent of the present research is to argue that rejection of a bidimensional model of pitch for the perception of melodies may be premature. The reason that is generally given for rejecting a bidimensional model is that melodies are perceived as a string of successive intervals. If intervals are computed between the absolute frequencies (tone height) of the component tones, then the octave similarities offered by the chroma dimension cannot be useful. For example, the absolute magnitude of the interval C_4 to D_5 is clearly larger than that of C_4 to D_4, minimizing the similarity between them. Yet a bidimensional model can be conceptualized in such a way that it avoids this consequence. If a convergence of pitch information occurs from tones standing in an octave relationship, then it seems tenable that all such tones may be collapsed into a single abstract representation. Thus, there would be one abstract A representing all As, one abstract A# representing all A#s, out to one abstract G# representing all G#s. Perception of a melody could then occur within the resulting abstract octave, represented by the circle in Figure 1. Under such a scheme, the intervals C_4 to D_4 and C_4 to D_5 would be equivalent, both simply representing the 2-semitone interval from C to D. Within such a framework, tone chroma could play as great a role in the perception of melodies as in the perception of single tones. While such a conceptualization is not a necessary correlate of a bidimensional model of pitch, it is certainly a viable one, suggesting that in no sense are the concepts of successive interval abstraction and tone chroma theoretically irreconcilable.

Yet it might be asked whether such a conceptual framework should not predict that subjects would be able to recognize the distorted versions of *Yankee Doodle* in Deutsch's (1972b) study, as all of the tones would converge into a single octave space. However, Deutsch's manipulation effectively eliminated the potential utility of an abstract octave. In displacing the component tones of the melody across octaves, Deutsch made no attempt to preserve the directions of the successive intervals in the original melody. It might be expected that the direction of approximately half of the successive intervals would be in the direction opposite to that found in the original melody. As a result, three sources of relational information would be distorted: successive intervals in terms of tone height, abstract intervals in terms of tone chroma, and the melodic contour as a pattern of pitch changes across a melody. Consider, for example, the interval C_4 to D_4, occurring in the original melody. Transformation of this interval into C_4-D_5 would distort the magnitude of the interval in terms of tone height, but would preserve the abstract interval in terms of tone chroma. The 2-semitone increase from C_4 to D_4 would be transformed into a 14-semitone increase from C_4 to D_5, both of these intervals being represented as an ascending second in the abstract octave. In contrast, transformation of C_4-D_4 into C_4-D_3 would violate not only the absolute interval, but also the abstract interval. The 2-semitone increase from C_4 to D_4 would be transformed into the 10-semitone decrease from C_4 to D_3, which would be represented as a descending seventh in the abstract octave. At the same time, reversing the directions of half of the intervals in the melody would distort the contour of that melody.

In order to provide a fair test of a bidimensional model of pitch in melodic perception, it is necessary to displace the tones of a melody by octave intervals while preserving the melody's contour. Both tone height and absolute interval magnitude would then be violated, while tone chroma and melodic contour would be preserved, maintaining the abstract intervals. However, recognition of a melody could be facilitated by preservation of contour independent of any contribution of tone chroma, as contour alone may provide critical information for perception. This possibility is entirely reasonable from a musical standpoint, as forms such as the fugue often involve transformations which preserve melodic contour but distort interval relations (David & Mendel, 1966). Moreover, empirical support for the role of contour has been obtained. Dowling and Fujitani (1971) found that subjects are quite accurate in identifying distorted melodies which preserve melodic contour, though their accuracy increases when successive intervals are preserved as well. Furthermore, melodic contour alone sometimes provides sufficient information for accurate perception. Dowling (1972), for example, has found that when melodies are subjected to various musical transformations, subjects are no more accurate in recognizing transformations which preserve exact interval relations than transformations which preserve only melodic contour. Similarly, Idson and Massaro (1976) have found that degraded tone patterns which can be uniquely specified by their contours are more accurately identified than patterns for which interval magnitude must also be perceived. Consequently, the contribution of contour alone must also be assessed.

The present research attempted to explore the utility of a bidimensional model of pitch for melodic perception. A melody recognition task was employed, in which the melodies were subjected to a variety of structural transformations designed to preserve or violate selective sources of information. The level of performance on these transformed melodies, with respect to the level of performance on the original untransformed melodies, provides an index of the importance of the manipulated information. To the extent that distortion of a source of information produces a performance decrement, that information can be thought to play a role in melodic perception. Similarly, if a given source of information can be eliminated without disrupting performance, then that information is not necessary for accurate melodic perception.

In the first experiment, five transformations were employed. Transformation O simply entailed playing the original melody, in order to provide a baseline for performance. Transformation OVC* replicated Deutsch's (1972b) manipulation. The component tones of the melody were displaced by octave intervals, violating the melodic contour by reversing the directions of approximately half of the successive intervals. In transformation OPC*, the component tones of the melodies were displaced by octave intervals, but now the melodic contour was preserved. Two additional transformations assessed the independent contribution of melodic contour by preserving the contour but violating all information as to absolute frequency and successive interval magnitude. These two transformations differed in terms of the precision with which the contour was maintained. In transformation PC, only the directions of the successive intervals were preserved. In transformation LT a linear transformation was taken of the contour, which halved the size of each of the component intervals. The resulting melodies would retain both the directions of the successive intervals and the relative magnitude of these changes. The unusually small intervals in these melodies should have no detrimental effects upon performance, as both Werner (cited by White, 1960) and White (1960) have found that once subjects obtain some practice with "microintervals" performance becomes quite accurate. Both of these contour-preserving transformations were included, since melody recognition studies (Dowling & Fujitani, 1971; White, 1960) have shown that subjects are quite sensitive to preservation of relative degrees of contour.

A comparison of performance in identifying melodies under these five types of transformations should make it possible to evaluate the contributions of tone height, tone chroma, and melodic contour in melodic perception. Since transformation OVC* is identical to Deutsch's (1972b) manipulation, performance is expected to be quite poor. If tone chroma is ineffective in melodic perception, then performance on transformation OPC* should be equivalent to that on transformation OVC*. However, if octave convergence does occur under the octave displacement transformation, but perception is nevertheless disrupted due to violation of melodic contour, then performance on transformation OPC* should be superior to transformation OVC*. Moreover, the contributions of octave convergence and melodic contour can be disambiguated by comparing performance on transformation OPC* to that on transformations LT and PC. If any performance advantage seen for transformation OPC* over OVC* is due solely to preservation of melodic contour, then performance should be equally good on transformations OPC* and PC/LT. If further information is being provided by octave convergence, however, transformation OPC* should yield better performance than transformations PC/LT. A comparison of performance on transformations O and OPC* allows an evaluation of the abstract perceptual octave proposed above. If this concept is valid, then performance on transformation OPC* should be comparable to that on transformation O, since the octave displacement manipulation preserving melodic contour maintains all of the intervals, in an abstract sense, which are found in the original melodies. To the extent that performance on transformation O is superior to that on transformation OPC*, absolute interval magnitude, rather than abstract interval identity, is playing an important role. Thus, the current study allows a direct evaluation of a bidimensional model of pitch in melodic recognition.

EXPERIMENT 1

Method

Subjects. The subjects were 25 University of Wisconsin students who received additional credit in an introductory psychology course for their participation. At the start of the experiment, the subjects were asked whether they knew all five melodies. Any subject who reported that he or she did not know one or more of the melodies was not continued in the study.

Apparatus. Four subjects could be tested simultaneously in separate sound-insulated rooms. All experimental events were controlled by a PDP-8L computer. The tonal stimuli were generated by a digitally controlled oscillator (Wavetek Model 155). The output of the oscillator was gated by two computer controlled audio switches (Iconic Model 0137) to separate amplifiers (McIntosh Model MC-50) for each ear. The tones were then presented over matched headphones (Grason-Stadler Model TDH-49). The visual display was presented over a digitally controlled array of light-emitting diodes (Monsanto Model MDA-III).

Design. The experimental design was a 2 by 5 by 5 factorial. The three factors were: the length of the melody, the identity of the melody, and the type of transformation performed upon the melody. The two lengths were either the first 8 or the first 12 notes of the melody. The five melodies were well known songs: *Happy Birthday*, *London Bridge*, *Pop Goes the Weasel*, *On Top of Old Smokey*, and *God Rest Ye Merry Gentlemen*. The five transformations (O, OPC*, OVC*, LT, and PC) involved changes in the frequencies of the component tones of the melodies, as described below.

Stimulus structure. The component tones of the melodies were 50-msec sine waves, having 10-msec rise/fall times. The intertone interval between successive tones was a constant 150-msec interval. No rhythm information was present, and a sustained note was mimed by repeating the tone (see Dowling, 1973). The intensity of each tone was adjusted individually, in accord with the Robinson and Dadson (1956) equal-loudness contours. All of the melodies began on A. For transformations OPC* and OVC*, either A_4 (435 Hz) or A_5 (870 Hz) was used for the starting note. For transformations O, PC, and LT, the melody could be played in any one of the seven octaves (A_1 = 54 Hz to $G\#_7$ = 6,570 Hz) spanned by the melodies in the cross-octave transformations. This insured that any differences which might be found between the within-octave (O, PC, LT) transformations and the cross-octave (OPC*, OVC*) transformations would not reflect simple differences in discriminability for tones in different octaves. This seems particularly important, as the difficulty of judging the direction of frequency shifts has been found to increase markedly for tones above about 4,000-5,000 Hz (Ward, 1970).

One instance of each of the five melodies was constructed for the four distorting transformations (OPC*, PC, LT, OVC*). These 20 melodies and 5 original melodies were used throughout the study, either the first 8 or the first 12 notes being played. The following heuristics were used to construct the melodies. In transformation O, the original melodies were played. In transformations OPC* and OVC*, each tone was replaced by an exact octave multiple, with the replacement tone being drawn from the octave adjacent to that containing the immediately preceding tone. In transformation OPC* the contour of the original melody was preserved, so that the direction of each interval in the transformed melody was identical to the direction of the corresponding interval in the original melody. Both tones of a unison interval were always played in the same octave. In transformation OVC*, the contour of the melody was violated, by reversing the direction of approximately every other interval. Tones of a unison interval were played in different octaves, with half of the second tones shifted up an octave and half shifted down an octave. In addition, an attempt was made to insure that for each of the individual melodies, transformations OPC* and OVC* spanned the same octave range and had approximately an equal distribution of notes in each octave.

In transformations PC and LT, the melodies were played entirely within a single octave, preserving the melodic contour while violating the absolute frequencies and successive interval magnitudes. In transformation PC, the component tones were chosen randomly, subject to the constraint that tones and intervals appearing in the original melody did not appear in the same position in the transformation. In transformation LT, the relative magnitudes of the intervals were preserved, by halving each interval. When the original interval involved an even number of semitones, halving that interval produced a new interval in which the second tone was not a note of the musical scale. The linear transformation insured that tones and intervals in the original melody would not appear in the same position in the transformed melody.

Procedure. The experiment was conducted on 2 consecutive days. Day 1 consisted of three sessions of 110 trials each, separated by 5-min rest breaks. Day 2 consisted of four sessions of 110 trials each. The first 10 trials of each session were treated as unscored practice trials, though the subjects were not aware of this. On each trial, the subjects heard one of the five transformations of the five basic melodies. The subjects were to identify which of the five basic melodies was presented by pressing one of five buttons labeled with the names of the five melodies. The subjects were given a maximum response interval of 8 sec from the onset of the melody. However, if all four subjects responded in less than this 8-sec interval, the experiment continued immediately upon the last response. Following the response interval, a .25-sec visual presentation of a row of three asterisks (***) was given over the light-emitting diodes to indicate the end of the response interval. A 1-sec intertrial interval was employed. All 50 (2 lengths × 5 melodies × 5 transformations) experimental conditions were completely random and occurred equally often within a session.

The same procedure was employed for all four experiments.

Results

The dependent measure is the percentage of correct responses. The percentage correct was calculated for each subject, in each session, at each length by melody by transformation condition. These percentages were then submitted to an analysis of variance on subjects, sessions, lengths, melodies, and transformations. The main effect of sessions was nonsignificant in the analysis ($F < 1$), as were all interactions involving sessions as a term. These results indicate that performance was not improving with increasing practice. Consequently, to increase the reliability of the individual scores, these data were pooled over sessions and reanalyzed. The same data analysis procedures were employed in all four experiments.

The longer melodies were identified approximately 6% more accurately than the shorter melodies [$F(1,24) = 53.52$, $p < .001$]. However, all interactions involving melody length as a factor were nonsignificant in the analysis. This indicates that the length of the melody was affecting only the overall level of performance and was not differentially affecting either the individual melodies or the individual transformations. Consequently, the data will be discussed in terms of the average results for the two melody lengths.

Figure 2 presents the percentage of correct identifications of the melody as a function of the type of transformation performed on the melody. It is apparent from the figure that the type of transformation performed upon the melody had a large effect upon performance [F(4,96) = 85.79, p < .001]. Displacing the component tones of the melodies by octave intervals in accord with the melodic contour (OPC*) yielded performance at essentially the same level as that observed on the original melodies (O). A difference of only 3% was found between these transformations, and this was shown to be nonsignificant by a specific comparison. When the component tones of the melodies were displaced by octave intervals violating melodic contour, performance was substantially poorer than that observed for the original melodies. A performance decrement of 40% was found for transformation OVC* relative to transformation O [F(1,96) = 420.57, p < .001].

The importance of melodic contour becomes even more apparent when the results on transformations LT and PC are considered. When only the pattern of pitch direction changes was preserved (PC), a 12% advantage was seen over transformation OVC* [F(1,96) = 34.16, p < .001]. Performance on transformation PC was 27% poorer than on transformations OPC* and O [F(1,96) = 189.30, p < .001]. When a linear transformation (LT) of the contour was taken, a 7% performance advantage was observed over a simple preservation of the contour (PC) [F(1,96) = 10.67, p < .005]. This result supports the dual claim that it is meaningful to talk about preserving relative degrees of contour and that the processing system is able to effectively use any additional specificity of the contour. However, this additional information is still insufficient for melodic perception. A 20% performance advantage was found for the original melodies (O) over those in transformation LT [F(1,96) = 99.73, p < .001]. This result argues that the equivalent performance observed on transformations O and OPC* is due to tone chroma, rather than an exact version of melodic contour.

Table 1 presents the percentage of correct identifications of the individual melodies, as a function of the type of transformation performed upon them. As can be seen from the table, the discriminability of the melodies differed [F(4,96) = 2.94, p < .005]. The best performance was seen on *God Rest Ye Merry Gentlemen* (70%), with a slightly lower level observed on *Happy Birthday* (64%), *London Bridge* (66%), and a still lower level on *Pop Goes the Weasel* (59%), and *On Top of Old Smokey* (55%). A significant Melody by Transformation interaction was obtained [F(16,384) = 13.06, p < .001], though there was no consistent pattern to these differences.

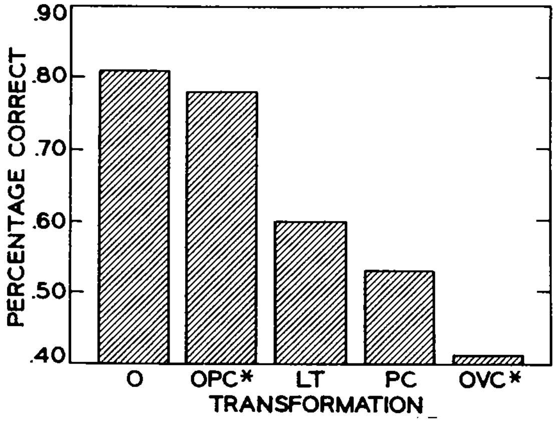

Figure 2. Percentage of correct identifications of the melody, as a function of the type of transformation performed upon the melodies (Experiment 1).

The levels of performance in the group data were sufficiently high to suggest the possibility that ceiling effects might be obscuring possible differences between transformations O and OPC*. Consequently, the subjects were divided into three groups on the basis of their level of performance on the original melodies. Figure 3 presents the percentage of correct identifications of the melody, as a function of the type of transformation, for each of these three groups of subjects. The top panel presents the average results for the four poorest subjects, whose identifications of the original melodies were at a level of between 41% and 50% correct. The middle panel shows the results for the 10 subjects whose performance was between 70% and 85% correct, while the bottom panel gives the data for the 11 subjects between 86% and 100% correct. It can be seen that subjects performing at quite different levels on the original melodies yielded essentially equivalent patterns of results, comparable to the pattern observed in the group data. Note, in particular, that even for the lowest performance group, essentially no difference was observed between transformations O and OPC*, suggesting that ceiling effects are not responsible for the group results. The only exception to the replication of the group results for the individual subject groups was the finding, for the poorest group of subjects only, that performance was slightly better (1%) on transformation OVC* than on transformation PC. However, this most probably reflects a floor effect for these subjects. Chance, in this task, is 20% and performance on transformations PC and OVC* was at only 25% and 26%, respectively.

Table 1
Percentage of Correct Identifications of Each of the Individual Melodies as a Function of the Type of Transformation in Each of the Four Experiments

	Transformation									
Melody	O	OPC*	LT	LT*	PC	PC*	EC*	CC*	OVC*	PI
Experiment 1										
Happy Birthday	77	70	69		69				43	
London Bridge	89	80	69		52				39	
Merry Gentlemen	89	86	65		73				36	
Pop Goes the Weasel	85	84	53		44				31	
Old Smokey	66	72	45		32				58	
Experiment 2										
Happy Birthday	97	90	87	78	67	68			42	
London Bridge	96	93	85	83	65	61			49	
Merry Gentlemen	76	72	69	64	49	46			46	
Pop Goes the Weasel	92	94	67	60	74	62			34	
Old Smokey	80	78	55	51	69	61			62	
Experiment 3										
Happy Birthday	87	79	78		61		66	54	51	
London Bridge	96	87	88		71		73	61	45	
Merry Gentlemen	90	79	72		81		56	67	33	
Pop Goes the Weasel	88	85	54		61		67	74	32	
Old Smokey	82	86	51		54		72	74	77	
Experiment 4										
Happy Birthday	78	74	71		66				42	26
London Bridge	93	88	74		62				52	30
Merry Gentlemen	81	79	64		65				47	46
Pop Goes the Weasel	91	89	75		76				47	05
Old Smokey	89	76	65		57				36	29

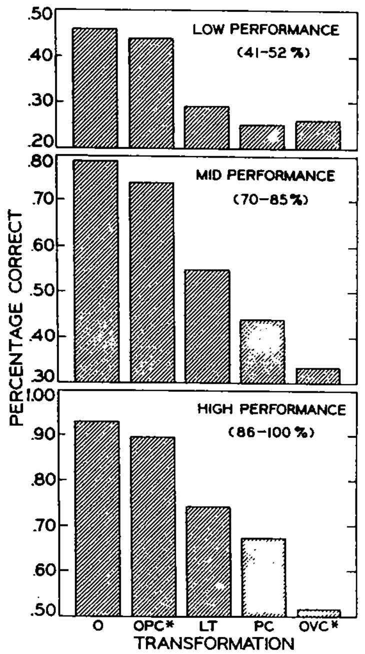

Figure 3. Percentage of correct identifications of the melody, as a function of the type of transformation, for three sets of subjects grouped on the basis of their overall levels of performance on the original melodies (Experiment 1).

Discussion

The results of the current study argue for the applicability of a bidimensional model of pitch to the perception of melodies by supporting the utility of the dimension of tone chroma. If chroma is functional in melodic perception, then tones which stand in an octave relationship should show unique generalization effects, such that the component tones of a melody can be displaced by octave multiples without disrupting perception. Direct confirmation of this prediction can be found in a comparison of performance under transformations O and OPC* of the current study. In transformation OPC*, the component tones of the melodies were displaced by octave intervals preserving the melodic contour. Subjects identified these transformed melodies as accurately as they identified the original melodies (transformation O). Yet, for the chroma dimension to be useful, contour must also be preserved. Though the only difference between transformations OPC* and OVC* was the preservation of contour in OPC*, performance was substantially poorer on OVC* than on OPC*. The better-than-chance performance on OVC* most likely resulted from preservation of half of the successive intervals of the original melody, providing partial information as to melody identity. However, the high level of performance seen on transformation OPC* cannot be due solely to preservation of melodic contour, as a large performance decrement was observed when only contour was available in both transformation PC and transformation LT.

The current results contradict Deutsch's (1969) model by demonstrating that the chroma dimension is functional in melodic perception. One possibility does exist for reconciling the results of transformation OPC* with her model. The argument has been made (Deutsch, 1972b, Note 1) that while tone chroma cannot be used directly in recognizing a melody, it can be used indirectly to confirm a specific hypothesis about the identity of a melody. Subjects in Deutsch's (1972b) study were unable to generate hypotheses, since they heard the melody only once and had to identify it from the set of all possible melodies. In contrast, our subjects, who worked with a small set of melodies for hundreds of trials, could presumably have generated and tested hypotheses, accounting for the excellent performance on transformation OPC*. However, Deutsch's potential explanation is inadequate, since the hypothesis-testing strategy could equally have been employed in transformation OVC*, where performance was quite poor. The more probable explanation is that by disrupting contour, Deutsch (1972b) eliminated the utility of tone chroma. The present results demonstrate that, when contour is preserved, identification of a melody on the basis of tone chroma is possible, but with contour violated, it is not. This conclusion is strengthened by the internal replication of Deutsch's results in transformation OVC*, which suggests that the accurate performance on transformation OPC* cannot be due to procedural differences between the two studies.

After the present studies were completed, Dowling and Hollombe (1977, Experiment 2) reported evidence relevant to the possible occurrence of octave convergence when the contour of a melody is preserved. A melody recognition task was employed in which the subjects were to identify which of 10 familiar melodies was presented on each trial. The first 32 quarter-notes of the melody were presented at a rate of 1.33 notes/sec. Three experimental conditions were employed. The melodies were either played entirely within a single octave, with the component displaced randomly across octaves, or with the tones displaced across octaves but with contour preserved. Although large differences were found between the 10 melodies, more accurate identifications were obtained when contour was preserved than when contour was violated. However, even with contour preserved, performance was far less accurate when the tones of the melodies were displaced by octave intervals than when the tones were played entirely within a single octave. The Dowling and Hollombe results appear to contradict the present findings. However, the preservation of contour in their study actually preserved only segments of the contour, since the tones of a unison interval were played in different octaves. For example, the interval C_4-C_4 could become C_5-C_4 or C_3-C_4. Thus, whenever the contour of a melody was flat (e.g., C_4-C_4), their manipulation introduced either a descending (C_5-C_4) or an ascending (C_3-C_4) interval. Dependent upon the number of unisons in a melody, this manipulation could radically disrupt the contour, greatly decreasing its utility.

The results of the first experiment argue that tone chroma is functional in melodic perception. Yet, the strength of this conclusion is somewhat minimized by the possibility that the excellent performance observed on transformation OPC* was due solely to the preservation of contour, independent of any contribution from tone chroma. The performance advantage on transformation OPC*, with respect to transformations LT and PC, rules out any simple explanation in terms of contour preservation. However, a somewhat more complex explanation still exists. The argument has been made (Idson & Massaro, 1976) that the discriminability of the contour of a melody will increase with the size of the changes in frequency. By displacing the tones of the melodies in transformation OPC* across octaves, the magnitude of most frequency changes was doubled, effectively stretching out the contour. It is possible that the superior performance observed on transformation OPC* over transformations LT and PC was due to the more discriminable contour. In this case, no conclusions concerning tone chroma would be justified.

Experiment 2 was designed to explore this "contour-stretching" hypothesis, by utilizing a transformation in which the contour is as discriminable as that in transformation OPC* but chroma is not preserved. This was accomplished by taking the melodies in transformations LT and PC and displacing their component tones by octave intervals, preserving the melodic contour. If the contour-stretching hypothesis is correct, then performance on these cross-octave contour-preserving transformations should be at a level equivalent to that seen for transformation OPC*. If chroma is critical, performance should be at a lower level than that observed on transformation OPC*. The strongest prediction which can be made from a bidimensional model is that the within-octave contour-preserving transformations (PC and LT) should yield performance at the same level as their cross-octave counterparts (LT* and PC*).

EXPERIMENT 2

Method

Subjects. The subjects were 18 University of Wisconsin undergraduates who received credit towards an introductory psychology course for their participation. None of the subjects had participated in Experiment 1.

Design. The experimental design was a 2 by 5 by 7 factorial. The three factors were melody length, melody identity, and type of transformation. The same two melody lengths, 8 or 12 notes, used in Experiment 1 were employed. The same melodies used in Experiment 1 were also employed: *Happy Birthday*, *London Bridge*, *God Rest Ye Merry Gentlemen*, *Pop Goes the Weasel*, and *On Top of Old Smokey*. Seven transformations of the melodies were now employed. The five basic transformations were retained from Experiment 1 (O, OPC*, PC, LT, OVC*) and two new transformations added (LT* and PC*).

Stimulus structure. The five transformations retained from Experiment 1 were constructed in the same way as in the earlier study. Transformations LT* and PC* were constructed by using the following algorithm. The melodies of both within-octave transformations were translated into cross-octave transformations by displacing the component tones of these melodies by octave intervals in accord with the melodic contour, in the same way that the original melodies were transformed in OPC*. Transformation LT* represents the cross-octave version of LT and transformation PC* represents the cross-octave version of PC.

Results

The data analysis was identical to that of Experiment 1. As the sessions factor again produced no significant effects, the data were pooled over sessions and reanalyzed. The main effect of melody length was nonsignificant in the analysis (F < 1), as were all effects involving melody length as a term. Consequently, the results given below are the average effects across the two melody lengths.

Figure 4 presents the percentage of correct identifications of the melody, as a function of the type of transformation. The type of transformation again had a large effect upon performance [F(6,102) = 62.92, p < .001]. Replicating the earlier results, a difference of only 3% was found between transformations O and OPC*. A specific comparison between these conditions revealed this difference to be nonsignificant. Once again, a substantial decrement was found for transformation OVC* relative to transformation O [F(1,102) = 255.33, p < .001]. Performance was significantly poorer on transformation LT than on transformations O or OPC* [F(1,102) = 40.76, p < .001]. Transformation LT again produced better performance than PC, which was, in turn, substantially better than OVC*. Of primary interest, comparable levels of performance were found for transformations LT and PC and their respective cross-octave counterparts, transformations LT* and PC*. A 5% performance advantage was found for transformation PC over transformation PC*, and a 6% advantage of transformation LT over transformation LT*. Specific comparisons of transformation LT against transformation LT* and transformation PC against transformation PC* were both nonsignificant.

The percentage of correct identifications of each of the individual melodies, as a function of the type of transformation, is shown in Table 1. The main effect of melody identity was significant in the analysis [F(4,68) = 4.76, p < .005]. The overall levels of performance for the individual melodies did not reflect those found in Experiment 1. The highest levels of performance were seen on *Happy Birthday* (75%) and *London Bridge* (76%), with slightly lower levels of performance found for *Pop Goes the Weasel* (70%) and *On Top of Old Smokey* (65%) with the lowest performance seen on *God Rest Ye Merry Gentlemen* (60%). The Melody by Transformation interaction was also significant [F(24,408) = 6.28, p < .001]. However, the pattern of results did not reflect that observed in Experiment 1, suggesting that the interactions are not due to any consistent underlying effects.

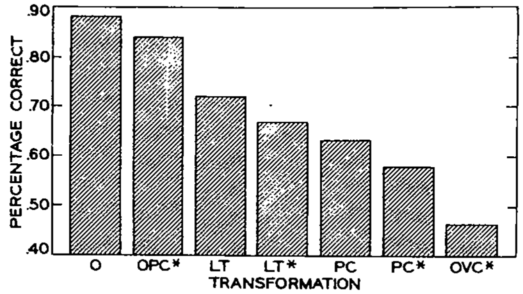

Figure 4. Percentage of correct identifications of the melody, as a function of the type of transformations performed upon the melodies (Experiment 2).

Discussion

The results of Experiment 2 replicate exactly those of Experiment 1, with respect to the five overlapping transformations: O, OPC*, LT, PC, and OPC*. The results of fundamental importance concern performance on transformations LT* and PC*, in which the melodies in transformations LT and PC were altered by dispersing the component tones across octaves in accord with the contour. This effectively stretched out the contours of the melodies in transformations LT* and PC* in a manner analogous to that for transformation OPC*. If chroma plays no role and the high level of performance seen on transformation OPC* is due to the increased discriminability of the contour, then performance on transformations LT* and PC* should have been at an equally high level. This result was not obtained; performance on transformations LT* and PC* was at a substantially lower level than performance on transformation OPC*. The levels of performance obtained on transformatons LT* and PC* did not differ significantly from the levels of performance obtained on their within-octave counterparts, transformations LT and PC. This result suggests that the chroma dimension made the melodies in transformations LT* and PC* identifiably equivalent to those in transformations LT and

PC. This is particularly interesting because halving the intervals in transformation LT resulted in tones in transformations LT and LT* which were not notes of the musical scale. The equivalent performance in these two transformations suggests that the similarity of tones in an octave relationship is not limited to notes of the musical scale.

Experiment 2 offers clear support for a bidimensional model of pitch, by eliminating an explanation for the results of Experiment 1 which was based upon the perception of contour. Unfortunately, as very little is known about the perception of contour, a large set of alternative explanations could be developed. Rather than evaluating each of these interpretations individually, a more parsimonious approach would be to test the entire class of contour-based explanations against the bidimensional model. Any contour-based explanation would argue that the superior performance on transformation OPC* is due to preserving, in some sense, an exact version of the contour of the original melodies of transformation O. In contrast, a chroma explanation would submit that it is the octave relationship between the tones of the melodies in transformations O and OPC* which is responsible. A clear test between these alternatives is to alter each note of the melodies in transformation OPC* by one or two semitones. This manipulation effectively eliminates the utility of the chroma dimension while preserving a highly accurate version of the contour. If some contour explanation is correct, then performance on such a transformation should be quite good. Alternatively, if a chroma explanation is valid, performance on such a transformation should be no better than performance on other contour-preserving transformations.

Experiment 3 was designed to carry out this test of the class of contour explanations, by employing two new transformations which altered the OPC* melodies by either one or two semitones. In transformation EC*, the absolute magnitude of each interval was increased by one or two semitones, effectively expanding the contour. In transformation CC*, each interval was decreased by a comparable amount, effectively contracting the contour.

EXPERIMENT 3

Method

Subjects. The subjects were 16 University of Wisconsin undergraduates who received credit in an introductory psychology course for their participation. None of the subjects had participated in either Experiment 1 or Experiment 2.

Design. The experimental design was a 2 by 5 by 7 factorial on melody length, melody identity, and type of transformations. The length of the melodies were either the first 8 or the first 12 notes of the melody. The same five melodies used in the earlier studies were employed: *Happy Birthday*, *London Bridge*, *God Rest Ye Merry Gentlemen*, *Pop Goes the Weasel*, and *On Top of Old Smokey*. The seven transformations included the five basic transformations from Experiment 1 (O, OPC*, OVC*, LT, and PC) plus the two new transformations (CC* and EC*).

Stimulus structure. The stimuli in transformations O, OPC*, OVC*, LT, and PC were identical to those in the previous studies. In transformations EC* and CC*, cross-octave versions of the original melodies were constructed just as for transformation OPC*. However, in these transformations, the intervals were either larger (transformation EC*) or smaller (transformation CC*) than those of transformation OPC*. In transformation CC* the absolute magnitude of each interval was decreased by either one or two semitones. These transformed melodies had intervals corresponding to the correct absolute interval + 10 or 11 semitones, rather than an exact 12-semitone octave. In transformation EC*, each interval was increased by one or two semitones so that the transformed melodies had intervals corresponding to the correct absolute interval + 13 or 14 semitones rather than an exact octave.

Results

The data analysis was equivalent to that of Experiment 1. The main effect of sessions was nonsignificant (F < 1), as were all interactions involving sessions as a term. Consequently, the data were pooled over sessions and reanalyzed. Since no overall difference was found between the two melody lengths, and since this factor did not interact significantly with any other factor, the data will be discussed in terms of the average results for the two melody lengths.

Figure 5 presents the percentage of correct transformations of the melody, as a function of the type of transformations performed upon the melodies. It is apparent from the figure that the type of transformation had a large effect upon performance [F(6,84) = 49.94, p < .001]. The original melodies were identified 5% better than those of transformation OPC*; a difference which was marginally significant [F(1,84) = 4.86, p < .05]. Replicating the earlier studies, transformation OVC* produced substantially poorer performance, yielding a performance

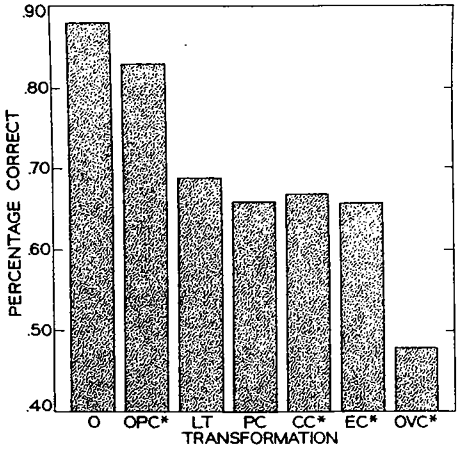

Figure 5. Percentage of correct identifications of the melody, as a function of the type of transformations performed upon the melodies (Experiment 3).