Bucknell University
Bucknell Digital Commons

Faculty Journal Articles | Faculty Scholarship

1984

Perception of Structure in Novel Music

Andrea R. Halpern
Bucknell University, ahalpern@bucknell.edu

Follow this and additional works at: https://digitalcommons.bucknell.edu/fac_journ

Part of the Cognitive Psychology Commons, and the Music Commons

Recommended Citation

Halpern, Andrea R.. "Perception of Structure in Novel Music." *Memory & Cognition* (1984) : 163-170.

This Article is brought to you for free and open access by the Faculty Scholarship at Bucknell Digital Commons. It has been accepted for inclusion in Faculty Journal Articles by an authorized administrator of Bucknell Digital Commons. For more information, please contact dcadmin@bucknell.edu.

Memory & Cognition
1984, 12 (2), 163-170

Perception of structure in novel music

ANDREA R. HALPERN
Bucknell University, Lewisburg, Pennsylvania

Two experiments demonstrated the way in which musicians and nonmusicians process realistic music encountered for the first time. A set of tunes whose members were related to each other by a number of specific musical relationships was constructed. In Experiment 1, subjects gave similarity judgments of all pairs of tunes, which were analyzed by the ADDTREE clustering program. Musicians and nonmusicians gave essentially equivalent results: Tunes with different rhythms were rated as being very dissimilar, whereas tunes identical except for being in a major versus a minor mode were rated as being highly similar. In Experiment 2, subjects learned to identify the tunes, and their errors formed a confusion matrix. The matrix was submitted to a clustering analysis. Results from the two experiments corresponded better for the nonmusicians than for the musicians. Musicians presumably exceed nonmusicians in the ability to categorize music in multiple ways, but even nonmusicians extract considerable information from newly heard music.

People encounter, produce, and reproduce many different kinds of music in everyday life. The music's familiarity can vary widely. At one extreme, most everyone has a mental storehouse of familiar tunes that they can recognize and reproduce with fair accuracy. This ability is often retained over a span of many decades (Bartlett & Snelus, 1980). At the other extreme, a number of common situations require efficient processing of novel music. Consider musical themes in movies: The theme's effectiveness often depends on its recognizability after one exposure. A good example is the use of the five-note "alien" theme in *Close Encounters of the Third Kind*. The audience was expected to recognize the theme after its first presentation, and also to recognize the complex variations "composed" by the aliens and the earth computer in the movie's climactic scene. One aid to memory for familiar music and to the learning of complex music is the nonmusical associations we attach to the sounds. For instance, patriotic songs tend to be grouped in memory regardless of their musical similarity (Halpern, in press). The purpose of the current experiments was to investigate the cognitive schemes used to comprehend and remember novel tunes when nonmusical contexts are unavailable.

Listeners presented with a new piece that uses a theme-and-variations technique are in just that kind of listening situation. That is, a musical idea is presented and then is elaborated by varying one or more of its musical parameters: rhythm, tempo, harmony, or combinations of these. To what extent do listeners hear and remember the relationships between themes and their variations? Pollard-Gott (1983) found that, after repeated listening, subjects classified passages from a Liszt sonata by their similarity of musical theme (underlying melody). But, when only briefly exposed to the passages, they used more elemental aspects of music to classify the passages: pitch range, loudness, etc. Musicians, more so than nonmusicians, made use of the more sophisticated thematic classification, and musicians intimately familiar with that particular piece used only the thematic classification.

Because Pollard-Gott (1983) used previously composed music, the exact nature of the relationships among the passages was sometimes difficult to specify. The current experiments used melodies purposely composed for use as stimuli. All the melodies were pleasant and natural sounding, but were similar to one another in exact, specifiable musical ways. To be more precise, a set of melodies was constructed by combining two levels of four factors (exact sequence of intervals, contour, rhythm, and mode) in a factorial manner. Thus, each melody had a well-defined musical relation to each other melody in the set. The basic question asked here was: Which relationships among the tunes would be most salient to new listeners? Both similarity ratings and a memory task were used to help answer this question.

In addition to their importance in music theory, there is evidence to suppose that each of the above factors could serve as an organizational scheme for the entire set. Previous research has demonstrated that listeners are aware of each of these factors in music. For instance, Attneave and Olson (1971) found that, when subjects transposed (changed the absolute frequency of) a sequence of intervals, they were quite accurate at maintaining the frequency relationships between the tones that the intervals comprised. This

Experiment 1 was included in a doctoral dissertation submitted to Stanford University. I would like to thank Charlotte Tsuyuki and Caroline Wilshusen for running Experiment 2, Jim Corter for discussions about additive clustering, and Owen Floody for comments on a draft of this manuscript. Address correspondence to Andrea R. Halpern, Psychology Department, Bucknell University, Lewisburg, PA 17837.

 Copyright 1984 Psychonomic Society, Inc.

suggests that the abstract nature of the interval was apparent to them (although a familiar pattern was needed by nonmusicians to achieve the same accuracy as musicians). Cuddy and Cohen (1976) showed that intervals formed by noncontiguous tones were also abstracted.

Given that a musical factor is perceived by a listener, too much attention to the aspect can lead to confusion among tunes related by sharing the factor. For instance, Dowling and Fujitani (1971) presented subjects with quasi-randomly generated five-note melodies. The sequences were then transposed. Subjects falsely recognized sequences that maintained the original contour but that differed in the exact note sequence. Similarly, Massaro, Kallman, and Kelly (1980) found that different tunes with the same contours were confused in an identification task.

Tunes that are very similar may not be perceived as such if their one difference is on a salient musical factor. Dowling (1973) showed that imposing a new rhythmic grouping on a tune decreases its recognizability. And in another domain, changing a tune's mode from major to minor (or vice versa) obscures the underlying similarity of the two versions (Francès, 1958). In each of these cases, either confusion between items that *share* a factor or inability to perceive the relatedness of items that *differ* by a factor suggests that the tunes were mentally grouped by that factor. In order to ascertain more about preferred ways of organizing tunes, the current experiments gave listeners tunes that could be grouped by many competing strategies.

Another interesting question addressed in these experiments was the extent to which musical training or aptitude contributes to the mental organization of novel tunes. Given the difficult tasks of making similarity judgments (Experiment 1) or of identifying recently learned tunes (Experiment 2), how do musicians differ from nonmusicians, if at all? One hypothesis is that an untrained listener would be unable to perceive much structure whatsoever in the music and would give an incoherent pattern of similarity judgments or identification errors. The musician might be cognizant of some of the relationships among the tunes and reflect that knowledge in orderly results. A second hypothesis is that even nonmusicians would be sensitive to the musical properties investigated here, and would give interpretable results. Given that nonmusicians are aware of the musical structure, then it stands to reason that musicians would have an acute sense of the musical relations being probed. If several or all of the musical relations were salient to them, then individuals would vacillate among the various grouping strategies. Overall, a disorderly pattern of group similarity judgments would emerge from these multiple schemes. The overtrained musician might also produce error-free (and thus, noninformative) performance in identification. An intermediate hypothesis is that both groups would be moderately sensitive to the structure of the tunes, but would differ in which musical factors were salient.

To summarize, a set of highly related tunes was constructed by factorially combining several musical attributes. Musicians and nonmusicians rated the similarity of every pair of tunes in Experiment 1. In Experiment 2, a different group of subjects was trained to identify the tunes, and the number of confusion errors was used as the dependent measure.

EXPERIMENT 1

In this experiment, pairs of melodies were presented to subjects for judgment of similarity.

Method

Subjects. Participants were 23 Stanford University undergraduates who volunteered for course credit. No subjects possessed "perfect pitch" or had hearing problems. The 11 musicians all had had at least 5 years of training on at least one instrument, and most were still active in music. The 12 nonmusicians had had fewer than 5 years of training, if any; the training had ceased long ago; and none was currently active in music.

Materials. Sixteen stimulus melodies, shown in Figure 1, were composed. Initially, Tunes A and I were composed. Both are simple eight-note melodies with the same up-down-up pitch contour. All tunes derived from Tune A have the same sequence of intervals and are called Note Pattern 1; those derived from Tune I are called Note Pattern 2. Next, each tune was written with either of two rhythms, seen by comparing melodies in the

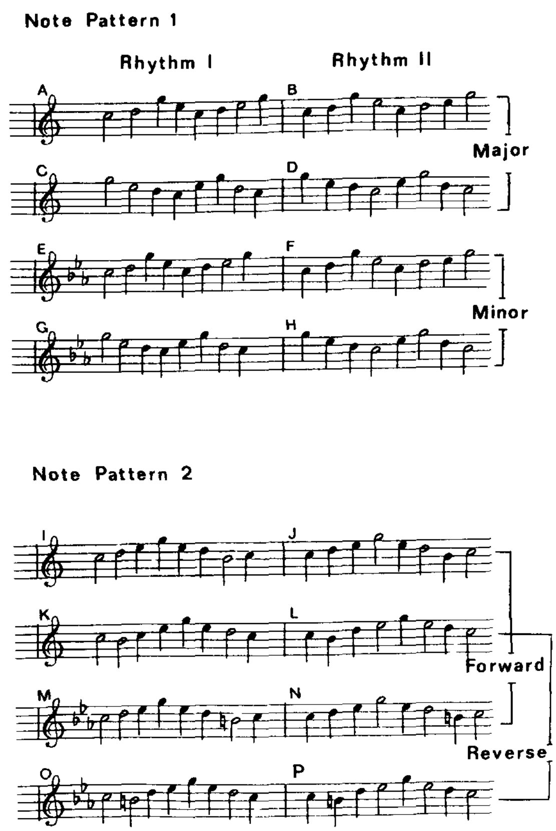

Figure 1. Stimulus melodies used in both experiments and their musical interrelationships.

first column (Rhythm I) with those in the second (Rhythm II). The original tune was then written with pitches in back-to-front order ("retrograde"), but with the original rhythm maintained. This gave a Forward and a Reverse version of each tune. For example, Tune C is the reverse of Tune A, Tune K is the reverse of Tune I, and so on. Note that, with these patterns, the reversal inverts the pitch contour from up-down-up to down-up-down. The final factor was mode: Each tune (originally in a major mode) was also written in a minor mode. Thus, Tune E is the minor version of Tune A, and Tune M is the minor counterpart of Tune I.

In summary, the 16 tunes resulted from factorially combining two note patterns, two rhythms, two directions, and two modes. All the resulting tunes sounded like plausible music in spite of their having been algorithmically generated.[1]

Procedure. Two stimulus tapes were prepared. Each tape presented the complete set of 120 similarity judgments in a different random order, that is, every possible pairwise comparison without regard to order. A professional pianist recorded the tapes on an electric organ. The tempo was set at 120 beats (quarter notes) per minute. Tapes were played monaurally on a Revox tape recorder through one speaker. Any pair presented in order AB on one tape was heard in order BA on the other tape. Each tape was presented to approximately half the subjects.

One member of each pair was played in the key of C, as notated in Figure 1. The other member was transposed to the key of F. Transposition slides all the notes of a melody up or down an equal number of pitch units. A melody and its strict transposition are presumed to be musically equivalent. The transposition prevented subjects from making judgments based solely on the number of overlapping notes of the two melodies. The key of each tune and the key of the first member of each pair were counterbalanced over tapes.

The subjects were tested in small groups. First, to familiarize the subjects with the stimulus set, all the tunes were played once in the key of C. The subjects then made the 120 similarity ratings. On each trial, the subjects heard one melody, followed by a 5-sec pause, and then heard the second melody. The subjects then had 7 sec in which to record similarity judgments about the two tunes. A 1-7 rating scale was used, such that 1 indicated that the passages were not at all similar and 7 indicated that they were very similar.

During a rest break that occurred after Trial 60, the subjects filled out a questionnaire about their musical backgrounds. The experimental session lasted about 1 h

Results

When questioned, no subjects claimed to be aware of the factorial structure of the tune set. Many commented on the similarity of the tunes or on the way in which some of the tunes differed (e.g., "Some tunes were major and some were minor"). Nevertheless, the results of both subject groups demonstrated more complete knowledge of the stimuli than is evident from their comments.

The dependent measure was the mean similarity score for each pair of tunes, combining data from both stimulus tapes. These data were submitted to a clustering program called ADDTREE (Sattath & Tversky, 1977). ADDTREE represents proximity data as additive trees. In such a diagram (see Figure 2), stimulus items appear to the far right. Similarity between items is represented as the interitem distance, calculated by summing the horizontal paths between pair members. Vertical distances are arbitrary. Items occurring close together are maximally similar to each other and minimally similar to other items. The branching pattern of the tree reveals the clusters. The length of a branch out to a cluster is a measure of the cluster's distinctiveness. The program does not label the branches, but ease of such assignments is one measure of how well the solution has represented the data. Proportion of variance accounted for is a more quantitative index of the acceptability of the solution.

The ADDTREE solution for nonmusicians appears in the left panel of Figure 2, and that of the musicians appears in the right panel. The letters at the right refer to the tune labels in Figure 1. Some of the clusters are labeled by their musical relatedness.

The major results are as follows: Both groups gave a fairly interpretable pattern of results and the pattern was similar for each group, but the nonmusicians' data were better captured by the ADDTREE solution than by the musicians' data. Each of these results will be discussed in turn.

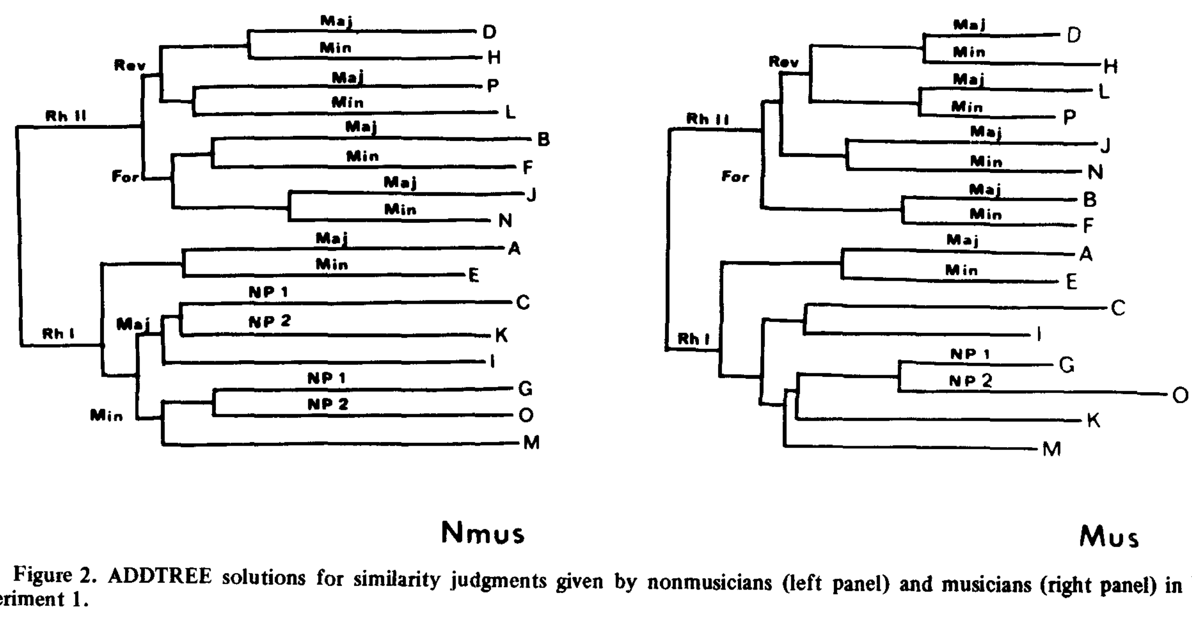

Figure 2. ADDTREE solutions for similarity judgments given by nonmusicians (left panel) and musicians (right panel) in Experiment 1.

As noted, most of the branches in each solution could be labeled with the name of a factor. Particularly for Rhythm II melodies, the branching pattern reflects the factorial stimulus set. The Rhythm I melodies are in a less orderly pattern, but the pattern is still coherent.

To be more specific, the major stimulus division for each group was rhythm: All the Rhythm I tunes were rated as being more similar to each other than any one Rhythm I tune was to any Rhythm II tune. For both groups, the Rhythm II melodies formed the more distinct cluster, with a clear-cut structure. Examining each diagram from right to left, we see that Rhythm II tunes identical except for being in a major versus a minor mode were considered to be most similar. Next, all tunes with the same direction (or contour) were clustered together. Finally, all the tunes with same rhythm were considered to be similar. The result can be restated by examining each diagram from left to right. Rhythm is the most discriminating factor, followed by direction (contour), followed by note pattern. (One of the four factors, in this case, note pattern, is left unlabeled for clarity because it is uniquely determined once the other three factors have been specified.)

The Rhythm I melodies are a little less orderly. One additional major-minor closest-neighbor pair is formed by Tunes A and E. The remaining six tunes have a different structure for each group. For the nonmusicians, they divide into two clusters: major and minor. Two members of each cluster are identical except for note pattern; the third is a singleton. For the musicians, the Rhythm I melodies are more fragmented, with one same-except-for-note-pattern pair and the others not classifiable by reference to a single shared factor.

The nonmusicians, more so than the musicians, agreed among themselves as to the preferred similarity orderings leading to a better ADDTREE solution. The solution for the nonmusicians accounted for 80% of the variance in the data, whereas the musicians' solution accounted for only 60%. The better solution is also reflected in the more complete and logical labeling system allowed in the nonmusicians' versus the musicians' diagram. And finally, the initial rhythm clusters are more distinct (have longer branches) for the nonmusicians than for the musicians. Again, this reflects the more structured results obtained from the nonmusicians.

Discussion

To summarize, there was clearly a coherent pattern to many of the similarity data from this experiment. When divested of their nonmusical contexts, tunes can be processed as purely musical percepts. Many orderly musical relations between tunes were reflected in the results. This occurred even though the subjects were not consciously aware of the complex structure of the melody set and were required to judge melodies after transposition.

The nonmusicians were at least as sensitive to the musical structure as were the musicians. Interestingly, they showed a more well-defined structure for the Rhythm I melodies than did the musicians. Musicians presumably have greater skill in formal analysis of music. They may have understood more clearly the multiple ways in which to group the tunes. Different individuals may have adopted different methods of grouping, thus preventing the consensus found among the nonmusicians.

The ordering of salience for the particular musical relations deserves some discussion. Rhythm was the most important distinguishing factor, being the main division in all the ADDTREE solutions. Direction or contour was next most important. In these particular stimuli, direction and contour were confounded. Thus, the subjects in this study could either elect to regard as being similar tunes that were retrogrades of one another, or they could group together tunes that shared the same contour. Both factors have been shown to operate in music perception. A tune's retrograde can be recognized as such with a moderate rate of success (Dowling, 1972). On the other hand, tunes can be mistaken for others sharing the same contour (Dowling, 1978). In this experiment, the subjects overwhelmingly preferred to group together tunes that shared the same contour rather than those that were retrograde pairs. In general, the psychology literature also has shown a much stronger effect of contour than of retrogression on melody recognition. And although composers have used retrograde motion as a musical device in many periods of Western music, it is relatively rare (Berry, 1966), perhaps for good reason.

The major/minor difference was the least salient for both groups. Those pair members share many of the same notes, after transposition is allowed for. Perhaps playing the melodies with an appropriate harmonic accompaniment would have differentiated their modes more clearly.

Rhythm II tunes were seen as being more distinctive than Rhythm I tunes and led to a neater classification. Perhaps this was due to the greater ease of processing tunes in the Rhythm II triple meter (waltz time) than in the Rhythm I duple meter (two-step). Triple meters, with their one strong and two weak beats, may instantiate a stronger rhythm schema than do more diffuse strong-weak duple meters. Beginning ballroom dancers often find a waltz in triple meter easier to keep time to than a fox-trot in a duple meter.

This experiment required that subjects judge the similarity between tunes. If the similarity structure were the limiting factor on tune perception, then subjects should confuse tunes they rated as being similar. This was pursued in the next experiment.

EXPERIMENT 2

In this experiment, subjects learned names for a subset of tunes from Experiment 1. Errors in a subsequent identification task provided information about the confusability of each pair of tunes.

Of interest is the correspondence between the similarity and the identification results. Are items that subjects rate as sounding similar confused when memory is tested directly? Shepard (1972) successfully used both similarity judgments and confusion errors to map out the mental organization of phonemes in English. Here, the different subject groups may differ in the extent of the correspondence. On the one hand, the correspondence may be equally good or poor in both groups. On the other hand, the correspondence may differ due to training. If the musicians are aware of the multidimensional structure of each melody, then they may be simultaneously aware of the similarities and differences between related tunes. For instance, musicians may rate as being highly similar two tunes that differ only in rhythm. However, they would still be able to distinguish the tunes by mnemonic: "same except for rhythm." Thus, for these subjects, similarity ratings and errors would not necessarily correspond. Accurate identification benefits from attention to both similarities and differences, whereas similarity ratings can be performed by using just similarities. Nonmusicians, lacking experience in analysis of musical relations, would be more likely to confuse items that also sounded alike to them.

Method

Subjects. The same criteria as in Experiment 1 were established for musicians and nonmusicians. The 11 musicians and 11 nonmusicians were Bucknell University undergraduates. They were unpaid volunteers. No subject reported any hearing problems or possession of "perfect pitch."

Materials. Melodies A through H from Experiment 1 served as stimuli. A reduction in the number of stimulus tunes was necessary in order to make a learning experiment feasible. Of the tune-construction factors, note pattern seemed to be the least interesting theoretically and was therefore eliminated.

Two training tapes and one test tape were prepared. A professional pianist played all materials on a well-tuned piano. A Superscope cassette recorder was used for both stimulus recording and playback to the subjects.

For each training tape separately, Melodies A through H were assigned randomly to the numbers 1 through 8. Melody 1 was presented twice, followed by two presentations of Melody 2 and then by several presentations of each in a random order. Melody 3 was then introduced with two presentations, followed by random presentations of Melodies 1, 2, and 3, and so on, until all melodies had been presented. The last melody was introduced by Trial 40. The 30 remaining trials presented all eight melodies in a random order, for a total of 70 training trials.

Melodies were separated by 8-sec pauses. All melodies on one tape were in the key of C, and on the other, in the key of F. After Trial 70, three additional tunes were presented in a new key (G and C for each tape, respectively) to familiarize the subjects with the idea of transposition.

The test tape was organized into three blocks of 32 trials, each block containing four presentations of each melody in a random order. Each block was in a single key: C, F, or G. Each melody was followed by an 8-sec pause.

Procedure. The subjects were tested individually in a quiet room. Upon arrival, the subjects filled out a questionnaire about their musical backgrounds. The subjects then heard one of the training tapes (each tape was heard by approximately half the subjects). During training, the experimenter announced each melody's number after the melody's presentation. The number was that assigned to each melody on the training tape that the subject had heard. The subjects were instructed to learn the number name/melody pairing for later testing. The subjects could request a rest break or pause at any time.

After the 70 training trials, the idea of transposition was explained, and then the three transposed melodies were presented without feedback.

During the subsequent test phase, the subjects were presented with two of the three test trial blocks. The order and occurrence of the trial blocks was counterbalanced over subjects. Thus, each melody was presented for identification eight times. After each stimulus trial, the subjects recorded the tune's number on an answer sheet. If unsure of a tune's identity, the subjects were required to guess. There was a rest break after the first block of trials. The complete experiment lasted between 45 min and 1 h.

Results and Discussion

For each subject, a confusion matrix that recorded how often a given response was made to a given tune was tabulated. These matrices were summed across subjects and training tapes separately for musicians and nonmusicians to provide the data for most of the analyses. Although tunes were presented an unequal number of times in the training tapes, the confusion between any two tunes by both subject groups was uncorrelated with presentation disparity for either tape.

For neither group was the task trivial or impossible. If the subjects were simply guessing, they would be correct on 12.5% of the trials. The musicians were in fact correct on 54% of the trials, and the nonmusicians on 31%. Because performance was far from perfect even for the musicians, it was reasonable to pursue an analysis of the error patterns. The range of frequencies in the "error" cells of the matrix was wide: from 0 to 24 for the musicians, and from 2 to 28 for the nonmusicians (maximum possible in each cell = 88, resulting from 11 subjects responding to eight presentations of each melody).

The initial full matrix recorded separately the frequency with which each Tune B was given in response to presentation of each Tune A (lower half-matrix) and the frequency of Tune A given as a response to Tune B (upper half-matrix). Correlations performed between all such AB/BA pairs showed a significant relationship between them: $r(26) = .56$ for the nonmusicians, $p < .01$; and $r(26) = .87$ for the musicians, $p < .001$. Because of their similarity, the two half-matrices were added together for all subsequent analyses.

The first analysis compared the confusability of items that differed by one, two, or three musical factors. For example, Tunes A and B, Tunes A and E, and Tunes A and C differ from one another only by rhythm, mode, and contour, respectively. There are 12 such one-factor pairs. Similarly, Tunes A and F, Tunes A and D, and Tunes A and G differ by two factors (12 such pairs); and Tunes A and H differ by all three factors (4 such pairs). If the stimulus structure were affecting error patterns, then one-factor pairs should be more often confused than two-factor pairs, which should be more often confused than three-factor pairs. Mean number of errors for each pair type are shown in Table 1, separately for musicians and nonmusicians. Overall, musicians made

Table 1
Mean Number of Errors for Each Pair Type in Experiment 2

Pair Type	Nonmusicians	Musicians
Differ by 1 Factor	24.7	20.6
Differ by 2 Factors	12.3	5.4
Differ by 3 Factors	9.8	1.5

Note–Maximum possible = 116.

fewer errors than did nonmusicians, and for each group, errors were in the predicted direction. This was confirmed by a two-way analysis of variance (unweighted means), with pair type and subject group as the factors. The musician/nonmusician difference was significant [$F(1,50) = 4.92$, $p < .05$], and errors differed reliably among the pair types [$F(2,50) = 12.97$, $p < .01$]. In terms of frequencies, the first and second most frequent error responses to each song by each group were almost always tunes differing by one factor (2 exceptions among 32 instances). There was no significant group x pair type interaction [$F(2,50) < 1$]. In other words, musicians and nonmusicians reacted similarly to the stimulus set, despite the overall superiority of the musicians.

The identification errors can be viewed as similarity measures, such that a high score implies close similarity. The error matrix was entered into an ADDTREE program in order to generate a more detailed view of the error patterns. The ADDTREE solution for nonmusicians is shown in the upper half of Figure 3, and that for musicians is shown in the lower half.

The solutions are easily interpretable, although slightly less so for the musicians than for the nonmusicians. The proportions of variance explained by the solution were .84 for the nonmusicians and .88 for the musicians.

As seen in Figure 3, the nonmusicians' solution exactly reflects the factorial structure of the tunes. The first factor division is by direction (contour), the second is by rhythm. This leaves the major/minor distinction as the least salient factor (only the first of the four pairs is labeled as "maj/min" for clarity). The lower half of Figure 3 shows that the musicians also find direction to be the easiest factor by which to distinguish melodies. The Forward tunes are then further distinguished by mode and then by rhythm. The Rhythm II Reverse melodies form a cluster, but the Rhythm I Reverse melodies do not cluster together.[2]

Comparing Figures 2 and 3 for the nonmusicians, we see that they genuinely seem confused about the major/minor distinction. In similarity ratings, the majority of closest-neighbor pairs are distinguished solely by mode. And similarly, all the most highly confused pairs in identification are major/minor pairs. There is a reversal in the order of salience of the other two factors across experiments. Although rhythm is the most discriminable factor in similarity ratings, followed by contour, the opposite is true for confusions.

The correspondence between Experiments 1 and 2 is weaker for the musicians than for the nonmusicians. Although musicians rated a number of major/minor pairs as being very similar (Figure 2), those pairs in general were not confused with one another (Figure 3). Conversely, tunes with different rhythms were rated as being relatively dissimilar (Figure 2), but were confused with one another more frequently than any other pair type (Figure 3). Perhaps as the distinction between tunes of different rhythms becomes stronger, the distinction within same-rhythm groups becomes relatively obscured, as may be the case in categorical perception of speech (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).

For both groups, the distinction of tunes by contour (direction) is robust, again demonstrating that the competing retrograde relationship between two tunes is not very discriminable. Among tunes differing by one factor, confusion of retrograde pairs was the least frequent error for both groups (9% of one-factor errors for musicians; 23% for nonmusicians). For no song was its retrograde the first or the second most frequently given error response. The same finding can be restated as showing the sensitivity that both groups have to contour: Items with different contours are not often confused.

The factorial nature of the stimulus set made intriguing a secondary analysis of the data by multidimen-

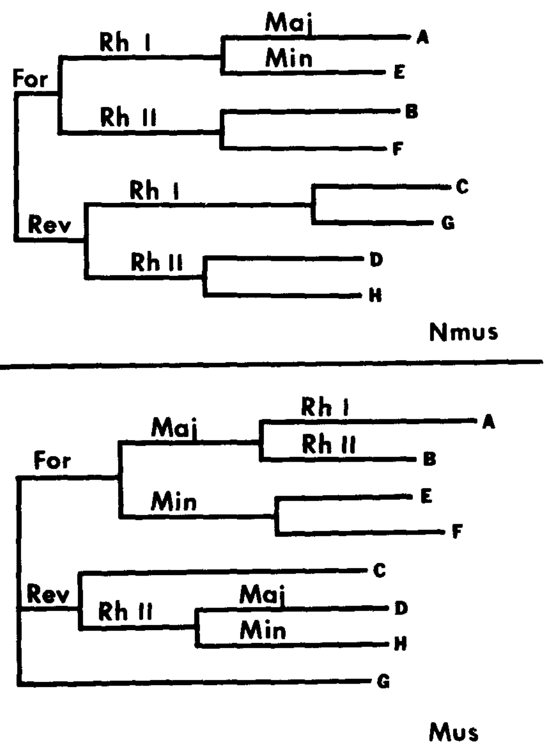

Figure 3. ADDTREE solutions for confusion errors elicited from nonmusicians (top) and musicians (bottom) in Experiment 2.

sional scaling. Both data sets were submitted to the KYST scaling program. Solutions were obtained for two and for three dimensions. If the subjects regarded each musical factor as being equally salient, then the three-dimensional solution should correspond to a cube, wherein every vertex (stimulus) is about equidistant from every other one. As may be inferred from the above results, the cubical representation was not found. The nonmusicians' solution (stress 2 = .024) in three dimensions greatly resembled the ADDTREE branching pattern. One plane separated forward from reversed tunes, and another separated Rhythm I from Rhythm II tunes. Tunes forming major/minor pairs were grouped closely together in the third plane. The four pairs were about equidistant in that plane. The musicians' solution (stress 2 = .040) revealed a Forward/Reverse plane, a major/minor plane, and an uninterpretable scattering of items in the third dimension.

In summary, musicians and nonmusicians created errors in the pattern logically predicted from the structure of the stimulus melodies. Although musicians made fewer errors than nonmusicians, the pattern of results was similar for both groups. Confusion of a tune with its retrograde pair was infrequent in both groups. Musicians confused same-except-for-rhythm pairs most often, whereas nonmusicians confused major/minor pairs most often. The nonmusicians' results more closely corresponded to the similarity ratings in Experiment 1 than did the musicians' results.

GENERAL DISCUSSION

When novel tunes were encountered without extramusical associations, listeners were quite good at detecting a subset of the complex musical relations among them. This knowledge was not consciously possessed; instead, musicians and nonmusicians gave consistently orderly results in a similarity rating and an identification experiment.

Regarding group differences, the intermediate hypothesis stated in the introduction is the one most strongly supported by these results. That is, both groups were fairly sensitive to the musical structure, but differed in which factors were most salient.

The contour and rhythm distinctions were the most obvious ones to the nonmusicians, but the major/minor distinction was not very evident in their results. Across the two tasks, the musicians varied in which factors were the most salient. They pretty much agreed with the nonmusicians' ordering of similarity judgments. They made many fewer identification errors than did the nonmusicians, but when they did make such an error, they tended to confuse a tune with its same-except-for-rhythm counterpart. This is a difficult result to understand in light of Deutsch's (1980) and Dowling's (1973) findings that giving notes a new rhythm slows down the processing of those notes. Future research should examine the musicians' rhythm confusions more closely.

The performance of the nonmusicians was quite good in both experiments, and certainly exceeded their own opinions of their ability. Many nonmusicians had had absolutely no musical training prior to the experiment. They claimed to have no awareness of the stimulus structure. Yet their performance was very interpretable in Experiment 1 and, in addition, considerably above chance in Experiment 2 (although both groups claimed to find the learning phase very difficult). Both experiments confirmed the transposition ability of even untrained listeners: The similarity judgments were performed across keys, and all subjects identified at least one block of tunes in a key other than the one on which they had been trained. Therefore, we may conclude that musical training is not necessary in order for listeners to perceive and categorize some of the musical infrastructure of a piece.

However, in both experiments, the confusability of the major/minor pair was a prominent and robust feature of the nonmusicians' results. In these particular melodies, changing a tune between major and minor meant changing only one note: the lowering of the third degree of the scale. In a literal sense, only one piece of information is thus altered, as compared with the many notes that are changed when rewriting a tune's rhythm. To the musician, the local change of lowering the third implies a very global change in scale system, implied harmony, melodic movement, etc. To the nonmusician, the local change may be perceived as only that—a small difference that is hard to detect. The development of appreciation of different modes may be a good place to focus on musician/nonmusician differences in future research.

Similarity ratings and identification errors corresponded better for nonmusicians than for musicians. This implies that the nonmusicians attended more to musical aspects that tunes shared and were less able to categorize and/or remember the tune differences that distinguished them. This is particularly true of the major/minor pairs, because the relative distinctiveness of contour and rhythm reversed from Experiment 1 to Experiment 2. Contour and rhythm may in fact be equally salient, so the order of their distinctiveness depends on the particular task being performed. And, indeed, replication of these tasks using other rhythms, contours, and patterns would aid the generalizability of all these results.

In conclusion, naive and trained listeners do very well at comprehending music heard for the first time. Contour and rhythm are very salient, but mode changes escape the notice of the nonmusicians to some extent. Least salient is the relation between a tune and its retrograde, although this is a device used by composers.

REFERENCES

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

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology*, **93**, 551-560.

Berry, W. (1966). *Form in music*. Englewood Cliffs, NJ: Prentice-Hall.

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

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics*, **28**, 381-389.

Dowling, W. J. (1972). Recognition of melodic transformations: Inversion, retrograde and retrograde inversion. *Perception & Psychophysics*, **5**, 417-421.

Dowling, W. J. (1973). Rhythmic groupings and subjective chunks in memory for melodies. *Perception & Psychophysics*, **14**, 37-40.

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

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

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

Halpern, A. R. (in press). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

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

Massaro, D. W., Kallman, H. J., & Kelly, J. L. (1980). The role of tone height, melodic contour, and tone chroma in melody recognition. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 77-90.

Pollard-Gott, L. (1983). Emergence of thematic concepts in repeated listening to music. *Cognitive Psychology*, **15**, 66-94.

Sattath, S., & Tversky, A. (1977). Additive similarity trees. *Psychometrika*, **42**, 319-345.

Shepard, R. N. (1972). Psychological representation of speech sounds. In E. E. David, Jr., & P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill.

NOTES

1. The reader will find it helpful to refer to Figure 1 frequently throughout the paper.

2. In an attempt to statistically validate the musician/nonmusician difference in salience of dimensions, an ANOVA was carried out. The dependent measure was the ratio of error responses preserving the same dimension (e.g., an error response using the same rhythm as the correct response) to total errors committed. Factors were subject group and dimension (mode, rhythm, and direction). The ratio should approach 1.00 if the subject considers the dimension to be highly salient (i.e., only makes errors within and not between items sharing that dimension). The ANOVA on the arcsin-transformed scores showed that musicians had higher ratios [$F(1,20) = 12.16$, $p < .01$] and the dimensions differed [$F(2,40) = 5.20$, $p < .01$]. The crucial interaction of subject group x pair type was not significant [$F(2,20) = 2.57$]; however, the means suggested that both groups found contour to be most salient, whereas the order of salience of rhythm and mode reversed for the two groups, as shown in Figure 3. Because there are, however, several problems in using an ANOVA for these stimuli, this analysis must remain secondary to the ADDTREE solutions as being the result of most interest here.

(Manuscript received August 19, 1983;
revision accepted for publication January 26, 1984.)