COGNITIVE PSYCHOLOGY 5, 322–337 (1973)

The Perception of Interleaved Melodies[1]

W. J. DOWLING[2]
University of California, Los Angeles

If the notes of two melodies whose pitch ranges do not overlap are interleaved in time so that successive tones come from the different melodies, the resulting sequence of tones is perceptually divided into groups that correspond to the two melodies. Such "melodic fission" demonstrates perceptual grouping based on pitch alone, and has been used extensively in music.

Experiment I showed that the identification of interleaved pairs of familiar melodies is possible if their pitch ranges do not overlap, but difficult otherwise. A short-term recognition-memory paradigm (Expt II) showed that interleaving a "background" melody with an unfamiliar melody interferes with same–different judgments regardless of the separation of their pitch ranges, but that range separation attenuates the interference effect. When pitch ranges overlap, listeners can overcome the interference effect and recognize a familiar target melody if the target is prespecified, thereby permitting them to search actively for it (Expt III). But familiarity or prespecification of the interleaved background melody appears not to reduce its interfering effects on same–different judgments concerning unfamiliar target melodies (Expt IV).

Grouping of auditory stimuli into perceptual "channels" and attentional selectivity in relation to such groups has been a major focus of interest in psychology for over a decade (see, e.g., Moray, 1969). In most experiments on the processing of complex messages, however, pitch difference as a basis for grouping has been mixed with other factors such as spatial location, timbre, or loudness as, for example, in the case of two different speakers, male and female. The studies reported here focus on pitch difference *per se* as a basis for grouping, where the stimulus consists of two messages (melodies) whose elements (tones) are temporally interleaved.

[1] This research was supported in part by grant MH-24942 from the U. S. Public Health Service and a grant from the UCLA Academic Senate. Computing facilities were provided by the Center for Cognitive Studies, Harvard University, and the Department of Psychology, UCLA. I thank Donald Norman, Jerome Bruner, Edward Carterette, Saul Sternberg, Irvin Rock, Kelyn Roberts, Caroline Dowling, and Roberta Greer for valuable suggestions and encouragement at various stages of the project.

[2] Requests for reprints should be addressed to Dr. Walter J. Dowling, Department of Psychology, California State University, Los Angeles, California 90032.

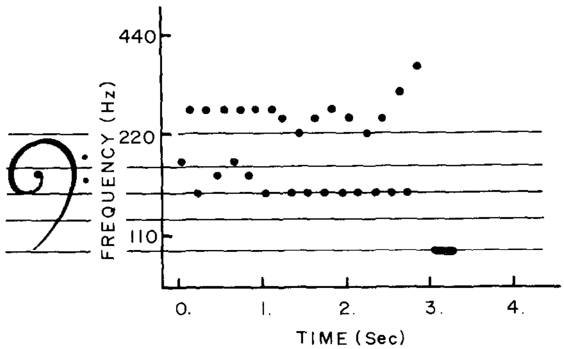

FIG. 1. Passage from J. S. Bach *Suite No. 3 for Cello Solo* (gigue) with fundamental frequency (pitch) plotted against time. A music staff is aligned with the log-frequency scale so that the top line of the staff represents the A below middle-C (220 Hz).

Composers of music have made use of pitch difference to permit the perception of interleaved melodies for the past six centuries, and especially during the Baroque period (ca. 1600–1750). Such passages as that from a Bach solo cello suite shown in Fig. 1 are typical. Here two melodic lines in nonoverlapping pitch ranges have been temporally interleaved. The upper and lower sets of tones split apart perceptually into separate contrapuntal lines, an effect to be called "melodic fission."[3] Here the effect produces a contrapuntal texture with only one instrument playing a single succession of tones. The popularity during the Baroque of music for solo instruments designed primarily to produce one note at a time—flute, violin, cello—coupled with the Baroque style of writing more than one melody at a time in counterpoint led to the extensive use of melodic fission during that period. But examples of interleaved melodies can be found in western music as early as 1325 (The Robertsbridge Fragment, Wooldridge, 1897–1913) and as recently as in the jazz of John Coltrane (Williams, 1967).

Miller and Heise (1950) described an effect related to melodic fission, but with single tones instead of melodies. In measuring what they called the trill threshold they found that as the frequency difference between two tones alternating at 10 tones/sec "is increased progressively, a point is reached at which the trill seems to break; then two unrelated and interrupted tones are heard. It is as if the listener's 'melodic tracking'

[3] Norman (1967) and Dowling (1967, 1968) initially called the phenomenon "rhythmic fission," and that usage seems preferable where the focus is on the perception of distinct rhythmic patterns (van Noorden, 1971). "Melodic fission" is used here because of the focus on distinct melodies.

could not follow a sudden change larger than a certain amount" (Miller & Heise, 1950, p. 637). In describing another case in which an intermittent tone was interpolated in a warbling pattern, Heise and Miller (1951, p. 72) characterize the grouping effect as "quite marked, as if the isolated tone came from a separated sound source completely independent of the background pattern." Miller and Heise found that the frequency separation necessary to produce melodic fission in simple trills is approximately 15%, or a little less than 3 semitones. This finding may provide a perceptual basis for the fact that in their solo sonatas for flute, violin, and cello both Bach and Telemann systematically avoid intervals between successive tones of less than 3 semitones in melodic fission passages (Dowling, 1967). The present experiments extend Miller and Heise's work to the perception of actual melodies, using stimuli much closer to the passage shown in Fig. 1.

Recent experiments by Bregman and Campbell (1971) demonstrate melodic fission phenomena somewhat more complex than those of Miller and Heise. Their first experiment was based on the assumption that the temporal order of tones within perceptual groups is more accurately perceived than the order of tones between groups. Given this assumption, they found that interleaved tones from two separated 8-semitone ranges were perceptually grouped by pitch. Their second experiment used a short-term recognition memory paradigm in which a three-tone target ("standard") pattern was presented alone, and the test ("comparison") pattern was then presented with additional interleaved tones. Recognition accuracy was higher when the three tones of the target pattern were drawn from just one of the two 8-semitone ranges (narrow-range melody, no overlap with range of interleaved background), rather than from both (wide-range melody, overlap with range of interleaved tones). This experiment does not permit distinguishing between the effects of the range of a target melody and effects of overlap between its range and the range of the interleaved background.

Experiment II below carries the observations of Bregman and Campbell further, by varying the overlap between ranges of melody pairs, without varying the ranges themselves.

EXPERIMENT I: IDENTIFICATION OF FAMILIAR INTERLEAVED MELODIES

Experiment I used a method of limits to determine the pitch separation between interleaved melodies necessary for identification.

Stimuli. The stimuli in Expt I were pairs of melodies. The tones of one melody were interleaved with the tones of the other melody as in Fig. 2j. Eight familiar melodies were used; they are shown in standard musical notation in Fig. 2. The tones (quarter notes) alternated with tones from other melodies presented in the rest. Repeated tones occurred wherever the familiar form of the melody had a sustained note, for example, on the third and fourth notes of Fig. 2c. Figure 2j shows "Frere Jacques" (stems down) and "Happy Birthday" (stems up) interleaved for simultaneous presentation. The odd-numbered set of alternate notes belongs to "Frere Jacques"; the even-numbered set to "Happy Birthday." The tones were .125 sec in duration so that the presentation rate of interleaved stimuli was 8 tones/sec.

A PDP-4 computer acting as a square-wave generator produced the stimuli. Machine timing controlled the period of the square wave emitted and an external millisecond clock controlled note durations and presenta-

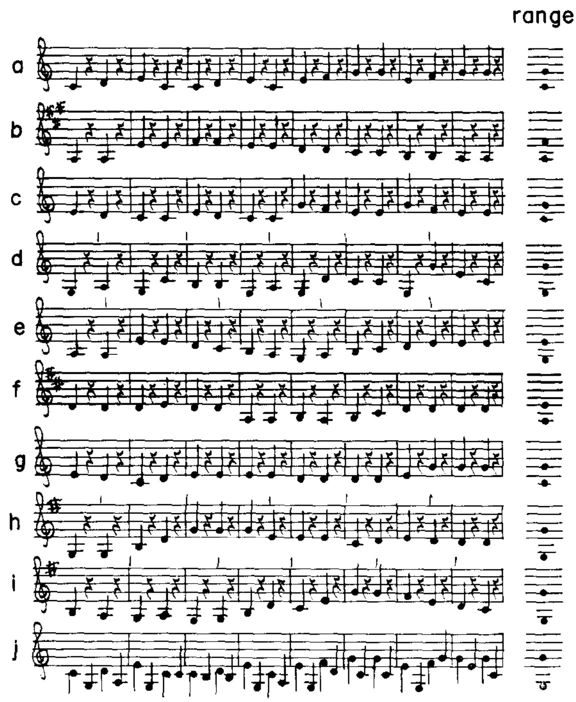

FIG. 2. Familiar melodies used in Expts I and III. (a) "Frere Jacques," (b) "Twinkle, Twinkle Little Star," (c) "Three Blind Mice," (d) "Happy Birthday to You," (e) "God Rest Ye Merry, Gentleman," (f) "Good King Wenceslaus," (g) "Mary Had a Little Lamb," (h) "On Top of Old Smokey," and (i) "Fair Harvard." (j) Combination of melodies (a) and (d), in which the rests in melody (a) are replaced with the notes of melody (d). Combined melodies were presented at the rate of 8 tones/sec. Marks over the staff indicate metric division of melodies with triple meter (d, h, and i) or a "pick-up" note (e).

tion timing. The computer-produced tones sounded very much like the low register of a clarinet. They were presented to *Os* over loudspeakers at comfortable levels. Which melody constituted the odd-numbered notes, and which the even, was randomly determined.

Observers. Six *Os* participated in single sessions. *Os* were Harvard University students (ages 20–25) who volunteered because of their interest in music. Four *Os* were proficient on a musical instrument. All six *Os* possessed recordings of Baroque music diplayed melodic fission effects, typically Bach's *Brandenburg Concertos.* All *Os* professed familiarity with the tunes of Fig. 2.

Procedure. In a pretest, *E* presented each tune for *O* to name. Two *Os* had to be reminded of the names of two and three tunes each. These *Os* accurately identified the tunes they missed in the pretest later in the experiment. *Os* had the list of names of the eight tunes in front of them throughout Expt I. *Os* were told they might hear some of the tunes more than once in the course of the experiment.

On each trial *O* listened to an interleaved pair of melodies presented over a loudspeaker, starting with maximally overlapping ranges (one range completely contained in the other) as in Fig. 2j. A 3-sec interval occurred between presentations. On every fourth presentation one of the melodies was transposed up by a semitone, so that over the series of presentations the frequency separation between the two melodies gradually increased. This increase began to occur with the first transposition for most of the melodies, but was delayed until the second or fourth transposition in the cases of the melodies of Fig. 2b, e, and f because of their narrower ranges. *O* was instructed to respond with the name of the first melody he recognized, and to identify the second melody as soon as possible after that. Correct identification of the second melody always followed correct identification of the first within two presentations, and typically occurred on the presentation immediately following the first identification. The trial ended when *O* correctly identified both melodies. Each *O* heard four to seven pairs this way. Each of the 28 possible pairs of melodies were presented in random order at least once to some *O*, and four were presented twice. The duration of each *O*'s participation was limited by his willingness to serve.

Results. Median presentations to correct identification of both melodies was 32.0 (IQR: 19.5, 41.0).[4] These correspond to transposition of one of the melodies by a median of 10.0 (IQR: 6.0, 13.0) semitones. The amount of transposition required for the melodies to be identified was

[4] Values in parentheses give first and third quartiles; the difference between them gives the size of the interquartile range, IQR.

generally that amount which left the ranges of the melodies not quite overlapping. For example, "Happy Birthday" (in Fig. 2j) would become recognizable when transposed up a little more than an octave—12 to 14 semitones—so that its lowest note would fall slightly above the highest note of "Frere Jacques." The median amount of overlap of the melodies' ranges at identification was −1.2 (IQR: −4.2, 0.8) semitones. (The negative overlap signifies a separation of ranges.) Thus three-fourths of the identifications occurred only after overlap had been eliminated completely. No one *O* contributed notably more than the others to the required range separation, nor did any single melody contribute inordinately. *O*s reported that they could select either the higher or the lower melody to attend to. Whether the melody consisted of odd- or even- numbered notes did not affect the order of identification.

The most plausible rival hypothesis is that number of presentations (or amount of transposition) rather than range overlap is the dominant variable in determining point of identification. Evidence against this alternative is provided by the finding that at the point of identification, variability of the amount of transposition (IQR: 6.0, 13.0 semitones) was greater than variability in the amount of overlap (IQR: −4.2, 0.8 semitones). A second alternative hypothesis is that mean pitch differences between temporally adjacent tones in the melodies would predict which pairs of melodies would be most easily distinguished from each other. However, a rank order correlation between melody pairs ranked for mean pitch separation between members of the pair and pairs ranked for speed of identification was not significantly different from zero.

EXPERIMENT II: SHORT-TERM RECOGNITION MEMORY OF UNFAMILIAR INTERLEAVED MELODIES

Experiment I showed that interleaved melodies can be identified when the separation between the pitch ranges of the two melodies is increased to the point where they no longer overlap. However, in that experiment pitch separation was confounded with number of presentations. This confounding is absent in Expt II, in which a short-term recognition memory paradigm was used to investigate the effects of pitch separation on recognition with a single presentation of each melody.

Stimuli. On each trial of the experiment *O* first heard a standard melody presented at a rate of 3 tones/sec. This standard began on middle-C (262 Hz), was five tones long, and was different on every trial. Each tone of this standard melody was .16 sec long, with .17 sec silent intervals between tones. After a 2-sec pause a comparison melody followed the standard. The comparison always began on the same frequency as the standard and was either identical to the standard or different. A .16-sec tone from a background melody was inserted after each tone of the comparison melody. Thus the combined tone sequence of comparison and background melodies proceeded at a rate of 6 tones/sec. Figure 3a illustrates a typical trial of the experiment in which the comparison (filled symbols) is the same as the standard. The background was chosen from one of two possible patterns (shown in Fig. 3a and 3b, open symbols). The background started on one of three possible frequencies: middle-C (262 Hz, the same as the comparison, shown in Fig. 3a), the F♯ 6 semitones above middle-C (370 Hz shown in Fig. 3b), or the C 12 semitones above middle-C (524 Hz). Only two backgrounds were used to minimize variability in the stimuli. Pilot studies indicated little difference in performance between using only two backgrounds and using four, and also failed to disclose any difference in the disruptive effects of the two backgrounds used.

A Hewlett-Packard 2112B computer generated the stimuli which consisted of sawtooth waves. Stimuli were recorded on tape and played to *Os* over loudspeakers at comfortable levels. On each trial the computer

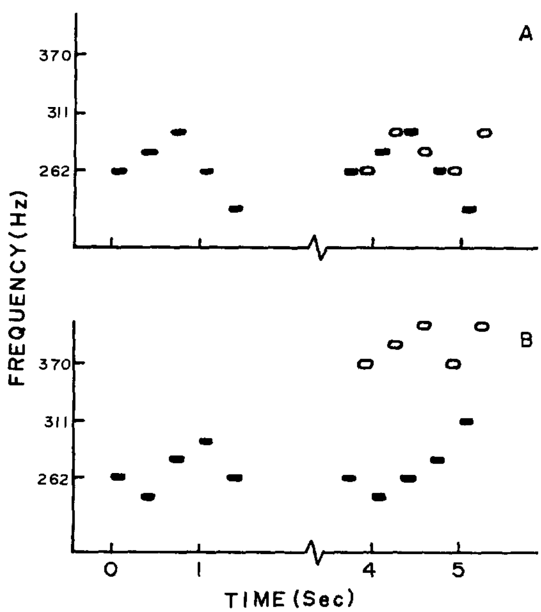

Fig. 3. Examples of trials in Expt II. (a) Trial with comparison melody identical to the standard and with one of the background patterns (open symbols) at zero pitch separation from the comparison. (b) Trial with comparison melody different from the standard, and the other background pattern (open symbols) at 6 semitones separation from the comparison.

generated a new standard melody by starting on middle-C and selecting succeeding notes such that intervals between successive notes occurred with the following probabilities: $p(\pm 1$ semitone$) = .50$, $p(\pm 2$ semitones$) = p(\pm 3$ semitones$) = .25$. Thus at 6 semitones separation the probability of overlap of ranges of comparison and background melodies was very small. Same comparisons were exact repetitions of the standard. Different comparisons were newly selected just as the standards with the stipulation that the melodic contour (the sequence of directions of pitch changes from one note to the next; see Dowling & Fujitani, 1971) of the comparison be different from the contour of the standard. Whether the background or the comparison started first was randomly determined.

Procedure. O's task was to recognize those comparisons which were identical to the standards. On each trial Os responded with one of four confidence levels labeled "Sure Same," "Same," "Different," and "Sure Different." Os had 5 sec to respond. A warning tone (4250 Hz) preceded the onset of the following trial by 2 sec. There were two sessions of 60 trials each. Thirty trials in each session had identical comparison melodies; the other 30, different. One-third (10) of each of these two trial types (same or different) had the background melody starting on the same note as the comparison, or zero separation; one-third at 6 semitones separation; and one-third at 12 semitones separation. Each of the two backgrounds occurred equally often on each of these six types of trial. This produced in all 12 different types of trial defined by type of comparison (same or different), pitch separation of background from comparison (0, 6, or 12 semitones), and background melody (one of two). The 60 trials in each session consisted of five consecutive random permutations of these 12 trial types. The sessions were separated by a 15-min break during which *E* collected data on Os' previous musical training.

The first session was preceded by a thorough explanation of the procedure, including 18 sample trials. The sample trials included eight trials at the largest pitch separation, and two series of five trials each progressing from large to small frequency separation.

Observers. Twelve ULCA undergraduates served in one group session. Os had a mean of 3.75 yr musical training, including lessons on an instrument or voice, or playing in an ensemble. Singing in choirs or taking music appreciation classes was not counted in this measure.

Analysis. Os responded by marking IBM cards which were scored by computer. For each O, areas under the memory operating characteristic (MOC) were computed for each pitch separation and session, and used as an index of recognition performance. (Norman & Wickelgren, 1965, provide a detailed description of a similar data analysis procedure.)

The area under the MOC can be taken as an estimate of the probability of correct response in a two-alternative forced choice procedure.

Results. Table 1 shows mean areas under the MOC for two sessions and three pitch separations. The main effect of pitch separation is statistically reliable ($F(2,22) = 7.36$, $p < .005$). As in Expt I recognition was easier when the ranges of comparison and background stimuli did not overlap. Performance at zero separation was better than chance during the second session ($p < .025$). Neither the main effect of sessions nor the Sessions × Separation interaction were significant. The only statistically significant correlation between performance and musical experience was for 6 semitones separation in session 1 ($r = .87$, $p < .05$).

An important question not answered directly by Expt II is whether an interleaved background whose range does not overlap that of the target melody has any effect on recognition performance. Dowling and Fujitani (1971) in one of their conditions presented pairs of five-note melodies at a rate of 6 tones/sec on trials otherwise identical in timing to those of Expt II. Pitch and timbre characteristics of stimuli were the same, the only difference being the absence of the background melody interleaved with the comparison. The mean area under the MOC for five Ss was .97, which represents considerably better performance than the .69 achieved under the 12 semitones separation condition here. This comparison shows that even at one octave separation the interleaved background interferes considerably with recognition.

The better-than-chance performance in Expt II (and in Expt III, below) in recognizing narrow-range melodies having no pitch separation from the interleaved background may be compared with the finding by Bregman and Campbell (1971, Expt 2) of chance performance with wide-range melodies. The perceptual basis for the fact that melodies throughout the world are restricted to fairly narrow pitch ranges in the sense of having narrow intervals between successive tones (Merriam, 1964; Dowling & Fujitani, 1971) may be that rapid wide-range melodies split apart into separate pitch-defined groups of tones.

TABLE 1
Areas Under the MOC: Experiment II

	Pitch separation (semitones)			
	0	6	12	Mean
Session 1	.57	.72	.69	.66
Session 2	.62	.65	.69	.65
Mean	.59	.69	.69	

EXPERIMENT III: OVERCOMING THE EFFECTS OF INTERLEAVING BY ACTIVE SEARCH FOR A FAMILIAR MELODY

On the last trial of Expt I the two most practiced *Os* continued to listen after correct identification as the pitch separation between the two melodies was reduced. *O* was instructed to attend to one of the melodies and to report when he could no longer hear it distinctly. For both *Os* the melody attended to remained clear and distinct even when the pair of melodies were again presented in their original overlapping ranges. These *Os*' responses on the last trial may be contrasted with their responses on the first presentation of the whole series when they could distinguish neither melody. Experiment II tests whether active search for a familiar target melody in an overlapping background can lead to recognition even when that melody has not been heard immediately before, but has merely been named for *O*. Under these circumstances *O* has only his knowledge of the pitch relations among the notes of the target melody as a basis for distinguishing it from the interleaved background tones.

Observers. The same six *Os* served in Expt III as served in Expt I. All had additional experience in similar studies with interleaved melodies (Dowling, 1967). *Os* served individually in single sessions.

Stimuli. Each *O* had eight trials, randomly divided into five *signal trials* and three *catch* trials. On each signal trial *O* listened to a series of presentations of one of the melodies of Fig. 2a–h, randomly selected, and repeated at 2.4-sec intervals. These familiar target melodies were interleaved with background melodies in the same pitch range. The stimuli were produced by computer and were similar to those of Expt I, except that the pitch ranges of target and background melodies overlapped throughout the series of presentations. The tones of the background melody on each presentation were randomly chosen with equal probability from a subset of tones (a C-major scale) of the same pitch range as the target. The background melody was permitted to change from one presentation to the next in order to provide maximal contrast to the fixed, familiar targets.

On the three catch trials the target named by *E* was absent and another stimulus was substituted. This substitute on one trial was another melody (not one of the eight targets, but the equally familiar "Believe Me if All Those Endearing Young Charms," the Harvard alma mater, Fig. 2i) and on the remaining two trials was a random sequence of tones (like the background) which changed from one presentation to the next. Both kinds of catch trial were included to allow for two possible kinds of response in addition to failing to report, or erroneously reporting, the presence of the target. First, *O* might correctly report noticing the substituted familiar melody, thus indicating accurate perception in spite of the misleading cue. That the substitute would have been noticed immediately in the absence of the interleaved background is suggested by the fact that all *Os* recognized it immediately when it was presented to them without the background after the experiment. Second, *O* might mistake the random substitute target for the absent target melody, indicating a tendency to follow the misleading cue in an ambiguous situation. The random background was changed on every presentation to enhance ambiguity.

Whether the background started first or second on a given presentation was randomly decided. *Os* listened with earphones monaurally with the preferred ear.

Procedure. O's task on each trial was to say whether he recognized the target melody. *E* told *O* the name of the target before the start of each trial, and the trial continued until *O* reported hearing the target distinctly or for 15–17 presentations, whichever occurred first. *Os* responded verbally "yes" or "no" after each presentation. *Os* were instructed to respond "yes" as soon as they recognized the target melody. After every few presentations *E* encouraged *O*'s continued attention with such remarks as, "Are you *sure* you don't hear it yet?" On presentation 15, *E* said, "Can you hear any pattern at all in the tones?" If *O* replied that he had not listened for any pattern other than the target the trial was extended to 17 presentations for *O* to try to hear other patterns.

No feedback of results was provided during the experiments. In particular, *E* did not inform *Os* of the deception until after the experiment was over. After each trial, regardless of outcome, *E* simply said "ok" and proceeded to introduce the next trial. The order of trials was randomized differently for each *O*.

Results. All six *Os* reported hearing distinctly the melodies they were told to listen for on all five trials in which those melodies actually occurred. The mean number of presentations required for recognition on those five trials was 3.6 (IQR: 2.0, 5.7). No *O* reported hearing the target melody on those trials in which it was absent.[5] Since trials with only "no" responses contained at least 15 presentations, stimuli not con-

[5] There are several plausible sets of assumptions under which one might compute the chance probability of this result. If the presence of the target had been irrelevant to *Os*' reports, and it were assumed that *Os* each had three "no" responses to distribute among the trials, the probability of this correlation of responses and trial types for the six *Os* would have been $56^{-6} < 10^{-10}$. Alternatively, one might assume that the *a priori* probability of a "yes" on any trial to be .50, in which case the probability of exactly this result would be $2^{-48} < 10^{-13}$.

taining the expected target melody were repeated well beyond the point at which recognition occurred on signal trials.

No *O* reported hearing any pattern at all on catch trials, whether the substitute melody was familiar or random. All *O*s were surprised at the actual composition of the catch-trial stimuli when it was disclosed to them after the experiment was completed.

Experiment III demonstrates the importance of knowing the target for recognition of a familiar melody when it is interleaved with an overlapping background. Performance on the catch trials shows that when *O* is cued with the name of a melody not contained in the stimulus this prevents a familiar melody from being noticed, but does not lead to false-positive identification. The name appears to have provided a basis for *O* to "search" for the target. This result is analogous to that obtained by Gottschaldt (1926) in a visual recognition task. After familiarizing *O*s at great length with a simple target pattern, Gottschaldt presented the target embedded in a more complex context. *O*s did not spontaneously recognize the familiar target, just as *O*s in the first trials of Expt I did not spontaneously recognize tunes familiar to them since childhood. Only when Gottschaldt told *O*s to search for the hidden figure did they report seeing it. As in Expt III, active search led to perception of the target.

EXPERIMENT IV: EFFECT OF KNOWLEDGE OF THE BACKGROUND ON RECOGNITION

In Expt I, identification of the second melody always followed identification of the first within two presentations. One explanation is that the same condition (elimination of overlap) permitted both identifications. An alternative is that identification of the second was facilitated by knowledge of its background. Experiment IV tests the importance of familiarity with an interleaved melodic background using a recognition memory paradigm as in Expt II. *O*s were given experience identifying the backgrounds separately before doing the recognition memory task. Half of these *O*s were informed on each recognition trial which of the two backgrounds would be employed on that trial (group FI). The other half were not informed but were simply familiar with the backgrounds (group F). A third group became familiar with a different set of backgrounds not used in the recognition memory task (group N).

Stimuli. Stimuli were produced in essentially the same way as in Expt II and were presented over loudspeakers.

Observers. Three groups, each consisting of seven *O*s, served in Expt IV. All groups served in both parts of the experiment. *O*s were drawn from the same pool as those of Expt II.

Procedure. During Part 1 of the experiment all *O*s performed an identification task of 60 trials with feedback. On each trial of Part 1 *O*s had to identify one of two five-note melodies. *O*s listened to the melody at a rate of 3 tones/sec and then had 5 sec to respond either "one" or "two" with pencil on an IBM card. After 5 sec *E*'s recorded voice gave the correct response. The next trial started 2 sec later. *E* monitored *O*s' behavior to make sure *O*s were responding before the feedback. The two melodies of Part 1 were the same as the backgrounds in Expt II (Fig. 3, open symbols) or their retrograde inversions (upside-down backward) beginning on the same note. Groups F and FI used the backgrounds as shown in Fig. 3; group N used the retrograde inversions. Half the trials of Part 1 used one of these backgrounds; the other half, the other, in random order. One-third of the trials for each background began on middle-C, one-third six semitones above middle-C, and one-third 12 semitones above middle-C. All *O*s scored 95% correct or better during Part 1.

During Part 2 *O*s did a recognition memory task of 60 trials like that of one session of Expt II. The only differences were that the warning tone on each trial preceded the standard melody by 4 sec, and that the warning tone for group FI was followed immediately by *E*'s recorded voice naming the background ("one" or "two") used in the subsequent trials.

Results. Table 2 shows mean areas under the MOC for the three groups in Expt IV for the three degrees of background separation. Analysis of variance disclosed no significant results for either main effects or interactions. Thus there is no evidence that knowledge of background facilitates recognition. Indeed, there is a hint that such knowledge may actually hurt performance: the failure to replicate the pitch separation effect of Expt II could be explained if a familiar background produced distraction at large separations. This conjecture is supported by the observation that group N, for which the backgrounds were

TABLE 2
Areas Under the MOC: Experiment IV

	Pitch separation (semitones)			
Condition	0	6	12	Mean
FI	.62	.70	.66	.66
F	.71	.73	.67	.70
N	.69	.70	.75	.72
Mean	.67	.71	.69	

not familiar as in Expt II, showed the clearest trend with pitch separation, as well as the best overall performance.

ADDITIONAL OBSERVATIONS AND CONJECTURES

Melodic fission and presentation rate. One issue arising in studies of melodic fission is its dependence on rate of presentation, a factor not varied in the present series of experiments. Bregman (1971) has studied this factor with a method similar to that of Bregman and Campbell (1971, Expt II). He found that a melody with a small pitch range (as in the present experiments) that is separated from the range of the interleaved background is easily perceived at rates from 1.4 to 20.0 tones/sec. There may, however, be optimal rates within this range for producing melodic fission effects with stimuli such as the passage in Fig. 1. The upper limit may be governed by the rate at which individual tones can be perceived; the lower limit may be reached when attention can be switched between one melody and the other as fast as the tones are presented. As that lower limit is approached it should become easier to fuse the pitch-separated melodies into a single pattern. Van Noorden (1971) has shown in a task similar to Miller and Heise's (1950) that if Ss are given a set for hearing fission between the sets of alternating tones, a fission threshold comparable to Miller and Heise's is obtained for presentation rates of 6.7–20.0 tones/sec. However, if Ss have a set for hearing fusion between the sets of tones at the greatest possible separation, the threshold for fission increases from a little more than the Miller and Heise trill threshold (3 semitones) at 20 tones/sec to about an octave at 6.7 tones/sec. *Os* in the present experiments (all recognition tasks) could be said to have a set of hearing fission at the least separation possible. A survey of 20 recordings of Baroque music for piccolo, flute, violin, and clavier disclosed a median presentation rate in melodic fission passages of 6.3 (IQR: 4.7, 8.2) tones/sec (Dowling, 1967). The rates of 6 and 8 tones/sec used in the present series of experiments falls within this range.

Loudness of target versus background. Neisser (1967, p. 212) points to the importance of distinguishing between the loudness of a stimulus and its vividness (cf. James, 1890, pp. 425f). Stimuli which are attended to are more vivid but not necessarily louder. To provide an empirical basis for the distinction Neisser suggests that *Os* could be asked to judge the loudness of stimuli presented to the rejected channel compared with the loudness of attended-to stimuli.

The six experienced *Os* of Expts I and III listened binaurally with earphones to pairs of melodies like those of Expt III. *Os* were instructed to attend to a familiar, prespecified melody which was on successive series of trials 5 dB more intense than, 5 dB less intense than, or equal to a rejected random background with overlapping pitch ranges. A random background which changed with each presentation was used to facilitate attention to the melody. *E* found it almost impossible to attend to such a background but easy to attend to the melody, and this was corroborated by *Os*' comments. A 125-msec comparison tone followed each presentation by 1 sec. There were four trials at each of the three intensities. *O* made an ascending and a descending setting of the intensity of the comparison tone to match the loudness of either the melody or the background, each on half the trials. *O* made each adjustment over several presentations spaced at 5-sec intervals, and the trial ended when *O* was satisfied. After 12 trials with melodies *O* made four more settings with a 4-sec steady tone replacing the 4-sec melody-background combination.

Os made their settings with the same accuracy and precision on each type of trial. There was a constant error of -3 dB which may be attributed to the time delay and the brevity of the comparison tone. Thus there appears to be no effect on their loudness of whether tones are contained in an attended melody or an unattended background (Dowling, 1967).

Analogies with ambiguous visual stimuli. The later trials of Expt I provide an auditory analog of the Necker cube. When the two melodies were in separate pitch ranges *Os* reported that they could attend to either one of them with the other as background, but not both simultaneously. Attention could easily be switched back and forth between them; and if *O* did not focus his attention on one of them it would, after a short time, switch abruptly and automatically to the other. This alternation between a pair of possible experiences of the same stimulus is characteristic of ambiguous figures such as the Necker cube and the Schröder staircase (Woodworth & Schlosberg, 1954).

Context effects. Guilford and Nelson (1937) found a marked contrast effect in the recognition memory for tones in two-note melodies, an effect which became more pronounced as the size of the melodic interval was increased to 6 or 8 semitones. Since fairly small pitch distortions of familiar tunes are sufficient to reduce their recognizability considerably (Dowling & Fujitani, 1971), *Os* in Expts I and III must have been able to ignore context effects on pitch produced by the interleaved background tones. Grouping may reduce context effects between tones that are perceived in different groups.

Effects of knowledge of musical form on perception of music. Experiment III suggests that active search for a well-known melody can lead to discerning it in a confusing context when it would go unnoticed by the passive listener. This suggests that the expectations of the listener who knows the structural principles of a form will lead him to hear different things in a piece than will be heard by the naive listener. The listener who knows the typical pattern of recurrences of a theme in a fugue, a rondo, adsonata-allegro, a Balinese dance, or an Indian raga, is able to perceive that theme more easily than the listener who doesn't know what to expect.

REFERENCES

BREGMAN, A. S. *Primary auditory stream segregation and the perception of tunes.* Unpub. mimeo. Ms., Department of Psychology, McGill University, Montreal, 1971.

BREGMAN, A. S., & CAMPBELL, J. Primary auditory stream segregation and perception of order in rapid sequence of tones. *Journal of Experimental Psychology,* 1971, **89**, 244–249.

DOWLING, W. J. *Rhythmic fission and the perceptual organization of tone sequences.* Unpub. Ph.D. Thesis, Harvard University, 1967.

DOWLING, W. J. Rhythmic fission and perceptual organization (Abstr.). *Journal of the Acoustical Society of America,* 1968, **44**, 369.

DOWLING, W. J., & FUJITANI, D. S. Contour, interval and pitch recognition in memory for melodies. *Journal of the Acoustical Society of America,* 1971, **49**, 524–531.

GOTTSCHALDT, K. Über den Einfluss der Ehrfahrung auf die Wahrnehmung von Figuren. *Psychologische Forschung,* 1926, **8**, 261–317. Excerpts translated in M. D. Vernon (Ed.), *Experiments in visual perception.* Baltimore: Penguin Books, 1966, pp. 29–44.

GUILFORD, J. P., & NELSON, H. M. The pitch of tones in melodies as compared with single tones. *Journal of Experimental Psychology,* 1937, **20**, 309–335.

HEISE, G. A., & MILLER, G. A. An experimental study of auditory patterns. *American Journal of Psychology,* 1951, **64**, 68–77.

JAMES, W. *Principles of psychology.* Vol. I. New York: Holt, 1890.

MERRIAM, A. P. *The anthropology of music.* Evanston, Ill.: Northwestern University Press, 1964.

MILLER, G. A., & HEISE, G. A. The trill threshold. *Journal of the Acoustical Society of America,* 1950, **22**, 637–638.

MORAY, N. *Attention.* London: Hutchinson Educational, 1969.

NEISSER, U. *Cognitive psychology.* New York: Appleton-Century-Crofts, 1967.

NOORDEN, L. P. A. S. VAN. Rhythmic fission as a function of tone rate. *IPO Annual Progress Report,* 1971, **6**, 9–12.

NORMAN, D. A. Temporal confusions and limited capacity processors. *Acta Psychologica,* 1967, **27**, 85–94.

NORMAN, D. A., & WICKELGREN, W. A. Short term recognition memory for digits and pairs of digits. *Journal of Experimental Psychology,* 1965, **70**, 479–489.

WILLIAMS, M. The legacy of John Coltrane. *Saturday Review,* 1957, **50**(36), 1969.

WOODWORTH, R. S., & SCHLOSBERG, H. *Experimental psychology.* New York: Holt, Rinehart and Winston, 1954.

WOOLDRIDGE, H. E. (Ed.) *Early English harmony.* 2 vols. London: Bernard Quaritch, 1897–1913.

(Accepted May 29, 1973)