Aiming attention in pitch and time in the perception of interleaved melodies

W. JAY DOWLING, KITTY MEI-TAK LUNG, and SUSAN HERRBOLD
University of Texas at Dallas, Richardson, Texas

Listeners succeeded in following a melody interleaved at 6 or 8 notes/sec with distractor notes in the same pitch range and of the same timbre. Their ability to perform this auditory "hidden-figures" task depended on the rhythmic control of attention on the basis of expectancies developed through perceptual learning with melodies in the listeners' culture. Listeners appear to have aimed expectancies in pitch and time at regions where events critical to the identification of melodies are likely to occur—regions defining "expectancy windows" through which target notes are perceived. Events in pitch and time regions outside these expectancy windows were not perceived as accurately as events within the window. Listeners discerned interleaved melodies whose notes fell *on* the consistent temporal beat of a pattern better than they did melodies whose notes fell *off* the beat. Expectancies could also be aimed *off* the beat: expected target notes occurring *off* the beat in a syncopated rhythm were judged more accurately than unexpected notes occurring *on* the beat. Listeners found it more difficult to judge the pitch of target notes that fell outside the expected pitch region than that of notes within the expected region. The interleaved distractor notes appear to be instrumental in narrowing attention to within the expectancy windows. When the interleaved distractors were removed, unexpected notes became *more* salient than expected ones.

In the 19th century, the active nature of vision was commonly contrasted with the passivity of hearing. Vision involved the active seeking out of stimuli, whereas hearing was characterized by passive reception. Schopenhauer (1844/1966) even went so far as to attribute the effectiveness of music to the passivity of hearing:

> The effects of music on the mind, so penetrating, so immediate, so unfailing, ... are explained by the *passive* nature of hearing.... The vibrations of the tones following in combined, rational, numerical relations, set the brain-fibres themselves vibrating in a similar way. (Vol. 2, p. 31)

Schopenhauer went on to contrast hearing with vision:

> It is just by reason of the *active* nature of the sense of sight that it is exceedingly keen in the case of hunting animals, that is, beasts of prey. (Vol. 2, p. 31)

We now realize that hearing also is active. The latter quote from Schopenhauer evokes the image of a visual hidden-figures task of the "find six lions in the jungle" sort. This article explores an auditory version of the hidden-figures task, in which the notes of a familiar melody are temporally interleaved with distractor notes in the same pitch

range and of the same timbre (Figure 1). To a listener who does not know what melody to search for, the combined rapid sequence sounds like a meaningless jumble of notes. For an active and informed listener, however, the familiar tune can be heard clearly in the midst of the interfering distractors. As long as listeners know what tune to listen for, and are familiar with that tune, they can attend to it successfully in the confusing context (Dowling, 1973a).

Although the alternation of notes between target tune and distractors is not common in music, selective attention to a melody woven into a complex musical texture is often required. Comprehending music usually depends on active attention to a series of events spread out in time (Pick, 1979). Actual music and the interleaved-melodies task both require the perception of temporally extended events, and thus differ from the static, visual hidden-figures task. Both resemble the selective-looking tasks described by Neisser (1979), in which a viewer attends to just one of two event sequences simultaneously superimposed on a video screen. In the condition closest to the present task, for example, two basketball games played by the same people were intermingled on the screen. Viewers were able to attend to one ball game at a time and ignore the other, as indicated by their pressing a key whenever the ball was thrown in the target game. In both the interleaved-melodies task and the selective-looking task, attention can be directed at a meaningful event, and elements of that event can be discerned accurately even though no simple perceptual filter could be imagined that would account for the selection of relevant stimulus features. The notes of Dowling's (1973a) interleaved melodies fell in the same pitch range and were of the same
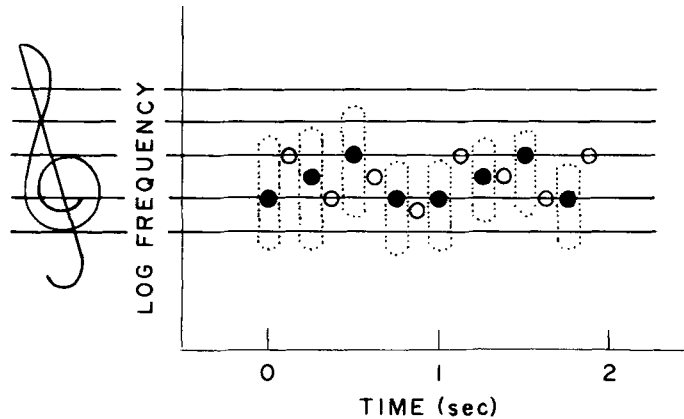
Figure 1. The notes of "Frère Jacques" (filled symbols) interleaved with distractor notes (open symbols). The ordinate is a logarithmic scale of fundamental frequencies on which the initial "G" of the target melody represents 393 Hz; the horizontal stave lines fall at 3 and 4 semitone intervals. A series of hypothesized "expectancy windows" (dotted lines) are aimed at the notes of the target "Frère Jacques." Events outside the windows are more difficult to perceive than events inside them.

loudness and timbre as those of the backgrounds. The events in Neisser's basketball games were spatially intermingled and the players were dressed the same. As Neisser (1979, p. 210) said, "Perceivers can follow the action of a natural event with kinetic information alone"—and listeners can follow a musical event with melodic and rhythmic information alone.

Selective attention in Neisser's and in Gibson's (1978) view is guided by the stimulus structure available to the perceiver. Rhythmic control of attention so as to pick up important events in the stimulus structure contrasts with attentional mechanisms that rely entirely on sensory filtering on the basis of simple stimulus features such as pitch or location. As Spelke (1979) has shown, such rhythmic control of attention is available for children even early in infancy. With perceptual learning, people acquire the ability to respond to more and more elaborate temporal patterns. One mechanism by which stimulus structure affects selective attention is in the development of expectancies (or, as Neisser called them, "specific anticipatory schemata") through perceptual learning.

The rapid auditory sequences involved in the interleaved-melodies task provide us, in their comparative simplicity, with the possibility of investigating in detail aspects of selective event perception that would be more difficult to address with events on the video screen. The stimulus is easier to describe, both in its temporal detail and in the structural invariants it shares with other melodic patterns in the culture. Furthermore, the auditory events provide us with the opportunity to study the effects of perceptual learning, both because rapid interleaved patterns are rare enough for listeners hardly ever to have heard them before entering the experiments and because listeners vary in amount of formal musical training.

Expectancy Windows

In the following experiments we explore perceptual effects of listener expectancies. The listener's knowledge of a particular tune guides attention to focus on just those times and pitches at which critical events for tune identification should occur, consonant with Neisser's (1979, p. 205) suggestion that "perception is based on continuously changing and structured anticipations." It is as though the listener had a series of "expectancy windows" aimed more or less accurately in pitch and time, through which expected events—target notes—could be clearly perceived. Such expectancy windows are suggested by the dotted lines around target notes in Figure 1. Events inside the windows are perceived relatively accurately, while events outside them tend to become lost perceptually. The fate of notes at unexpected pitches and times is similar to the fate of unexpected events in Neisser's video display. When a lady carrying an open umbrella crossed the screen during the game, viewers tended to miss her. Similarly, notes outside the pitch and time windows defined by expected melody events tend to be missed.

Listeners can attend to critical events in the interleaved melodies task because they have expectancies based on knowledge of the pitch range and times at which those events are likely to occur. Such effects of pitch and time expectancies have been demonstrated previously for simpler tasks. Massaro (1975) and Sekuler and Blake (1985) cite numerous studies demonstrating that, for a variety of sense modalities, knowing what to expect improves sensory performance. Kahneman (1973, p. 176) described what he calls the "expectancy effect," in which processing is most efficient for signals arriving exactly when they are expected.

For pitch in particular, Green and Swets (1966, p. 327ff) cited evidence showing that performance is better when a listener in a detection experiment knows what pitch to expect than when pitches are unexpected. Watson and Kelly (1981) described a similar effect for pitches embedded in melodies. Indeed, not long after Schopenhauer, Helmholtz (1877/1954) described such an effect in qualitative terms in pointing out the efficacy of focusing attention on the pitch of an upper harmonic of a complex tone that one is trying to "hear out" of the blended whole:

> In commencing to observe the upper partial tones, it is advisable just before producing the musical tone itself which you wish to analyze, to sound the note you wish to distinguish in it, very gently, and if possible in the same quality of tone as the compound itself. (p. 50)

Immersed in music, Helmholtz was well aware of the importance of active search guided by expectancies in auditory perception.

Temporal expectancies, as well as pitch expectancies, are important in simple detection tasks. Green and Swets (1966, p. 265ff) reported studies by Egan, Greenberg, and Schulman showing that a listener's detection performance improves the shorter the time interval within which a target tone might occur. Similarly, a warning signal that the target event is about to occur is effective in improving performance.

Music is rhythmically organized, and that organization helps determine the perceptual groupings the listener makes. At the organizational level of musical phrases, of the order of seconds in length, rhythmic grouping guides the listener in parsing the melodic event into perceptual units that are stored and remembered as wholes (Deutsch, 1980; Dowling, 1973b). On the level of finer-grained rhythmic organization of events measured in 10ths or 100ths of a second, perceptual grouping and the control of attention also take advantage of regularities of pattern structure. Jones, Kidd, and Wetzel (1981) set listeners the task of discerning a two-note pattern in the midst of a confusing context. The target pattern was presented at a rate of 12.5 notes/sec, and was surrounded by distractor notes in the same pitch range that proceeded either at the same rate as the target or at one-third that rate (4.17 notes/sec). Jones et al. found better performance when the target was rhythmically differentiated from the distractor sequence than when it was not (62% vs. 48% correct). Other conditions in which the pitch range of the distractor notes was varied provided further evidence of the effects of rhythmic grouping on the listener's ability to focus attention on the target. As Jones (1976) suggested, the listener uses both higher and lower order rhythmic patterns to control attention and pick up critical information from stimulus events.

Experiment 1 was designed to demonstrate listeners' reliance on temporal expectancies in discerning a familiar tune interleaved with distractor notes. When the distractor notes were remote in pitch, it was relatively easy to hear the tune. When the distractor notes were interleaved in the same pitch range as the target, it was more difficult but relatively easier when the notes of the target occurred *on* the implicit beat of the stimulus pattern than when they occurred *off* the beat. That is, a pattern organization that facilitated the aiming of attention in time made the melody easier to perceive. It is as though the listener picked out the notes of the target by aiming a grating of expectancy windows at critical points in time, as suggested in Figure 1. Performance is better when those points are easy ones upon which to align the grating.

Experiment 2 was designed to demonstrate the restriction of the expectancy window in the pitch dimension. Listeners were familiarized with a brief melody and then instructed to notice when a particular one of its pitches moved up or down, when the melody was presented interleaved with distractors. On most trials the target pitch stayed within the range of the rest of the melody, and performance was relatively good. However, performance declined when the target note moved to more remote pitches, outside the pitch region in which it had been expected.

What we have been calling an "expectancy window" refers to listeners' ability to control their attention by bringing it into line with the pattern invariants of a particular well-known tune—the pattern of pitches and their timings that occurs whenever the tune is played. Perceptual learning (beyond the original learning of the familiar melody) is required to do that. Dowling (1973a) found that listeners experienced with interleaved-melodies tasks had little difficulty detecting the presence of a cued target melody. However, listeners with no experience on these tasks require an opportunity for perceptual learning with interleaved patterns before they perform well enough to attempt the tasks posed in the experiments. Therefore, the typical experiment was conducted in two sessions: a warm-up session (which sometimes ended with a relatively simple experimental task) and a session in which data from more difficult tasks were collected.

Experiments 3 and 4 provided control conditions with which to compare the results of Experiment 2. In Experiments 3 and 4, listeners performed the same tasks as in Experiment 2, but without interleaved distractor notes in the stimuli. The result was that without distractors the remote pitches outside the expectancy windows became highly salient and easy to judge correctly. The presence of the distractor notes forced the listener to focus the field of attention more narrowly, with the result that events outside that field were sometimes lost to perception.

In Experiments 1 and 2, expectancies of notes of target melodies coincided with the implicit beat in the rhythmic structure of the melodies. Experiments 5 and 6 demonstrated that expectancies can focus on target notes even when those notes occur *off* the beat, provided those note locations are sufficiently probable.

The methods used in the experiments were very similar, and so before introducing the particular experiments we will provide a description of features common to all of them.

GENERAL METHOD

Subjects

Undergraduates at the University of Texas at Dallas served in the experiments for partial course credit. About half of the subjects were categorized as musically experienced, having had 2 or more years of formal musical training (defined as lessons on an instrument or voice, or ensemble participation, but excluding music classes and choral singing). Experienced subjects had a mean of 5.2 years of training. "Experienced" here thus refers to moderate levels of training, compared to professional levels. The typical experienced subject had had about 5 years of music lessons 15 years previously. Inexperienced subjects had had less than 2 years of training (mean = .46 year). The mean age of the subjects was 32.4 years, and 73% were female. Subjects who failed to reach criterion during the warm-up phase of each experiment were dropped from further consideration, as noted below.

Stimuli

A Commodore 64 computer produced the stimuli using its 6581 sound interface device. Stimuli were presented to subjects on-line by the computer, binaurally, at comfortable levels via stereophonic amplifier and headphones (Experiment 1), or they were recorded on tape and presented to groups of subjects via loudspeaker. Distractor notes were temporally interleaved among the notes of the target melodies, so that test stimuli alternated between target notes and distractors. In Experiment 1, timing was controlled by the computer with an accuracy of .017 sec; in subsequent experiments, timing was controlled to better than .001-sec accuracy by means of an assembly-language program component and the stratagem of turning off the CRT video display during stimulus generation. Interleaved stimuli were produced at rates of 6 or 8 notes/sec. Unless otherwise noted, the tones were generated using a triangular waveform. All target tunes were in the key of C major and used middle C as the tonic. Middle C was tuned to a fundamental frequency of 256 Hz in Experiment 1 and to 262 Hz thereafter. Fundamental frequencies of other pitches were tuned in accordance with equal temperament.

Procedure

Each subject served in two sessions, usually spaced 2 or 3 days apart, but sometimes up to a week apart. The first session typically consisted of an exposure and warm-up phase followed by data collection for a simple task; the second session consisted of further data collection. In Experiment 1, each session was introduced with the same warm-up procedure.

EXPERIMENT 1

Experiment 1 was in part a replication of Dowling's (1973a) Experiment 3. In that study, after being cued with the title to a familiar tune, relatively experienced listeners heard either that tune or another one interleaved with distractor notes in the same pitch range. The listeners' task, which they performed very well, was to report whether the target melody was present or absent. In that study, the target was presented sometimes as odd-numbered notes in the interleaved sequence and sometimes as even-numbered notes. It occurred to us that a target presented as odd-numbered notes might be easier to discern, since those notes would fall on stronger implicit beats in the whole pattern. If the listener were in effect aiming a series of expectancy windows at the tones sequence, as depicted in Figure 1, then that series of windows might

be easier to align with the series of beats than with the gaps between the beats. This was suggested to us because students learning to play musical instruments typically find it easier to play a sequence of notes on the beat than to play a sequence off the beat. Tuba players learning the sequence of OOMs in an OOM-PAH-OOM-PAH pattern have a much easier time of it than French horn players learning to produce the sequence of PAHs. If motor behavior is easier to control when aimed at stronger temporal beats, then it seemed plausible that covert attentional processes might be similarly easier to control. Therefore, in each session of Experiment 1, we presented half the target melodies on the beat (odd-numbered notes) and half off the beat (even-numbered notes). This is illustrated in Figure 2, along with the other conditions included in the design.

We did not expect our listeners' performance to be as close to perfect as that of Dowling's (1973a) listeners, who had had considerable experience in performing tasks with interleaved melodies. Dowling's task was easier than ours, too, since he presented interleaved targets over and over until the listener responded, whereas we presented the pattern once and required a response. Therefore, in order to compare the performance of our less experienced listeners on the difficult task of detecting a target in the midst of distractors interleaved in the same pitch range with performance on a task in which the distractors were in a separate pitch range, we included for one group of listeners a condition in which the interleaved distractor notes were outside the pitch range of the target. Unlike previous studies (Dowling, 1973a, Experiment 1), the distractors outside target pitch range straddled the range of the target (as shown in Figure 2B). Listeners found it relatively easy to focus their attention on the range of the target, with distractors above and below.

A third variable that we included in the design of Experiment 1, but that proved to have little effect on performance, was the tonality of the distractor notes in relation to that of the target. We thought that if effectiveness in focusing attention on the target were due in part to attuning expectancies to the pitch set of the target's key, then "nontonal" distractors in a different key from that of the target might be more disruptive than "tonal" ones in the same key, because nontonal distractors would jar the listener's concentration on the target key. Alternatively, the nontonal distractors might have been easier to ignore, since they were more distinguishable (in terms of the feature of key membership) from the notes of the target. Perhaps both of these mechanisms were operating, but we found no consistent evidence for effects of tonality of distractors.

Because, in Experiment 1, we ran subjects individually with on-line stimulus generation and data collection by the computer, it was convenient to measure reaction times (RTs). The RT data were, however, quite noisy, given the small number of trials in a session and the differences in time within melodies at which our target melodies differentiated themselves from one another. No consis-
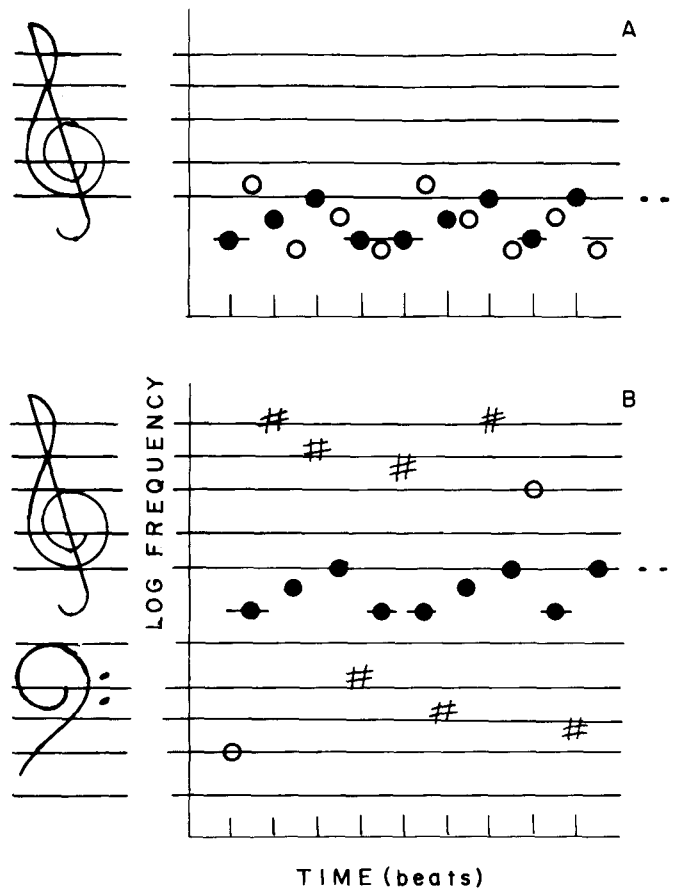
TIME (beats)

Figure 2. Sample stimuli from Experiment 1, illustrating different conditions: (A) target notes (filled symbols) *on* the beat, with tonal distractors (open symbols) interleaved in the same pitch range as the target; and (B) target notes *off* the beat, with nontonal distractors (open symbols and # signs) in separate pitch ranges. Implicit beats are indicated on the abscissa. (Targets were actually 16 beats long, with sustained notes in the familiar versions broken into separate notes on successive beats.)

tent and interpretable results arose from the RT data except for such obvious overall effects as longer RTs to reject a lure than to accept a target. We did check to make sure that the differences in proportion of correct responses reported for Experiment 1 were not the systematic beneficiaries of speed-accuracy tradeoffs, and they were not. (In general, the more difficult judgments in terms of accuracy took longer to make.) Therefore, we shall not report the RT data here.

Method

Subjects. Twenty subjects served in Experiment 1; 12 were musically experienced and 8 were inexperienced. Seven additional subjects (3 experienced and 4 inexperienced) failed to complete the pretest described below, and were dropped from the experiment.

Stimuli. The target stimuli consisted of the first 16 beats of each of eight familiar tunes: "Frère Jacques," "Twinkle, Twinkle, Little Star," "Mary Had a Little Lamb," "Three Blind Mice," "Yankee Doodle," "Old Macdonald Had a Farm," "I Dream of Jeannie with the Light Brown Hair," and "London Bridge Is Falling Down." Whether interleaved with distractor notes or not, these tunes

proceeded at a rate of 4 notes/sec. When interleaved the stimulus presentation rate was 8 notes/sec. Each note was 75 msec in duration, with an onset time of about 24 msec followed by a decay to a level .27 of peak amplitude (decay constant = 48 msec). Inter-note intervals were 50 msec. All tunes were in the key of C major, used middle C as the tonic, and had an overall range extending from the G below middle C to the A above. The computer presented instructions and melody cues visually via its CRT video display.

The conditions of Experiment 1 depended on the pattern of distractor notes, which varied in three ways. First, targets were either on or off the beat. When target notes were presented on the beat, they occupied note positions 1, 3, 5, and so forth (with distractors in positions 2, 4, 6, etc.). When target notes were off the beat, they occupied positions 2, 4, 6, and so forth (with distractors in positions 1, 3, 5, etc.). Equal numbers of trials within each session, randomly determined, had targets on and off the beat, so this was a within-subjects comparison.

Second, the range of distractors was either in or out of the pitch range of targets. Distractors in the same pitch range as targets consisted of two random permutations of 8 pitches in the range from F♯ below middle C to the G♯ above. Distractors outside the range of targets consisted of two random permutations of 8 notes occupying a pair of ranges straddling that pitch range (from C up to F♯ in the octave below middle C, and from C to F♯ in the second octave

above middle C). Different groups of subjects performed the test sessions with distractors in or out of the target range, making that a between-groups comparison. Nine subjects served in the distractor-in condition, and 11 served in the distractor-out condition, with approximately equal numbers of experienced and inexperienced subjects in each group.

Third, distractors either belonged to the same key (C major) as the target melodies (tonal distractors) or were foreign to it (nontonal distractors). Tonal distractors in the range of targets were generated from the set of 8 pitches between the G below middle C and the G above (G, A, B, C, D, E, F, G). Nontonal distractors in that range consisted of a set of pitches each 1 semitone lower than those of the tonal set, excluding the E and adding a G♯ at the top (F♮, G♯, A♯, B, C♯, D♯, F♮, G♮). Tonal distractors out of the range of targets were generated from 8 pitches in the range of C to F in the octave below middle C and in the second octave above middle C (C, D, E, F in each octave). Nontonal distractors out of target range consisted of 8 pitches each 1 semitone lower than the tonal pitches, but excluding the E and adding an F♯ (B, C♯, D♯, F♯ in each octave). Each subject served in two sessions, and in each session the distractors were either tonal or nontonal. The order of tonality conditions was counterbalanced across subjects.

Procedure. Each subject served in two individual sessions, spaced 2 or 3 days apart. Each session consisted of an exposure phase, a pretest, and a test phase. The sessions were identical in form, except for the tonality of distractor tones.

In the *exposure* phase, the computer displayed the titles of the eight target melodies one at a time, and while the title was displayed played the melody interleaved with distractors that were remote in pitch (consisting of a repeated C 2 octaves below middle C). The target melodies were easy to recognize under those conditions. The subjects were instructed to press key "f1" at the right-hand end of the keyboard as soon as they recognized each tune.

The *pretest* consisted of from one to four blocks of eight trials each. On each trial, a visual title cue (consisting of the title of one of the tunes) was presented and remained on the screen for the duration of the auditory stimulus. One second later, a fanfare beginning on middle C (C—E—G—E—C) was presented at 2 notes/sec to establish the tonality. After a 2.5-sec pause, a test stimulus was presented, consisting of one of the eight target melodies interleaved with tonal distractors remote in pitch (the low C). Each title cue appeared once in each block, in random order. On half the trials, randomly selected, the target matched the title cue, and on the other half, it did not (but was a random selection from the other seven possible targets). The subject was instructed to hold his/her right thumb and forefinger over keys "f1" and "f7" at the right end of the keyboard, and to press key "f1" (at the upper right—forefinger) if the target matched the title and to press key "f7" (lower right—thumb) if it did not. The computer provided feedback consisting of the words "Correct" or "Incorrect" for 6 sec following each trial. Subjects who scored at least seven trials correct in the first block of the pretest proceeded to the test phase. Otherwise they repeated the pretest. Subjects who did not achieve the seven-correct criterion by the fourth block of the pretest were dismissed. Twenty subjects achieved the criterion in the first or second block of their first session (and also in the second session), and seven did not achieve it by the end of the fourth block.

The *test* phase consisted of 40 trials arranged in five blocks of 8 trials each. Each block consisted of a random permutation of the eight possible targets, as in the pretest. On half the trials within each block, the target presented matched the visual title cue. Presentation was organized in the same way as in the pretest, except that distractors were generated according to the conditions of the experiment. As noted above, distractors were either in or out of the range of the targets, and different subjects performed the task in each of those conditions. Tonality of distractors was varied between each subject's pair of sessions, and half the subjects performed the

two conditions in each of the two orders, determined blindly. Timing of target (on vs. off the beat) varied from trial to trial within the session. Within each block, half the trials of each type (matching or mismatching the title cue) had targets on the beat and the other half were off the beat, determined randomly. The subjects responded and were given feedback as in the pretest, and the computer recorded the data.

Results

From each subject's proportions of hits and false alarms under each stimulus condition, we calculated A' as an unbiased estimate of the proportion of correct responses, where chance performance equals .50 (Grier, 1971). Those values of A' were evaluated by means of a four-way analysis of variance (ANOVA): 2 experience levels (between groups) × 2 distractor ranges (in vs. out—between groups) × 2 target timings (on vs. off the beat—within subjects) × 2 tonalities of distractor (tonal vs. nontonal—within subjects). The main effects of distractor range $[F(1,16) = 8.02, p < .02]$ and target timing $[F(1,16) = 12.08, p < .01]$ were significant, as was the interaction between them $[F(1,16) = 6.25, p < .025]$. No other effects were statistically significant. The task was easier with distractors outside of target range. Off-beat targets were more difficult to identify than on-beat ones, and that was especially true with distractors inside the target range. In fact, as can be seen in Table 1, there was hardly any effect of target timing when distractors were in a separate pitch range.

Discussion

When distractor notes were in pitch ranges separate from target melodies, listeners performed quite accurately $(A' = .89)$ whether or not the notes of the target fell on the beat with respect to the distractors. In that case, the target established its own temporal scheme and temporal coherence between target and distractors was lost. This result is similar to the loss of temporal coherence between auditory events in different pitch ranges described for several types of auditory pattern by van Noorden (1975).

However, when distractors were interleaved in the same pitch range as targets, performance declined $(A' = .78)$, although it remained much better than chance. In that case, it was important that target notes fall on the implicit beat in the combined stimulus pattern, where listeners could aim their expectancies more easily. When distractors moved into the same pitch range, performance for on-beat targets declined by only .06 but performance of off-beat targets declined by .17. This suggests that listeners were using their knowledge of the temporal organization

Table 1
Performance (A') as a Function of Target Timing and Pitch Range of Distractors in Experiment 1

	Target Timing	
	On Beat	*Off* Beat
Distractor Range:		
Inside Target Range	.84	.71
Outside Target Range	.90	.88

of the target melody to aim their expectancies at times when target notes should occur. Then they could match the expected pitches of the target with the pitches that actually occurred during those time intervals, and if a sufficiently good fit was obtained, respond positively.

This theory, involving hypothesis testing in the evaluation of targets against stimulus patterns (proposed by Dowling, 1984, for interleaved melodies and by Deutsch, 1972, in connection with octave-scrambled melodies), in turn suggests that listeners might be focusing their expectancies not only in time but also at just that region of the pitch spectrum in which target notes are expected. Experiment 2 tested this possibility by having listeners attempt to follow a pitch (occurring at an expected time) that moved from trial to trial within a target melody with which they had become familiar. On most trials, the critical pitch moved only a few semitones and remained within the range of the other notes of the target. However, on relatively few trials, it moved further and went outside the range of the target. If listeners were focusing their expectancies in pitch as well as in time, then the critical note should not be as well perceived on trials when it moved outside the range of those expectancies—outside the expectancy window.

EXPERIMENT 2

Listeners performed two types of task in Experiment 2: a contour-judgment task and a pitch-judgment task. Since the pitch-judgment task was considerably more difficult, all listeners did it second, in Session 2. The contour-judgment task followed the warm-up phase of Session 1. We wanted the results from both types of tasks because of our contention that something like an expectancy window was involved in listeners' perceptions. In the pitch-judgment task, expected target pitches were necessarily diatonic scale steps. The critical note, whose pitch was to be judged, moved to a variety of other pitches. Those pitches had to include nondiatonic pitches both for symmetry in the design and to make the task sufficiently difficult. However, we knew from previous studies (Bartlett & Dowling, 1980; Dowling, 1978; Francès, 1958/in press; Krumhansl, 1979; Shepard & Jordan, 1984) that nondiatonic pitches caused difficulties for judgment in a tonal context. We imagined that a nondiatonic pitch might produce errors of judgment even though it fell within our hypothesized window, simply because it was nondiatonic. Therefore, we included the contour-judgment task. Even if a nondiatonic pitch were not judged accurately in detail—even if it were assimilated to the category of the neighboring diatonic pitch—it should still produce a correct contour judgment, since none of the categories we employed fell across contour-changing boundaries. (We used two types of nondiatonic pitches in Experiment 2: semitones falling between the diatonic scale steps [D♯ and F♯], and quarter steps falling between diatonic steps and semitones [here symbolized D+ and F+].) The contour-judgment task was thus a control for the possibility that

pitch changes in target notes might be poorly judged for pitch, but not because of the effects of an expectancy window.

A typical trial from Session 1 of Experiment 2 is illustrated in Figure 3a. First a cue melody was presented, in which the next-to-last note was either above the notes on either side of it or below them. Then listeners heard the cue melody again, interleaved with distractor notes. In the interleaved pattern, the next-to-last note could move to a new pitch level. That new pitch level could be either above or below that of its neighbors in the target pattern, thus either leaving the contour (the pattern of ups and downs) unchanged or changing the contour. The listeners' task was to say whether the contour had been changed with respect to the cue.

On most trials, the moving pitch stayed within the range of the target melody. For example, if it had been presented as an F in the cue it could move to an F♯ (preserving the contour) or to a D or D♯ (below the neighboring Es and altering the contour). It could also move to the quarter steps above the D or the F, here symbolized D+ and F+. On a few trials, the critical pitch moved farther, to an A above or below the range of the target melody. Again, such a remote move could either alter the contour or leave it the same, and the listener had to say which of those alternatives had occurred. If listeners were focusing their
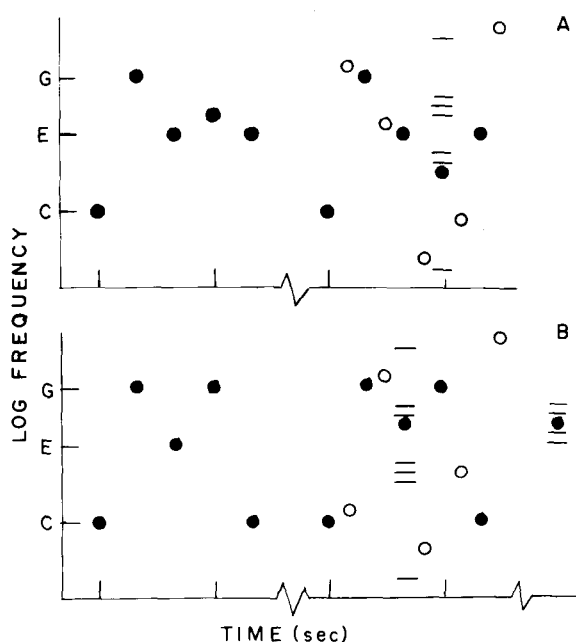

Figure 3. Sample trials from Experiment 2. Each trial began with presentation of a cue without interleaved distractors. (A) Contour task in Session 1: The next-to-last note of the target (filled symbols) moved to a new pitch in the interleaved pattern. Here it moves from an F to a D. Subjects judged whether the contour was altered as a result. The possible new pitches are indicated by short horizontal lines. (B) Pitch-judgment task in Session 2: The middle note (E) moved to a new pitch in the interleaved pattern (here an F). Subjects judged whether the comparison tone following the interleaved pattern matched that pitch or was higher or lower.

expectancies on the pitch regions in which the moving note usually appeared, then they should be less accurate in evaluating the contour of those targets in which the critical note moved beyond the expected regions.

In Session 2, this same possibility was tested using similarly constructed target melodies in which the middle note moved (Figure 3b); here listeners judged its new pitch in relation to that of a comparison tone following the interleaved pattern. Again, performance in pitch judgment of remote locations of the wandering tone should be poor to the extent that listeners focus their expectancies in a relatively narrow pitch region.

Since the comparison probe tones in Session 2 were either equal in pitch to the critical notes in the target melody or a quarter step higher or lower, we thought that, for the sake of symmetry, the pitches of the critical notes should also fall on quarter steps on some trials. This led, as indicated above, to the discovery that quarter-step target notes became assimilated to neighboring semitones.

Method

Subjects. Twenty-eight subjects served in Experiment 2—10 musically experienced and 18 musically inexperienced. Five additional inexperienced subjects, who failed to perform at better than 75% correct in the pretest described below, were dropped from the experiment.

Stimuli. All targets in Experiment 2 were in the key of C major and used middle C (fundamental frequency = 262 Hz) as the tonic, with an overall range extending from the G below middle C to the A above. The target stimuli in the warm-up consisted of the first 8 beats of each of six familiar tunes ("Frère Jacques," "Twinkle, Twinkle, Little Star," "Mary Had a Little Lamb," "Three Blind Mice," "Yankee Doodle," and "Old Macdonald Had a Farm"), plus two 5-note melodies designed for this experiment (like the one in Figure 3A). The latter two stimuli were the only targets used in the test phase (the contour-judgment task) of Session 1. Two additional 5-note melodies (like the one shown in Figure 3B) were used as targets in Session 2.

Whether interleaved with distractor notes or not, target patterns proceeded at a rate of 3 notes/sec. When interleaved, the stimulus presentation rate was 6 notes/sec. As in Experiment 1, interference notes were temporally interleaved among the notes of the target melodies, so that test stimuli alternated between target and interference notes. Targets were always presented on the beat, in the sense of Experiment 1. Interleaved stimuli were produced at a rate of 6 notes/sec. Each note was 137 msec in duration, with an onset time of about 24 msec followed by a decay to a level .27 of peak amplitude (decay constant = 48 msec). Internote intervals were 30 msec.

The whole experiment was replicated with all stimuli transposed up .5 semitone (a quarter step), so that semitone pitches in the "chromatic" scale on the piano became quarter steps and vice versa. Approximately equal numbers of subjects performed the experiment in each version, and each version used a different randomization of order. Since the quarter-step transposition had no effect on the results (as assessed by an ANOVA including that variable), the data are reported for the two versions combined. This replication at two pitch levels provided for a test of the possibility that at the level of assimilation to pitches in the tonal material subjects might display absolute-pitch-type behavior, especially since the absolute tuning standards of musical instruments have remained essentially constant in the culture during the subjects' lifetimes. The replication

also served as a control for possible uneven response characteristics across frequencies in the tone-generating equipment, which, if the quarter steps happened to be accidentally weaker acoustically in one version than in another, might account for less efficient perception.

Procedure. Each subject served in two group sessions spaced 2 or 3 days apart. Session 1 consisted of a warm-up phase, a pretest, and a test phase involving a contour-judgment task. Session 2 consisted of a second test phase, which involved a pitch-judgment task. Instructions were presented to subjects in a booklet along with answer sheets.

During the *exposure* phase of Session 1, subjects read the titles of the eight target melodies and listened to them one at a time, presented using a square waveform (a highly salient timbre) and with no interleaved interference tones. Then the subjects had eight trials in which they first heard a cue melody in the square-wave timbre (one of the eight possibilities) and then a square-wave target interleaved with distractors consisting of triangle waves (a contrasting, much "softer" timbre). The targets were easy to recognize under these conditions. The subjects judged the targets with respect to the cues, using a 6-point confidence-level scale that went from *very sure same* (6) to *very sure different* (1).

The *pretest* consisted of three blocks of eight trials each. On each trial, subjects first heard a cue melody presented as in the exposure phase and then, after a 2.0-sec pause, an interleaved comparison pattern. In contrast to the exposure phase, cues, targets, and distractors here were all in the same triangle-wave timbre. Test stimuli consisted of one of the eight target melodies interleaved with distractors in the same pitch range. Each cue melody appeared once in each block, in random order. On half the trials (randomly selected), the target matched the title cue, and on the other half, it did not (but was a random selection of one of the other target melodies). The subjects responded with the 6-category confidence-level scale as in the exposure phase.

The *contour-judgment* task of Session 1 consisted of 64 trials arranged in four blocks of 16 trials each. Each trial began with one of the two 5-note melodies (such as the one in Figure 3A) presented without interleaved tones; C−G−E−F−E or C−G−E−D−E. Notice that these cue melodies differ only in the direction of pitch change on the 4th note, which went either up or down in relation to the surrounding notes. After a pause of 1.6 sec, the comparison stimulus was presented. The comparison stimulus consisted of the melody C−G−E−X−E, which was presented with a distractor note following each melody note and in which X was replaced by one of eight possible target pitches: A(low), D, D+, D♯, F, F+, F♯, A(high). (D+ and F+ denote pitches .5 semitone above D and F, respectively.) When the cue melody C−G−E−F−E was followed by a comparison in which X was replaced with one of the latter four possibilities, the contour (the pattern of ups and downs) remained unchanged. However, when the comparison melody replaced the X with one of the first four possibilities on the list, the contour changed. Similarly, the contour of the melody C−G−E−D−E would have been affected in complementary fashion by those pitch substitutions.

The distractor notes interleaved among the notes of the target all fell on quarter-step intervals with respect to the cue stimuli, and were spaced so as to fall at least 2.5 semitones from the target pitch on a given trial. Each target was paired with a set of five interference tones spread evenly across the pitch range of the cue patterns, with a pair of tones straddling the target, 2.5 or 3.0 semitones above and below it. Distractor patterns were generated on each trial using a random permutation of the five possible interference pitches.

Each block of 16 trials in the contour-judgment task consisted of a random permutation of 16 possible trial types: 2 cues × 8 target pitches. On each trial, the subjects had 8 sec to respond, using

the 6-category confidence level scale in judging whether the contour of the comparison pattern was the same as or different from the cue.

During Session 2, subjects performed the *pitch-judgment* task consisting of 120 trials. Each trial began with one of two 5-note cue melodies (C−G−E−G−C [shown in Figure 3B] or C−G−E−C−G) presented without interleaved tones. Notice that these melodies each contained two Cs and two Gs, and just one E in the middle position. The comparison stimulus consisted of a repetition of the cue melody, presented with interleaved tones following each melody note, but with the E replaced by a new pitch. Comparisons were thus of the form C−G−X− ..., in which X was replaced by one of eight possible target pitches: A(low), D, D+, D♯, F, F+, F♯, A(high), as in the contour task of Session 1. The subjects' task was to judge whether the pitch of the probe tone was lower than, equal to, or higher than the pitch of the target. Interleaved distractors were distributed with respect to target tones, as in the contour task. The 120 trials were arranged in five blocks of 24 trials each, and were introduced with eight examples accompanying the instructions.

After an interval of 1.6 sec following the comparison stimulus there was a 1-sec-long probe tone that was either equal to the target tone in frequency or .5 semitone (a quarter step) higher or lower. The subjects were instructed to listen very closely to where the middle note, the E, in the comparison pattern moved to and to remember that pitch. Then they were to judge the pitch of the probe in relation to that pitch, marking their answer sheets with a "−," "=," or "+." Each block of 24 trials consisted of a random permutation of 24 possible trial types: 8 targets × 3 probe pitches for each target. The selection of cue melody on each trial was randomized.

Results

Pretest. Responses were collapsed across confidence-level categories to produce a proportion of correct "same" and "different" responses for each subject on the pretest. Only subjects achieving better than .75 correct on this measure were retained in the subsequent data analysis.

Contour-judgment task. Confidence-level judgments of contour were used to generate receiver operating characteristics (ROCs) for each subject, and area under the ROC was taken as an estimate of unbiased proportion correct where chance is .50 (Swets, 1973). These area scores were evaluated by means of a two-way ANOVA: 2 experience levels × 4 target pitch types. Only the main effect of target pitch type was significant [$F(3,78) = 5.63$, $p < .01$]. Mean contour-judgment-area scores for the four types of target pitches (diatonic scale steps, D and F; semitone intervals, D♯ and F♯; quarter steps, D+ and F+; and remote diatonic pitches, high and low A) are shown in Table 2. With the exception of the semitone intervals, contour responses to targets within the expectancy window were around 72% correct, whereas responses to the remote diatonic pitches outside the window were only 59% correct. A planned comparison between expected and remote diatonic steps yielded a significant difference [$t(27) = 2.45$, $p < .05$].

Pitch-judgment task. To evaluate pitch-judgment performance, we calculated A' for each subject for each of the four types of target pitch, taking "equal" judgments when probes were actually equal to targets as hits and "equal" judgments when probes were different from tar-

Table 2
Performance on the Contour-Judgment (Area Under ROC) and Pitch-Judgment (A') Tasks of Experiment 2

Task	Pitch Category of Critical Note			
	Diatonic D, F	Semitone D♯, F♯	Quarter Step D+, F+	Remote Diatonic A, A'
Contour judgment	.72	.62	.72	.59
Pitch judgment	.58	.53	.44	.52

gets as false alarms. As with the contour task, these A' scores were evaluated by means of a two-way ANOVA: 2 experience levels × 4 target pitch types. Only the main effect of target pitch type was significant [$F(3, 78) = 3.28$, $p < .01$]. Mean pitch-judgment A' scores for the four types of target pitches are shown in Table 2. Again, responses to diatonic targets within the expectancy window were superior to those outside it: 58% correct versus 51%. A planned comparison between expected and remote diatonic steps fell short of significance [$t(27) = 1.46$, $p < .10$], although with this more difficult task performance seems to be encountering a floor effect.

Note, in the pitch-judgment results, the worse-than-chance performance with quarter-step targets, resulting from their assimilation to neighboring semitone and diatonic probes. That is, when the target was an expected diatonic step, "equal" judgments to actually equal diatonic probes exceeded "equal" judgments to its quarter-step neighbors. However, when the target was a quarter step, "equal" judgments to its actually different diatonic and semitone neighbors exceeded "equal" judgments to actually equal diatonic probes. This phenomenon agrees with previous results (Krumhansl & Shepard, 1979), and we will explore it with further experiments in a separate article.

Discussion

Especially in the contour-judgment task, but to some extent in the pitch-judgment task, diatonic pitches outside the expected range of the melody were difficult to perceive. This result agreed with the experience of the experimenters in listening to the stimuli and with the comments of some of the subjects. A typical observation was that on some trials the critical note seemed simply to disappear. The listeners were sure that *something* had happened at that point in time, and that whatever pitch had occurred was not within the region they had been attending to. But they were not able to say what pitch it was. The phenomenon was not so much one of "masking" (in that listeners would have been able to detect the presence of a stimulus in the critical time slot) as it was akin to what Divenyi and Hirsh (1975) called "blanking." We have replicated the contour-judgment results several times because that task was used in Session 1 of most of the subsequent experiments, including the ones reported below.

The quarter-step assimilation that led to worse-than-chance performance in the pitch task did not appear in the contour-judgment task, presumably because when the

quarter steps were assimilated to neighboring semitones and diatonic steps the result in terms of contour was the same as if they had not been assimilated. Quarter steps produced just about the same level of accuracy in that task as did diatonic steps. What caused listeners difficulty in the contour task were the nondiatonic semitones (D♯ and F♯). It may be that the fact that those pitches were jarringly out of key disturbed the encoding or judgment of contour—a result that converges with results of a quite different contour-judgment task reported by Bartlett and Dowling (1987).

In Neisser's (1979) selective-looking experiment described above, in which he introduced an unexpected event—the umbrella lady—into the display of two intermingled ball games, the observers concentrating on the game generally missed the umbrella lady. This is analogous to our listeners' tending to miss the unexpected remote pitches. Presumably, if Neisser had introduced the umbrella lady in a condition in which viewers were presented with just *one* game to follow, the viewers would not have had to narrow their attentional focus to pick up the target game and more viewers would have noticed the unexpected event. This is analogous to what we did in Experiments 3 and 4, which were designed to illustrate the attentional effects of the interleaved distractor notes. Experiments 3 and 4 replicated the contour-judgment (Session 1) and pitch-judgment (Session 2) tasks of Experiment 2, respectively, but without distractors in the stimuli. Without the narrowing of attentional focus brought about by the distractors, the unexpected remote pitches should become more salient.

EXPERIMENT 3

Method
Subjects. Twenty-four subjects served in Experment 3—13 musically experienced and 11 inexperienced. One additional experienced subject failed to follow the instructions and was dropped from the experiment.

Stimuli. The stimuli for the contour-judgment task were essentially the same as those of Experiment 2, except that there were no distractors. The warm-up and pretest phases were identical to those of Experiment 2.

Procedure. The procedure was the same as in Experiment 2, Session 1.

Results
We calculated area under the ROC for each subject for each stimulus type, as in Experiment 2. The data are shown in Table 3. We subjected those scores to a 2 experience levels × 4 stimulus types (diatonic, semitone, quarter step, remote diatonic) ANOVA. There were no significant effects. Contour was equally easy to judge for all stimulus conditions, and considerably less difficult than in Experiment 2 (Table 2). In particular, there was no indication that the remote diatonic steps that were more difficult to judge in Experiment 2 were more difficult than the other pitch levels when distractors were removed.

Table 3
Performance on the Contour-Judgment (Area Under ROC) Task of Experiment 3 and the Pitch-Judgment (A') Task of Experiment 4

| Task | Pitch Category of Critical Note | | | |
	Diatonic D, F (D♯, F)*	Semitone D♯, F♯	Quarter Step D+, F+	Remote Diatonic A, A'
Contour judgment	.85	.88	.84	.85
Pitch judgment	.63		.60	.76

*For Experiment 4, the data were collapsed over these two categories so that equal numbers of trials would contribute to each cell. Pitch judgments of the F and D♯ were close to equal.

EXPERIMENT 4

Experiment 4 was a simplified replication of Experiment 2 with the distractors eliminated from the pitch-judgment task of Session 2. Listeners performed Session 1 in exactly the same manner as in Experiment 2, with the same results.

Method
Subjects. Thirteen subjects served in Experiment 4—8 musically experienced and 5 inexperienced. Two experienced subjects and 2 inexperienced ones failed to perform better than 75% correct on the pretest and were dropped from the experiment.

Stimuli. Session 1 was identical to Session 1 of Experiment 2. The stimuli for the pitch-judgment task in Session 2 were essentially the same as those of Experiment 2, except that there were no distractors and there were only six critical pitches to which the middle note could move: A(low), D+, D♯, F, F+, A(high). Each critical pitch was tested with five possible probe comparisons spaced at .5-semitone (quarter-step) intervals in the range of ±1 semitone around the critical pitch. For data analysis, the critical pitch types were collapsed into three categories: diatonic and semitone (D♯, F), quarter step (D+, F+), and remote diatonic (the As). Performance with the D♯ and F was essentially equivalent.) We did this so that there would be a comparable number of trials upon which each cell mean was based. There were 150 trials arranged in five blocks of 30 each. Within each block, each of the six critical pitches was tested with each of the five probes.

Procedure. The procedure was the same as in Experiment 2.

Results
We calculated A' for each subject for each stimulus type, as in Experiment 2. The data are shown in Table 3 and were subjected to a 2 experience levels × 3 stimulus types (diatonic and semitone, quarter step, remote diatonic) ANOVA. Only the effect of stimulus types was significant [$F(2,22) = 5.70$, $p < .02$], with performance being distinctly better on remote diatonic steps than on pitches within the expected pitch range. Again, removing the distractor notes appears to free the attentional mechanisms from the necessity of focusing narrowly on the expected range.

Discussion
There is a striking contrast between performance in judging the remote diatonic pitches in Experiments 3 and 4 and that in Experiment 2. In Experiment 2, with dis-

tractor notes present, performance was worse with diatonic pitches outside the expected range than for those within the expected range, for both the contour- and the pitch-judgment tasks. In fact, when distractors were removed, performance on the pitch-judgment task went from around chance to .76—significantly better than performance with expected pitches. The remote pitches were perceptually very salient and easy to judge with distractor notes absent. Discerning the target melody with distractors present, however, requires a focusing of attention in pitch-time expectancy windows within which events critical for target identification are likely to occur. As a result of that narrowing of attention, events that would otherwise be highly salient were not distinctly perceived and were difficult to judge accurately.

EXPECTING AN OFF-BEAT NOTE

We return now to the dimension of time that we explored in Experiment 1. There listeners performed better in judging target melodies whose notes fell on the implicit temporal beat in a stimulus pattern than they did in judging those whose notes fell off the beat. We suggested that listeners used the temporal structure of the stimuli to aim a series of expectancy windows at times when target events were likely to occur. Beat structure lies at a fundamental level of temporal organization in music, but it is typically overlaid with a more complex organization of surface rhythms (Dowling & Harwood, 1986). We were curious about whether those more complex rhythms could be used to guide expectancies in the same way that beat structure did. That is, would listeners be able to aim their expectancies at a critical note in a target that was expected, but *off* the beat? Accordingly, in Experiments 5 and 6 we familiarized listeners with a target melody that had a dance-like rhythm as illustrated

in Figure 4. The second note in this pattern was syncopated—that is, it fell off the beat. Our first question was: Would listeners be able to follow the syncopated note when the target was interleaved with distractor notes? That is, could listeners aim their attention at an off-beat position in time?

To test this, we gave listeners the task of plotting the position in pitch and time of the critical note of an interleaved pattern such as that in Figure 4. Listeners put an X in a grid indicating the position of the note in the comparison stimulus. The critical note occurred most often in its syncopated center position, but on some trials it wandered into one of the other locations. We were interested in learning how effective listeners would be in picking up the critical note in its expected, but syncopated, position.

The second question we addressed in Experiments 5 and 6 was whether pitch or time was more effective in providing a framework for the listener to organize expectancies. That is, if you are listening for a critical note that wanders in pitch and time, do temporal expectancies or do pitch expectancies help more in focusing attention? Are you better at judging the pitch of the note when it occurs at the expected time than when it does not? Are you better at judging the time of the note when it occurs at the expected pitch? And for which dimension does an expected event have the greatest advantage over an unexpected one?

EXPERIMENT 5

Method

Subjects. Twenty-one subjects served in Experiment 5—12 musically experienced and 9 musically inexperienced. Three inexperienced subjects and one experienced subject failed to follow the instructions and were dropped from the experiment.

Procedure. The procedure was the same as in Experiment 2 and covered two sessions 2 or 3 days apart. Session 1 of Experiment 5
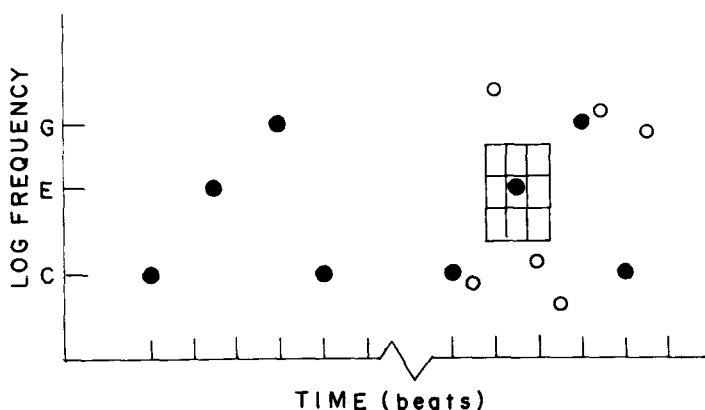

Figure 4. A sample trial from Experiments 5 and 6. The second note (E) in the target (filled symbols) was the critical one, and occurred *off* the beat in the cue and on most trials in the interleaved pattern. On some trials it moved to one of the eight other positions in pitch and time indicated by the 3 × 3 grid. Interleaved distractors (open symbols) occurred at times not occupied by target notes. Subjects responded either by marking a similar grid (Experiment 5) or by judging just the pitch or timing of the critical note (Experiment 6).

was identical to Session 1 of Experiment 2. In Session 2, subjects performed a task with trials structured as in Figure 4. On each trial, they heard a cue melody C—E—G—C, in which all the notes, except for the E, fell on the beat; E fell midway between note positions 3 and 5 (beats 2 and 3). After an interstimulus interval of 1.17 sec, the subjects heard the interleaved comparison pattern. In that pattern, the target melody became C—X—G—C, where X could be the expected E or D or F. X could also occur at the expected time or .17 sec (one-half beat) earlier or later. All .17-sec intervals in which target notes did not occur were filled with distractor notes in ranges that were on the edges of the target melody. There were thus nine possible positions that the critical note could occupy, including the expected syncopated E.

There were 96 trials in Session 2, preceded by 12 examples that illustrated the positions to which the critical note could move. On half the trials, the critical note was the same as in the cue (the syncopated E). On the other half, it moved to one of the other eight positions with equal probability. In the first 32 trials, the target melody was given a salient timbre by being played with a square waveform, while the distractors were played in the softer triangle-wave timbre. In the remaining 64 trials, the targets were hidden by being played in triangle waves as were the distractors. The listeners responded to each trial by marking an X in a 3 × 3 grid similar to the one drawn in Figure 4.

Stimuli. The trials in Session 2 were arranged in six blocks of 16 trials each. Within each block the critical note was in the expected position eight times and in each of the other positions once, in random order. The six distractor notes on each trial were a random permutation of six quarter steps clustered around the C and the G at the edges of the target melody: B+, C+, C♯+, F♯+, G+, G♯+. The distractors filled in the temporal spaces between target notes, occurring in different temporal arrangements, depending on the location of the critical note.

Results

To test for accuracy of pitch and time judgments, especially as they related to expectancies, we reduced the data to conditional probabilities of correct pitch and time judgments given whether the position of the critical note on the other dimension was expected or not. For example, we took the probabilities of correct pitch responses on those trials on which the critical note had occurred at its syncopated, expected time (that is, in column 2 of the grid in Figure 4). A response was counted correct for pitch if the response grid was marked in the row in which the pitch occurred, even if the subject had the column wrong. For the probability of correct pitch judgment given a critical note at an unexpected time, we took correct pitch judgments in cases where the critical note had occurred earlier or later than the expected time (that is, in columns 1 or 3 of the grid). We obtained corresponding correct-response probabilities in a similar manner for time judgments, conditional upon the pitch (grid row) of the stimulus, again counting any mark in the correct column as correct. Thus, we obtained the probability of a correct pitch judgment given a critical note at the expected time and the probability of a correct pitch judgment given a critical note at an unexpected time. We treated time judgments given expected and unexpected pitches similarly. Each subject's performance was characterized by those four proportions of correct responses.

The conditional probabilities of correct responses were subjected to a four-way ANOVA: 2 experience levels × 2 target timbres (salient, hidden) × 2 dimensions judged (pitch, time) × 2 expectancy for other dimension (expected, unexpected). There was a significant main effect of experience [$F(1,19) = 6.56, p < .02$], with experienced subjects performing better than inexperienced ones (.60 correct vs. .37). This task, in which the required response was very similar in form to musical notation, was quite difficult for the inexperienced subjects. Experience did not interact with any other variable, and so the data shown in Table 4 are collapsed across experience levels to display the other effects more clearly.

The only two other significant main effects were those of salience [$F(1,19) = 4.43, p < .05$] and expectancy for other dimension [$F(1,19) = 7.22, p < .02$]. Salient targets (.55) were judged more accurately than hidden ones (.45). And performance with expected stimuli (.54) was superior to that for unexpected ones (.45). The effect of expectancy is seen mainly in the judgment of *pitches* of notes at expected (vs. unexpected) *times*—.58 versus .42 in the column means of Table 4—producing the significant interaction of dimension judged × expectancy [$F(1,19) = 9.40, p < .01$].

The overall interaction shown in Table 4 of target salience × dimension judged × expectancy was significant [$F(1,19) = 7.23, p < .02$]. There was little effect of expectancy on time judgments for either salient or hidden targets, but the effect of expectancy on pitch judgments was greater for salient than for hidden targets. The interaction of salience × dimension judged was significant [$F(1,19) = 8.50, p < .01$]. Time judgments gained more going from hidden to salient targets (from .41 to .59) than did pitch judgments (.49 to .51). And the salience × expectancy interaction was significant [$F(1,19) = 4.87, p < .05$]. Expectancy had a greater effect with salient targets (.62 for expected vs. .48 for unexpected) than it did for hidden targets (.46 vs. .44).

There were no other significant effects; in particular, there was no main effect of dimension judged. Pitch and time judgments were virtually equal in difficulty (.50). However, temporal expectancy had a much greater impact on pitch judgments than did pitch expectancy on time judgments.

Discussion

The fact that overall performance in Experiment 5 was .50 correct—distinctly superior to a chance level of around .33—shows that listeners are able to focus attention on an expected target note even when it falls off the beat. Furthermore, although performances in judging pitch and time were roughly equal overall, pitch judgments benefited much more from temporal expectancies than did time judgments from pitch expectancies. This suggests that in stimuli constructed like the present ones, the expectancy window is more narrowly focused in time than in pitch, somewhat as depicted in Figure 1. In Experiment 2, notes

outside the range of the melody (7 semitones in this case) fell outside the expected range. The present results suggest that the 3-semitone range in which critical notes fell in Experiment 5 lay entirely within the expected window. That is, all the pitches were more or less expected, whereas all the times were not. With such rapid stimuli, time became narrowly focused.

It is curious that the effect of temporal expectancy on pitch judgments was most pronounced for salient targets. We anticipated the opposite effect—that hidden targets would force a narrowing of the window. In connection with the enhancement of time judgments for salient targets, it may be that accentuating the rhythm of the target by making its notes more salient may have facilitated aiming the windows in time. We are not entirely satisfied with that explanation for the attenuation of the expectancy effect with hidden targets, but are pleased to note that the effect of temporal expectancy appeared to be fairly strong with hidden targets in Experiment 6.

EXPERIMENT 6

Experiment 6 was in the first place designed to simplify the task of Experiment 5. Nonmusicians had found the task of Experiment 5, which resembled melodic dictation, to be extremely difficult. Therefore, in Experiment 6 different groups of listeners judged just one dimension: either pitch or time. This made the task easier, especially for the inexperienced listeners, and the main effect of experience disappeared.

By letting listeners focus on just one dimension of the critical notes, we also wished to find out whether the failure of listeners to make use of pitch expectancies to aid time judgments would disappear with an easier task. In the simpler task, we anticipated that time judgments might be better at expected pitches, just as pitch judgments had been better at expected times in Experiment 5. The results confirmed this supposition.

Finally, we were interested in the effects of providing some listeners with the opportunity for additional perceptual learning with the materials. Thus, about half the listeners participated in an additional session interposed between the warm-up session (the same as in Experiment 2) and the final test session. In that additional training session, listeners performed the same type of pitch- or time-judgment task that they would do in the test session, but with stimuli in which the nonjudged dimension was held constant at its expected value.

Method

Subjects. Thirty-five subjects served in Experiment 6—17 musically experienced and 18 musically inexperienced. Four inexperienced subjects failed the pretest and were dropped from the experiment.

Procedure. The procedure was essentially the same as in Experiment 5 and covered two or three sessions within a week. Session 1 of Experiment 6 was identical to Session 1 of Experiment 2. For about half the subjects (16), there was an additional training session between Session 1 and the test session. There were at least 4 subjects in each experience × training condition. In the test ses-

sion, subjects performed a task similar to that in Experiment 5, except that each group of subjects judged only one dimension of the stimuli.

The test session was introduced by 11 examples that illustrated the positions to which the critical note could move. In the first six examples, the target was in the salient square-wave timbre. In the remainder of the test session, the targets were hidden and had the same triangle-wave timbre as the distractors.

The additional training session was constructed exactly like the test session except that (1) the first block of 24 trials had targets in the salient timbre, and (2) all stimuli had the same, expected value, on the stimulus dimension (pitch or time) not being judged.

Stimuli. The 72 trials of the test session were arranged in three blocks of 24 trials each. In each block, the critical note appeared 18 times at the expected value of the dimension not being judged and 6 times at the unexpected values. Expected critical notes appeared 6 times in each of the positions being judged, and unexpected notes once in each position. Thus, when pitch was to be judged, the critical note appeared 18 times at the expected, syncopated time. There were 6 of those appearances at each pitch level—high, middle, and low (F, E, and D). Six times the note to be judged for pitch appeared at unexpected times—three early and three late, with one at each pitch level for each time. The time-judgment conditions were arranged similarly, with dimensions exchanged.

The 6 distractor notes on each trial were a random permutation of 6 semitones lying just outside the range of the target melody: A, A♯, B, G♯, A, A♯. Moving the distractors further out from the target was intended to make the task somewhat easier than that in Experiment 5.

Results

The data were analyzed as for Experiment 5. Conditional proportions of correct responses given expected and unexpected values of the nonjudged dimension were subjected to a four-way ANOVA: 2 experience levels × 2 training conditions × 2 dimensions judged (pitch, time) × 2 expectancy levels for other dimension (expected, unexpected). Only the main effect of expectancy [$F(1,27) = 8.14, p < .01$] and the dimension judged × training condition interaction [$F(1,27) = 6.42, p < .02$] were significant. Stimuli at the expected value of the nonjudged dimension were judged more accurately than those at unexpected values (.60 correct vs. .52 correct).

The dimension judged × training interaction is shown in the partial row means of Table 5, along with results to compare with those of Table 4. The interaction arose from the fact that additional training improved pitch judgments considerably, but did not improve time judgments (which actually declined somewhat for the subjects with the additional session).

Table 4
Proportion Correct Responses Judging Pitch and Time
of Critical Notes in Notation Task of Experiment 5

| Target | Dimension Judged | | | | |
| | Pitch | | Time | | |
	Time Expected	Time Unexpected	Pitch Expected	Pitch Unexpected	Mean
Salient	.64	.37	.60	.58	.55
Hidden	.51	.47	.42	.40	.45
Mean	.58	.42	.51	.49	

Table 5
Proportion Correct Responses Judging Pitch and Time
of Critical Notes in Experiment 6

| Training Sessions | Dimension Judged | | | | | |
| | Pitch | | | Time | | |
	Time Expected	Time Unexpected	Mean	Pitch Expected	Pitch Unexpected	Mean
One	.51	.45	.48	.60	.56	.58
Two	.70	.60	.65	.58	.46	.52
Mean	.60	.53		.59	.51	

Discussion

With the simplified task of responding on only one dimension, with the slight adjustment in distractor pitches, and with half the subjects receiving an additional training session, the effect of musical experience observed in Experiment 5 disappeared. We should note that the experience effect in Experiment 5 was only a main effect—that the behavior of inexperienced listeners was *qualitatively* the same as the behavior of the moderately experienced, just at a lower overall level. Here the behavior of the two groups was both qualitatively and quantitatively equivalent.

It is interesting that the interactions of dimension judged with other variables disappeared when the task allowed for focusing attention on just one dimension. This suggests that expectancy along either dimension can serve to focus attention. When making the more complex judgments in Experiment 5, listeners generally selected the strategy of focusing more narrowly on time than on pitch. However, the present results show that the focusing system is flexible, and that when the task is simplified a capacity to focus expectancies narrowly in either dimension appears. The broad end of the continuum of flexibility was seen in Experiments 3 and 4, in which, without distractors, attention was not particularly focused at all.

The fact that pitch judgments benefited more from additional perceptual learning than did time judgments probably reflects the more limited experience even our moderately experienced listeners have in judging pitch in their everyday lives. In contrast, time judgments of one sort or another are common. Thus, when given an opportunity to practice pitch judgments, listeners improved their performance, in contrast to performance with the already more practiced time judgments. It also seems likely that the extra training session with a constant, expected pitch led to a narrowing of pitch expectancies around that pitch, so that time judgments for notes at unexpected pitches actually declined in accuracy. At the very least, we can say that pitch and time expectations and judgments were differently affected by highly similar perceptual learning experiences.

GENERAL DISCUSSION

The hidden-melodies tasks presented here are similar to the "embedded" conditions described by Divenyi and

Hirsh (1978), who used a three-note target surrounded by distractor notes in the same or different pitch regions. In Divenyi and Hirsh's stimuli, the notes of the target remained temporally contiguous, and were not interleaved with distractors. (Divenyi and Hirsh were concerned that the target would lose its properties as a meaningful, whole gestalt if it were split up by interleaved distractors; but the present results show that targets remain identifiable even when temporally interleaved with distractors.) Consonant with the results of Experiment 1, Divenyi and Hirsh found that target identification was more difficult when distractor notes occurred in the same pitch range as the target. Divenyi and Hirsh suggested that listeners might be pursuing one of two strategies in identifying the hidden targets. They might be focusing attention on the particular pitch region in which the target was to occur (similar to our "expectancy window" strategy) or they might be trying to find the contour of the target in the complex pattern. Divenyi and Hirsh (1978, Experiment 7) devised a task in which the target changed pitch region while the distractor notes remained in the same pitch region on each trial. Target identification remained better when the targets and distractors were in separate pitch regions, even though target pitches were not expected. Divenyi and Hirsh therefore favored a contour-identification strategy in their interpretation of their results. However, the present results from tasks in which target pitches remained in the same region strongly suggest a pitch-focusing strategy. We believe the correct conclusion is that the auditory system is flexible enough to adopt either strategy, depending on task requirements.

The "expectancy window" metaphor of attention that we have adopted is similar to the "zoom lens" metaphor of Eriksen and St. James (1986). There is, however, one difference in the implications of the two metaphors that leads us to prefer our model for the present tasks, which we believe arises from the direction of attention in time required in our experiments. One feature of the zoom-lens model is that as the attentional focus is narrowed, perceptual resolution becomes better. However, our results show that as the attentional focus is narrowed in time, discrimination worsens. (Compare results in Table 3 for the pitch-judgment task in Experiment 4 with those in Table 2 for Experiment 2.) Pitch and time do not seem to be directly analogous to two spatial dimensions in vision, and they differ largely because time is different. As

Massaro (1975) suggested, enforcing haste in processing by the introduction of interfering stimuli produces more errors—an effect that the narrower focusing of attention does not overcome. It appears to take more information extracted from a stimulus to tell *what* it was than to tell *when* it occurred. One aspect of the expectancy window notion that we like is that it captures the distinction between our ability to *expect* more than one possible event within a given region and our ability to *attend* to only one event when it occurs.

Performance on these hidden melodies tasks demonstrates the active nature of auditory attention. Listeners with and without musical training can control the rhythms of their attention to match the rhythmic structures of familiar stimuli. In doing that, it is as though they had set up a series of "expectancy windows" aimed at critical events in the ongoing stimulus pattern. The windows can be aimed easily at beats in the temporal structure, and can also be aimed at expected events off the beat. The windows are typically narrow in time and relatively broad in pitch, encompassing a range of the order of 7 semitones in these studies. However, we found indications that the size and shape of the window is flexible, and varies depending on task demands.

REFERENCES

BARTLETT, J. C., & DOWLING, W. J. (1980). The recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception & Performance*, 6, 501-515.

BARTLETT, J. C., & DOWLING, W. J. (1987). Tonality and the asymmetric similarity of melodies. Manuscript in preparation.

DEUTSCH, D. (1972). Octave generalization and tune recognition. *Perception & Psychophysics*, 11, 411-412.

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

DIVENYI, P. L., & HIRSH, I. J. (1975). The effect of blanking on the identification of temporal order in three-tone sequences. *Perception & Psychophysics*, 17, 246-252.

DIVENYI, P. L., & HIRSH, I. J. (1978). Some figural properties of auditory patterns. *Journal of the Acoustical Society of America*, 64, 1369-1385.

DOWLING, W. J. (1973a). The perception of interleaved melodies. *Cognitive Psychology*, 5, 322-337.

DOWLING, W. J. (1973b). Rhythmic groups 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. (1984). Musical experience and tonal scales in the recognition of octave-scrambled melodies. *Psychomusicology*, 4, 13-32.

DOWLING, W. J., & HARWOOD, D. L. (1986). *Music cognition*. New York: Academic Press.

ERIKSEN, C. W., & ST. JAMES, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, 40, 225-240.

FRANCÈS, R. (in press). *The perception of music* (W. J. Dowling, Trans.). Hillsdale, NJ: Erlbaum. (Original work published 1958)

GIBSON, E. J. (1969). *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

GIBSON, E. J. (1978). How perception really develops: A view from outside the network. In D. LaBerge & S. J. Samuels (Eds.), *Basic processes in reading*. Hillsdale, NJ: Erlbaum.

GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

GRIER, J. B. (1971). Nonparametric indices for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.

HELMHOLTZ, H. L. F. VON (1954). *On the sensations of tone* (A. J. Ellis, Trans.). New York: Dover. (Original work published 1877)

JONES, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 83, 323-355.

JONES, M. R., KIDD, G., & WETZEL, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception & Performance*, 7, 1059-1073.

KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

KRUMHANSL, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 11, 346-374.

KRUMHANSL, C. L., & SHEPARD, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception & Performance*, 5, 579-594.

MASSARO, D. W. (1975). *Experimental psychology and information processing*. Chicago: Rand McNally.

NEISSER, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development* (pp. 201-219). Hillsdale, NJ: Erlbaum.

PICK, A. D. (1979). Listening to melodies: Perceiving events. In A. D. Pick (Ed.), *Perception and its development* (pp. 145-165). Hillsdale, NJ: Erlbaum.

SCHOPENHAUER, A. (1966). *The world as will and representation* (E. J. F. Payne, Trans.; 2 vols.). New York: Dover. (Original work published 1844)

SEKULER, R., & BLAKE, R. (1985). *Perception*. New York: Knopf.

SHEPARD, R. N., & JORDAN, D. S. (1984). Auditory illusions demonstrating that tones are assimilated to an internalized musical scale. *Science*, 226, 1333-1334.

SPELKE, E. S. (1979). Exploring audible and visible events in infancy. In A. D. Pick (Ed.), *Perception and its development* (pp. 221-235). Hillsdale, NJ: Erlbaum.

SWETS, J. A. (1973). The relative operating characteristic in psychology. *Science*, 182, 990-1000.

VAN NOORDEN, L. P. A. S. (1975). *Temporal coherence in the perception of tone sequences*. Eindhoven, The Netherlands: Institute for Perception Research.

WATSON, C. S., & KELLY, W. J. (1981). The role of stimulus uncertainty in the discrimination of auditory patterns. In D. J. Getty & J. H. Howard, Jr. (Eds.), *Auditory and visual pattern recognition* (pp. 37-59). Hillsdale, NJ: Erlbaum.