The Processing of Temporal and Nontemporal Information in the Remembering of Event Durations and Musical Structure

Marilyn G. Boltz
Haverford College

Within the field of cognitive psychology, one issue of theoretical debate concerns the relationship between an event's temporal (e.g., rhythm, rate, total duration) and nontemporal (e.g., sequence of pitch intervals) information. Some argue these are jointly encoded into the cognitive system, whereas others claim they are encoded in an independent fashion. This issue was investigated in 2 experiments that systematically manipulated participants' attention to the temporal and pitch properties of melodies and then examined participants' subsequent remembering of these structural dimensions. The results indicate that the nature of encoding is strongly dependent on the structure of environmental events and the degree of learning experience. This relationship in turn has implications for theories of both music cognition and time estimation behavior.

One of the most general goals in cognitive psychology is the determination of how events from the natural environment are encoded into the cognitive system. In investigations of this issue, it has traditionally been assumed that the distinction between an event's temporal and nontemporal information is an important one. Temporal information refers to the rate, rhythmicity, and temporal ordering of items as they unfold over a given time span, whereas nontemporal information refers to the items themselves. These items are typically assumed to have spatial properties as exemplified by different objects in the visual environment (e.g., cars, people) or pitch relations within the auditory modality (e.g., notes of a melody, words of a conversation).[1]

The primary issue of concern involves the relationship between temporal and nontemporal information. From a structural perspective, it is clear that the definition of any given event necessarily requires both types of information: All events are dynamic in nature and unfold over a given time span, but in order for a time span to exist, it must be marked by nontemporal information. Music, for example, displays nontemporal information in the form of pitch relations that not only mark the beginning and end of a melody but also unfold with a given rate and rhythmicity over the total time span. On a structural basis, then, an event's temporal and nontemporal information are inextricably entwined. However, from a cognitive perspective, the picture is less clear. Is it the case that temporal and

nontemporal information are jointly encoded into the cognitive system such that attending to one dimension results in the incidental learning of the other? If so, then this results in a parsimonious system because it means that both dimensions can be equally well learned and later remembered regardless of whether an individual initially attended to one or the other dimension alone or to both dimensions in tandem. Alternatively, others have argued that this type of encoding is not possible because it would exceed the overall resources of the cognitive system. Instead, it is claimed that temporal and nontemporal information must each be selectively attended to independently for learning to occur. Although each type of processing would result in a memorial representation that reflects both the temporal and nontemporal information of an event, the latter type of processing would require more attentional effort. Arguments for both types of processing have been offered in the cognitive literature and are of particular interest to at least two areas of research—namely, those of music cognition and time estimation behavior.

Music Cognition

In theories of musical behavior, the debate has focused primarily on the processing of pitch information relative to the temporal dimension of rhythm. It is well-accepted that both play important roles in the perception and remembering of melodies, but again, one finds disagreement on how each is encoded into the cognitive system. According to one

[1] Within the visual modality, the spatial dimension of events is rather self-evident in that people and other objects move over varying distances from one point to another. Similarly, it is generally well accepted that pitch relations form an analogous type of structure within the auditory environment (Julesz & Hirsh, 1972; Kubovy, 1981; Marks, 1978). Pitch relations are categorized as high or low, can span different distances over successive intervals, and form a particular contour over a total time span (i.e., the pattern of "ups and downs" in pitch).

perspective, pitch and rhythm are processed independently such that the influence of one dimension does not depend on the characteristics of the other dimension (Palmer & Krumhansl, 1987). For example, certain types of rhythms may facilitate attending and remembering regardless of the accompanying pitch pattern. Evidence for this type of relationship occurs primarily when individuals are asked to judge the similarity of different melodies (Monahan & Carterette, 1985) or the perceived completeness of melodic phrases (Palmer & Krumhansl, 1987). In contrast, others have argued that pitch and rhythm exert an interactive influence on behavior, in which the effect of one dimension is dependent on the other dimension (Boltz & Jones, 1986; Jones & Boltz, 1989). For example, a given rhythm may facilitate the perception and memory of some pitch patterns but not others. This type of relationship was in fact found when individuals were asked to recall melodies (Boltz, 1991; Boltz & Jones, 1986) or detect deviant pitch intervals in a melody recognition task (Boltz, 1993; Jones, Boltz, & Kidd, 1982). In sum, then, the independence–interactive debate has not been resolved, although the processing of pitch and rhythm does seem to depend on task demands.

Time Estimation Behavior

A second area of relevance is that of time estimation behavior, where the assumed theoretical relationship between temporal and nontemporal information has been more extensively developed. In this literature, the primary issue of interest is the determination of how the total durations of environmental events are encoded and retrieved from memory. Although several different theoretical models have been offered, most rely on the notion of attentional resource allocation as the underlying mediational mechanism at play. Zakay (1989) developed one of the most comprehensive models not only by incorporating the ideas of other theorists but by integrating these ideas with certain memory-based heuristics that have been suggested to influence time estimation behavior (e.g., Block, 1985; Ornstein, 1969).

The primary assumption of this approach is that when individuals are performing any given activity, such as listening to a melody or doing a set of math problems, attention is split between the event's temporal (i.e., total time span) and nontemporal (e.g., notes or numbers) information. Each dimension is assumed to be independent of the other such that temporal information is encoded via a cognitive timer and nontemporal information is encoded through its own separate mechanism. Both compete for a limited pool of resources, such that increased attention directed toward one dimension will decrease performance on the other.

This relationship in turn leads to the assumption that time estimation behavior will vary according to the way that duration judgments are assessed. In prospective situations, research participants are informed in advance that a time judgment will be required, and so these estimates are assumed to reflect experienced duration. Given that participants are consciously monitoring time's passage, more attentional resources will be allocated toward the cognitive timer. This should result in highly accurate and reliable

estimates of a task's total duration but in relatively poor performance on the task involving nontemporal information. However, as task complexity increases, because of either greater difficulty (e.g., more sophisticated math problems) or the performance of two concurrent tasks (e.g., doing math while watching television), more resources will be allocated toward the processing of nontemporal information. The result is that time judgments will become not only more variable and inaccurate but also shorter because less temporal information is being encoded into the cognitive system. However, a different pattern of results is predicted to occur in retrospective designs. These are assumed to reflect remembered duration because participants are not told that a time judgment will be required until after an event has transpired. Here, attentional resources will be directed primarily toward the processing of nontemporal information because there is no apparent reason for encoding durational information into memory. Relative to the situation with prospective designs, then, remembered duration judgments should be less accurate and more variable. In addition, increases in task complexity are predicted not only to decrease the accuracy and reliability of these estimates but also to yield an overestimation bias that is due to an "availability heuristic." This concept is one that has been borrowed from Ornstein (1969) and Block (1985), who asserted that remembered duration is inferred from the amount of nontemporal information stored in memory. Given that people do not encode event duration under retrospective circumstances, they presumably rely on the only source of information that is available—namely, the number of chunks (Ornstein, 1969) or changes (Block, 1985) within the array of nontemporal information.

This overall approach to time estimation behavior has received empirical support. Comparisons of the accuracy and variability of estimates across the two types of designs reveal that performance is in fact better under prospective conditions than under retrospective conditions (Brown, 1985; Brown & Stubbs, 1992; Miller, Hicks, & Willette, 1978). In addition, there is evidence that increases in task complexity exert a differential effect on experienced versus remembered duration. For example, Hicks, Miller, and Kinsbourne (1976) found that when participants were asked to sort a deck of cards into two (by color) versus four (by suit) categories, the latter yielded significantly shorter estimates than the former under the prospective paradigm. In contrast, complexity manipulations within retrospective designs often reveal overestimation biases that are presumably due to the availability heuristic. McClain (1983), for example, varied both the difficulty and number of words presented to participants in a semantic processing task and found that more words led to longer retrospective judgments.

Although these types of studies support the underlying assumptions of the resource allocation model, there are other empirical findings that are more problematic. In particular, the model assumes that people should be incapable of remembering event duration unless duration was selectively attended to during the encoding phase. That is, duration should be well remembered in prospective situations but not in retrospective ones. This, however, is not the case. In

several recent studies, it was found that when listeners were repeatedly exposed to a set of melodies (Boltz, 1992, 1995) or sounds from the natural environment (Boltz, 1992, 1994) and were later unexpectedly asked to recognize or reproduce their total durations, performance was in fact highly accurate and reliable. In other words, provided that there has been sufficient learning experience, event duration can be encoded into the cognitive system under retrospective situations.

In sum, then, the literatures on music cognition and time estimation behavior yield conflicting evidence on the processing of temporal and nontemporal information. Some studies suggest the two dimensions are processed independently of one another, and others suggest they are jointly encoded into memory. One possibility is that both types of processing can occur and that this in turn depends on certain factors within the surrounding context. Two such factors are suggested here and involve the structure of the event and the individual's degree of learning experience. Both have been incorporated into a model called the *structural remembering approach*, which was developed primarily to address time estimation behavior but also has implications for music cognition.

The Structural Remembering Approach

The structural remembering approach stems from a more general theory of attending (Jones & Boltz, 1989) that has been extended to address the remembering of temporal and nontemporal information within environmental events (Boltz, 1992, 1995). In contrast to other frameworks, the structural remembering approach places the primary emphasis on the role of the environment in determining both the nature and the success of attending and remembering activities. In particular, the structural remembering approach assumes that the degree of coherence between the temporal and nontemporal information of a given event will influence encoding and the ability to subsequently recapitulate the event's structure from memory.

If one were to systematically analyze the structure of different environmental events, one would find that in many cases, the arrangements of temporal and nontemporal infor-

mation directly coincide. In particular, it can be shown that an event's rhythmic structure is compatible with and reinforces the hierarchical arrangement of its spatial information. Although many such events can be identified in the natural environment, one of the best illustrations can be found in Western music. Nontemporal, or melodic, structure is based on the underlying key of a melody, which often can be derived from the strategic placement of tonic triad members (i.e., the tonic, the mediant, and the dominant). The tonic not only is used to end a melody but, along with the mediant and the dominant, often serves to mark phrase ending points within a melody's total time span (Berry, 1976; Cone, 1968; Kramer, 1982). Melodic phrases represent themes and variations; if, before shifting to new themes and variations, one marks the endings of these melodic phrases with tonic triad members, momentary points of resolution are provided within a tune as well as a reaffirmation of the underlying tonality scheme. For example, consider the opening bars of "The Star Spangled Banner," which are shown in Figure 1. This melody is written in the key of B major, and the phrase endings, which coincide with the words *see, light, hailed,* and *gleaming,* therefore correspond to the tonic, dominant, tonic, and tonic, respectively. Notice that these phrase endings also contribute to an overall hierarchical arrangement in that pairs of melodic phrases are nested within higher order phrases of a longer time span (i.e., 12 beats).

Although the nontemporal information (i.e., the array of pitch relations) of this type of melody is itself quite lawful, its overall coherence emerges from the pattern of temporal information within the tune's total time span. All melodies contain a regularly recurring beat or pulse that tends to remain relatively constant throughout. In musical terms, this is rendered by the underlying meter of the tune. The melody in Figure 1, for example, is based on a ¾ meter in which the 4 indicates that a quarter note (♩) is the referent beat and therefore receives one beat of time, and the 3 indicates there are three beats per measure. Against this background beat, temporal patterning is introduced through rhythm. Rhythm refers to durational changes from the referent beat in which longer values, arising from either pauses or prolonged note

Figure 1. An example of a musical composition, "The Star Spangled Banner" by Francis Scott Key, that displays a highly coherent arrangement of temporal–nontemporal information. Within the melody's total time span, sequences of notes are nested within four melodic phrases (p) marked by tonic triad members (the tonic, dominant, tonic, and tonic, respectively). These phrase ending points are attentionally highlighted by temporal accents (♩), and both regularly recur in time to outline higher order phrases.

durations, create temporal accents that capture a listener's attention. What is important is where these accents occur. In most Western music, temporal accents tend not only to recur after a predictable number of beats but also to coincide with melodic phrase endings. In Figure 1, for example, notice that temporal accents, arising from the relatively longer half notes ($\downarrow$), regularly recur after every six beats and correspond with phrase boundaries. Structurally, then, the overall pattern of temporal beats forms its own hierarchical arrangement whose nestings directly coincide with those formed by nontemporal information to yield a joint accent structure. Within this structural array, the two dimensions codefine one another: The hierarchical nesting of temporal periodicities is defined by different levels of pitch change, and the nesting of pitch changes is defined relative to certain temporal periodicities (Boltz, 1991; Boltz & Jones, 1986; Lerdahl & Jackendoff, 1983).

Although this example of high coherence may seem idealized and specific to music, it generalizes to a wide variety of natural events. Conversational interaction, for example, displays prolonged pauses that reinforce grammatical phrase structure (Goldman-Eisler, 1972; Grosjean, Grosjean, & Lane, 1979) as well as intonational accents that coincide with high-content words (Martin, 1972; Shields, McHugh, & Martin, 1974). Similarly, the boundaries of constituent action units are often marked by reductions in amplitude and overall rate of movement (Stetson, 1905; N. Todd, 1985).

From a cognitive perspective, highly coherent events facilitate encoding and remembering in a number of different ways. First, and most important, given that the event's temporal and nontemporal information structurally coincide, both can be encoded into memory with no extra effort—attending to one should lead to the incidental learning of the other. Hence, even though participants may have initially attended to one dimension alone, the other dimension should be equally well learned and later remembered. In addition, in contrast to the predictions of the resource allocation theory, attending to both dimensions at once should not produce any decrement in performance. This encoding process should be facilitated by the event's structure because the regularly recurrent pattern of rhythmic accents not only allows one to easily attend to and track the event as it unfolds but also highlights the event's overall organization for learning. Similarly, just as accents can be used to guide attending, they also offer cognitive referents for reconstructing the event's structure from memory. This then leads to the prediction that both the event's temporal information and its nontemporal information will be remembered with a high degree of accuracy. Given that both structural dimensions jointly characterize the event, remembering an event's time span means retrieving both the temporal and nontemporal information within this span.

Although many natural events are highly coherent, others are not. In the case of music, incoherence may arise when the arrangement of either temporal or melodic structure alone is relatively unpredictable in nature or, alternatively, when the array of temporal accents conflicts with the organization of melodic relations. These types of events, because of their apparent lack of organization, not only will be difficult to attentionally track but will require additional effort to encode both temporal and nontemporal information. Given that the two dimensions are structurally incompatible, attending to both at once will lead to interference effects that impair both learning and remembering. Instead, participants must selectively attend to each dimension alone and independently of the other dimension in order to successfully encode these events into memory.

The degree of structural coherence that is both produced and perceived within an event is, to some extent, determined by the individual's degree of learning experience. At early stages of learning, events that result from the motor behavior of an individual are often relatively incoherent. For example, the music of a novice pianist or the speech of a foreigner learning a new language often displays a dissociative rhythm in which pauses fail to coincide with grammatical phrase boundaries. Given that the temporal patterning of behavior does not coincide with the underlying hierarchical arrangement of nontemporal information, the two dimensions have become relatively independent of one another. With additional practice, however, the two will eventually display a joint accent structure. Similarly, a structurally coherent event may not appear to be so when it is initially encountered. Because there has not been sufficient opportunity to discern the lawful relationship between temporal and nontemporal information, the overall structure of the event may be perceived as relatively incoherent.

In sum, the approach being presented here is one suggesting that the encoding and remembering of natural events are strongly dependent on both learning and the structure of the event itself. In the case of highly coherent events, there is no need to divide attending between temporal and nontemporal information, as the resource allocation theory has claimed. Because the two dimensions are jointly encoded into memory, attending to temporal or nontemporal information alone should result in the incidental learning of the other dimension. Both dimensions, then, will be represented in memory with the same amount of processing effort required to input one dimension alone. This facilitating effect should be particularly enhanced with greater learning experience, because the lawful organization of the event will be more apparent. From a behavioral standpoint, this means that even though participants may be primarily focused on an event's nontemporal information during the encoding phase, the retrieval of this information should reflect the overall rate and rhythmicity of items as well as the event's total time span. Incoherent events, on the other hand, display temporal and nontemporal information that are structurally dissociated. Although both dimensions can be made available in memory, this availability will require more effort because each dimension needs to be independently attended to in order for encoding to occur. If participants are required to divide their attending between both dimensions at once, interference effects will arise that lead to trade-offs in performance such that increased attending to one dimension will decrease performance on the other dimension.

These ideas have direct implications for those issues previously discussed in the areas of music cognition and time estimation behavior. In musical behavior, pitch and rhythm would be expected to exert an interactive influence on one another if melodies display a high degree of coherence and the listener-performer has had sufficient learning experience—otherwise, an independent relationship should emerge. With respect to time estimation behavior, the constructs of learning and event structure may help to explain why the resource allocation model (Zakay, 1989) assumes that people are relatively incapable of remembering event durations with any degree of accuracy. Much of this research has assessed retrospective time judgments after a single learning trial in which the inherent structure of the event was not yet apparent. In order to encode both the event's temporal and nontemporal information into memory, the participant must therefore attend to each dimension independently. Hence, in retrospective circumstances, where participants are presumably not attending to event duration alone, remembered duration judgments will be quite poor, as in fact has been reported in the previous literature. However, a very different set of results should emerge if participants are asked to remember the durations of relatively coherent events after a sufficient learning period. Here, given that both structural dimensions can be jointly encoded into the cognitive system, duration memory should be quite accurate regardless of whether attending was primarily directed toward temporal (i.e., prospective designs) or nontemporal (i.e., retrospective designs) information during the encoding phase.

In a recent set of experiments (Boltz, 1992), some support for these ideas was observed. In contrast to previous retrospective duration studies, in this set of experiments a new technique was developed that enabled the collection of multiple time judgments from an individual participant (Boltz, 1992). With this procedure, participants were first asked to perform a series of perceptual rating tasks on a set of musical patterns that varied in their structural coherence. These sounds contained a single pitch value that (a) regularly recurred in time and was thereby coherent, (b) was continuous in duration, or (c) was irregularly timed to yield an incoherent pattern. After performing the set of rating tasks for a varying number of trials, participants were then unexpectedly asked to either recognize or recall the total durations of those sounds heard earlier. Results indicated that participants were able to learn and remember the characteristic durations of events but that performance was strongly influenced by both learning and the events' structure. Overall, the durations of coherent and continuous events were well remembered, with a high degree of accuracy and low variability, and performance significantly improved with greater learning experience. In other words, event duration was encoded into memory even though participants were not expecting any sort of upcoming memory test. In contrast, people were relatively incapable of remembering the durations of incoherent events—a finding that is consistent with the resource allocation model under retrospective circumstances.

Rationale of the Present Experiments

The purpose of the present set of experiments was to extend this research by examining the processing of temporal and nontemporal information in a more rigorous fashion. Although a previous study (Boltz, 1992) indicated that event durations can be remembered in retrospective circumstances, it did not allow a determination of whether the encoding of temporal and nontemporal information varies as a function of event structure. One way to investigate this issue is to systematically direct participants' attending to different aspects of an event's structure and then examine the corresponding influence on duration judgments. This strategy is used in Experiment 1, which relies on a design that is very similar to that of a previous study (Boltz, 1992). Participants were presented with a set of melodies during an initial learning phase in which tunes varying in their overall coherence were rated on several perceptual dimensions. The main manipulation involved the type of instructions administered to participants during the encoding phase and the type of memory task they expected to later receive. In particular, there were four between-subjects conditions. One was a control condition (i.e., no attending) in which participants were not expecting any memory task at all. However, in the three remaining conditions, participants were asked to attend to (and remember) either the melodies' temporal (i.e., total duration) information or nontemporal (i.e., array of pitch intervals) information alone or in tandem. After performing the set of rating tasks, all participants were then required to perform only one memory task—namely, reproducing the melodies' total durations.

The manipulations of this experiment allow one to directly compare predictions from the structural remembering approach with those from the resource allocation model of time estimation behavior. Whereas the former assumes that a differential pattern of results will occur for coherent versus incoherent melodies, the latter assumes that variations in selective attending will exert the same influence on all events. According to the structural remembering approach, the temporal and nontemporal information of highly coherent events are jointly encoded into memory with no extra effort, and attending to one dimension will result in the incidental learning of the other dimension. If this is true, then participants should show highly accurate and reliable duration judgments regardless of whether they attend to duration or pitch alone during the encoding phase. In addition, attending to both dimensions at once should not produce a decrement in performance. Last, the control condition, in which no explicit attending instructions are given, is representative of the methodological procedure used in past studies of retrospective (remembered) duration judgments. The fact that participants are free to attend to any dimension(s) they choose allows one to assess how performance in this condition compares to performance in conditions where the focus of attending is more rigorously controlled. In particular, if the temporal and nontemporal information of coherent events are in fact jointly encoded into memory, then performance in this control condition should be comparable to that in the remaining conditions.

Incoherent events, on the other hand, are predicted to yield results consistent with resource allocation theory. Given that these events display temporal and nontemporal information that are structurally dissociated, each dimension must be selectively attended to in order for encoding to occur. This in turn should lead to trade-offs in performance as a function of the attentional manipulations. Although time judgments should be quite accurate in the prospective condition, in which participants are expecting a duration memory task, the same will not be true in the remaining conditions. In the pitch-alone and no-attending conditions, participants do not know a time judgment will later be required and so there is no apparent reason to selectively attend to durational information during encoding. Duration memory performance, then, should be highly inaccurate and unreliable. In addition, performance should decline even more when participants are asked to divide their attention between both pitch and duration. Given that the temporal and nontemporal information of these melodies are structurally incompatible, this type of dual attending activity should result in interference effects that disrupt the learning and remembering of both dimensions.

Experiment 1

Method

Design and Participants

The design was a 2 × 2 × 4 mixed factorial. All participants listened to a set of folk tunes that varied in their degree of structural coherence (coherent or incoherent). The two between-subjects variables were melody set (A or B) and the particular structural dimension of melodies to which participants were instructed to attend (pitch alone, duration alone, both pitch and duration, or no attending).

Forty-eight students from an introductory psychology course at Haverford College participated in the experiment for course credit. Each had normal hearing but no formal training in music. Six students were randomly assigned to each of the eight between-subjects conditions.

Stimulus Materials

A set of folk tunes from other countries was selected from a music composition book.[2] In order to ensure that these melodies were unfamiliar to participants, a preliminary rating study was conducted with an independent group of 10 experienced musicians; any tune that was judged even slightly familiar to listeners was eliminated from the experiment. The four melodies selected as experimental stimuli are depicted in Figure 2. Although these melodies represent diverse keys (C, G, and F major) and overall pitch contours (the ups and downs of pitch change), all contain 24 notes and end on the leading-tone followed by the tonic.

Within their original compositions, each of the four melodies displayed a hierarchical arrangement of melodic structure that coincided with its inherent rhythm. As can be seen in the top half of Figure 2, melodic phrase endings (p) were marked by a tonic triad member (i.e., tonic, dominant, or mediant) to provide momentary points of resolution and a reaffirmation of the melody's underlying tonality scheme (Berry, 1976; Cone, 1968; Kramer, 1982). These phrase ending points occurred after every sixth note in two of the melodies ("The Rabbit and the Hare" and "In My Dreams") and

after every eighth note in the remaining two melodies ("Lithuanian National Hymn" and "Fond Regrets"). In addition to displaying this lawful arrangement of melodic structure, each melody also displayed its own unique rhythm of short and long notes that then repeated throughout the melody's total time span. In Figure 2, for example, notice that "The Rabbit and the Turtle" is based on the recurrent pattern of ♪♪♩♪♪♩. More important, temporal accents arising from prolonged dotted half notes (♩.) always coincide with phrase ending points to yield highly coherent melodies with joint accent structures: Temporal accents occurred on every sixth note for those tunes displaying six-note phrases and on every eighth note for those melodies displaying eight-note phrases.

Incoherent melodies were obtained by assigning each of the four rhythms in the accent-compatible condition to a different melody such that temporal accents no longer coincided with phrase ending points. These are shown in the bottom half of Figure 2. For example, notice that the rhythms associated with "The Rabbit and the Turtle" and "Lithuanian National Hymn" were exchanged with each other such that temporal accents occurred on every eighth note for a tune based on six-note phrases and on every sixth note for a melody based on eight-note phrases. A similar exchange of rhythms was applied to the melodies of "In My Dreams" and "Fond Regrets."

Across all rhythmic conditions, the placement of temporal accents on every sixth note always yielded tunes based on a ¾ meter, and those accents occurring on every eighth note led to melodies based on a ⁴⁄₄ meter. In addition, tonal duration (on-time) and intertone interval (off-time) were held constant such that the value of an eighth note (♪) always corresponded to 150 ms on and 75 ms off; the value of a quarter note (♩), to 300 ms on and 150 ms off; and the value of a dotted half note (♩.), to 900 ms on and 450 ms off. All melodies displayed a total duration of 10.35 s.

For the initial learning phase of the experiment, melodies were sorted into two different sets (A and B), each containing four melodies. The four melodies within a set were then presented as a block of trials for each of the four perceptual rating tasks. Melodies within a set were randomly ordered, and this randomization varied across the different tasks. Given that the four rating tasks were repeated a second time, there were a total of 32 learning trials—meaning that participants heard each of their four tunes on eight different occasions.

Within each melody set (A and B), two of the four melodies displayed a highly coherent accent structure and the remaining two melodies showed a relatively incoherent accent structure. The two sets differed from each other in terms of the particular melody assigned to a given rhythmic condition. In Set A, the two accent-compatible melodies were "The Rabbit and the Hare" and "Lithuanian National Hymn," and the two accent-incompatible melodies were "Fond Regrets" and "In My Dreams." In Set B, those tunes that were previously accent compatible were now accent incompatible and vice versa. This arrangement meant that each of the four melodies within a given set displayed its own unique rhythm.

Apparatus

All melodies were constructed and generated with the Midilab software system (R. Todd, Boltz, & Jones, 1989). During each experimental session, the melodies were presented on-line with a Yamaha TX81Z FM tone generator controlled by an IBM AT

---

[2] The original versions of these folk tunes were selected from *Folk Songs of Many People* (Vol. I) edited by Florence Hudson Botsford (1921).
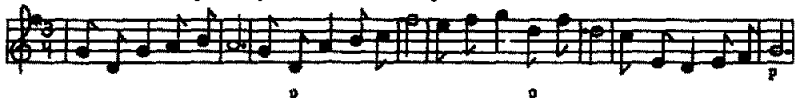
Figure 2. The set of melodies used in Experiments 1 and 2. The upper panel depicts tunes displaying a highly coherent structure in which temporal accents (♩) regularly recur and always coincide with melodic phrase endings (p). Incoherent melodies, shown in the lower panel, display a recurrent pattern of temporal accents but ones that fail to coincide with phrase ending points. (These folk tunes were selected from *Folk Songs of Many People,* Vol. 1, edited by Florence Hudson Botsford, 1921, New York: Woman's Press. In the public domain.)

computer with a Roland MPU-401 Midi interface unit. Sequences of tones were amplified by a Kenwood KR-4010 receiver and played over Koss-Pro 4AAA Plus headphones at a comfortable listening level (approximately 60 dB SPL). The timbre of all melodies was that of a grand piano.

Procedure

Recorded instructions informed participants of pattern presentation details and task requirements. For the initial learning phase, participants were told the experiment involved a study of perceptual judgments and that they would be asked to rate a set of four

melodies on four different dimensions. In addition, different groups of participants were instructed to attend to different aspects of the melodies because they would later receive a memory test on these dimensions. In particular, participants in the *pitch-alone* condition were told to attend to the pitch properties of all tunes, whereas participants in the *duration-alone* condition were asked to attend to their total durations. In the *both* condition, participants were instructed to attend to both the pitch properties and the total duration of all tunes, and participants in the *no-attending* condition were given no information at all about an upcoming memory task.

All participants then received the set of learning trials, in which a separate trial block was presented for each rating task and the

instructions for each task were administered before the presentation of four trials. On each trial, a 1-s warning tone (C_7, or C in the seventh octave) preceded the presentation of a melody by 2 s. The name of this melody was always indicated in the response booklet that had been distributed to all participants. Two seconds after the melody's offset, this same tune was presented again. During a 5-s response period which followed, participants were then asked to provide a rating on a 10-point scale by circling an appropriate value in their response booklet. For ratings of familiarity (1 = *very familiar*), participants were asked to decide whether the tune seemed at all familiar and to be one they had heard in the past. Ratings of predictability directed attention toward the pitch and timing qualities of each melody; participants were asked to judge whether these seemed very regular, such that it was easy to track the melody and anticipate its future course (1 = *very predictable*), or unpredictable in nature, such that it was very difficult to follow the tune's ongoing and future course. The third task involved pleasantness ratings and required participants to decide whether each melody was pleasing to listen to or aversive in its sound qualities (1 = *very pleasing*). Last, the ease-of-learning ratings required participants to decide whether each melody would be easy or more difficult to learn and remember (1 = *very easy*). All participants performed these tasks in the same order as outlined above and were then asked to make this set of ratings a second time with a separate response booklet. A previous study (Boltz, 1992) had shown that this degree of learning experience was sufficient for producing relatively accurate levels of duration memory.

Following this initial phase, all participants were then given a recall task for melody duration only. A handout was distributed that indicated the name of the particular melody that was to be recalled on each trial. During a given trial, a 1-s warning tone (C_7) was played, and after its offset, participants were asked to "mentally imagine the duration of the melody until it reaches its original ending point." Participants were to indicate this ending point by a press on a handheld push button which was automatically recorded by the computer. Participants were asked to recall the duration of each of the four melodies from the initial learning phase on two different blocks of trials, and trials were randomized both within and between blocks.

Participants were tested in groups of 3–4 individuals, and the duration of an entire experimental session was approximately 45 min.

Results

Data from the rating and duration recall tasks were analyzed separately through a series of analyses of variance (ANOVAs). Each is discussed in turn.

Rating Data

Table 1 depicts the mean ratings of familiarity, pleasantness, predictability, and ease of learning as a function of the melodies' degree of structural coherence and attended-to dimension. Means have been collapsed over melody set because this variable exerted negligible effects ($F < 1$).

For ratings of familiarity, the overall ANOVA revealed no significant effects and indicated that the sets of melodies were not recognizable by participants. There were also nonsignificant results for the pleasantness ratings, and as can be seen in the second column of Table 1, these means all clustered around the mid-range of the scale. The set of predictability ratings, however, varied as a function of the

Table 1
Mean Ratings of Familiarity, Pleasantness, Predictability, and Ease of Learning for Melodies in Experiment 1 as a Function of Event Coherence and the Attended-to Dimension of the Melody's Structure

Attended-to dimension	Familiarity	Pleasantness	Predictability	Ease of learning
No attending				
Coherent	8.64	5.42	2.54	3.37
Incoherent	8.51	5.19	8.18	8.05
Duration alone				
Coherent	8.79	5.23	2.52	3.24
Incoherent	8.33	5.47	8.01	7.91
Pitch alone				
Coherent	8.07	5.17	2.73	3.41
Incoherent	8.61	5.31	8.12	7.89
Both				
Coherent	8.57	5.33	2.68	3.29
Incoherent	8.42	5.28	7.97	7.83

Note. Lower values indicate a higher degree of each perceptual dimension.

melodies' degree of structural coherence, $F(1, 40) = 68.89$, $MSE = 2.01$, $p < .001$. Across all of the attended-to dimensions, highly coherent melodies were judged significantly more regular than were less coherent tunes. Last, the ease-of-learning ratings showed a similar pattern of results. A main effect for structural coherence, $F(1, 40) = 42.17$, $MSE = 3.27$, $p < .001$, revealed that coherent melodies were also judged to be easier to learn and subsequently remember than were those tunes that contained conflicting accent structures.

In sum, these data indicate that the melodies' familiarity and pleasantness were comparable across the different experimental conditions. Melodies within the different accent conditions, however, were judged to vary in their degree of internal predictability and ease of learning, which thereby validates the main manipulation of this experiment.

Absolute Errors Scores

In order to assess reproduction accuracy, all duration judgments (shown in the second column of Table 2) were transformed into absolute error scores. These scores are considered to be the most sensitive measure of judged duration accuracy (Michon, 1972) because they calculate the extent to which participants' responses deviate from an event's actual duration irrespective of the deviation's sign (i.e., over- or underestimation). Absolute error is determined by first subtracting a participant's reproduced duration from a melody's actual duration, dividing this absolute difference by the melody's actual duration, and then multiplying by 100. These values are depicted in the first column of Table 2 as a function of structural coherence and the attended-to dimension. Once again, scores were collapsed over melody set because this variable exerted no significant effects on performance.

The overall ANOVA revealed two significant effects. First, a main effect for structural coherence indicated that

Table 2

Mean Performance in the Duration Reproduction Task of Experiment 1 as a Function of Event Coherence and the Attended-to Dimension of Melody Structure

Attended-to dimension	Mean percentage absolute error		Reproduced durations (in seconds)		Ratio scores	
	M	SD	M	SD	M	SD
No attending						
Coherent	12.78	3.87	11.472	0.401	1.11	0.06
Incoherent	20.18	24.12	12.139	2.082	1.17	0.28
Duration alone						
Coherent	6.92	2.23	11.066	0.231	1.07	0.09
Incoherent	9.66	9.07	11.350	0.939	1.10	0.20
Pitch alone						
Coherent	8.59	4.11	11.239	0.425	1.09	0.06
Incoherent	29.03	22.84	13.355	1.950	1.29	0.31
Both						
Coherent	7.00	3.09	11.075	0.320	1.07	0.07
Incoherent	34.54	26.75	6.975	3.105	0.67	0.51

when temporal accents coincided with melodic phrase endings, duration was more accurately recalled than when temporal–melodic accent structures conflicted with one another, $F(1, 40) = 78.15$, $MSE = 26.69$, $p < .001$. This effect was a reliable one that applied to 42 of the 48 participants in the experiment.

Second, results also indicated that the attended-to dimension of a melody's structure exerted a differential effect on conditions of varying coherence, $F(3, 40) = 11.19$, $MSE = 26.69$, $p < .01$. First consider the coherent melodies of Table 2. Notice that the reproduction of total duration was least accurate in the no-attending condition ($M = 13\%$ error), in which participants had no knowledge of any sort of upcoming memory test. In contrast, performance was markedly enhanced in the three remaining conditions, in which participants were aware they would later be tested on the melodies' structure. More important, there were no significant differences among these latter three conditions. Those participants who attended solely to a melody's temporal or nontemporal information during the initial learning stage performed at the same level as participants who attended to both structural dimensions. This overall pattern of results was confirmed by a set of Tukey post hoc comparisons with p set at .05.

A very different pattern of results appeared with the remembering of incoherent melodies. Here, the classical prospective condition, in which participants were informed of the upcoming time estimation task, yielded the highest level of performance, with an average of 10% error. Relative to error in this condition, absolute error significantly increased when participants were not aware of any sort of subsequent memory task. Last, the percentage of error in duration memory further increased when participants were either attending to nontemporal information alone or were dividing their attention between both structural dimensions. Participants in the pitch-alone condition showed a 9% increase in error relative to participants in the no-attending condition, and an additional 6% increase relative to participants in the both condition. Again, these overall differences

were supported by a set of Tukey post hoc comparisons ($p = .05$).

This interaction was a highly robust one that applied to all four melodies in the experiment. By looking at the standard deviation of the absolute error scores in Table 2 one can also see that these effects were reflected in the variability of participants' responses across the different conditions.

Ratio Scores

A second analysis was conducted in order to assess the directionality of participants' time judgments. Recall that the absolute error scores measure reproduction accuracy by ignoring the signed difference between a participant's response and the actual duration of a melody. However, it may be that the manipulations of temporal accent structure in this experiment exerted a systematic influence on the direction of temporal judgments such that some conditions were more apt than others to yield over- or underestimations of duration recall. Hence, for each participant, a set of ratio scores (RD/AD) was calculated that evaluated reproduced duration (RD) relative to the melody's actual duration (AD). Relative to a ratio of 1 (representing no bias), ratios less than 1 indicate relative underestimations whereas ratios greater than 1 indicate relative overestimations. These mean ratio scores are depicted in the third column of Table 2.

Results from an overall ANOVA revealed a significant interaction between structural coherence and the attended-to dimension, $F(3, 40) = 27.12$, $MSE = 0.12$, $p < .001$. As can be seen in Table 2, ratios for the highly coherent melodies did not vary across the different instructional groups, and although they suggested a tendency toward overestimation, a set of Tukey post hoc comparisons indicated that they did not significantly deviate from a value of 1. In effect, bias was absent in these conditions. Incoherent melodies, on the other hand, yielded systematic over- and underestimations as a function of the attentional manipulations. Relative to participants in the prospective (duration alone) condition, which produced no significant bias, participants in the no-attending

condition overestimated event durations ($p < .05$). In addition, the magnitude of this bias increased ($p < .05$) when participants focused on the nontemporal information of their set of melodies (i.e., pitch-alone condition). This bias was a relatively consistent one that applied to 75% of participants in the former condition and 83% in the latter. Conversely, underestimations arose when participants were attending to both temporal and nontemporal information during the initial learning phase ($p < .05$). This too was a robust finding that occurred for 75% of the participants in this condition.

Last, it is informative to note that the variability of responses was also significantly greater in the accent-incompatible condition than in the accent-compatible condition.

Discussion

The results of Experiment 1 indicate that although variations in selective attending can influence duration memory, these effects are also dependent on the coherence of event structure. The incoherent melodies of this study, which displayed temporal accents that failed to coincide with melodic phrase endings, yielded a pattern of time judgments that was consistent with predictions of the resource allocation theory (Zakay, 1989). Overall, these judgments suggest that the temporal and nontemporal information of these events must be selectively attended to for encoding to occur. Here, the highest degree of accuracy and reliability was observed in the prospective situation in which participants knew that a duration recall task was forthcoming. This therefore enabled participants to direct attending solely toward this dimension and to engage in conscious time-keeping strategies to aid memory performance (e.g., tapping to or counting the number of beats within melodies). However, when attending was directed elsewhere or was shared with the processing of nontemporal information, performance markedly declined. In the no-attending condition, in which participants were not aware of an upcoming memory task, participants were free to direct attending toward pitch or timing information or both. Although duration accuracy declined here relative to that in the prospective condition, performance was nonetheless better than when participants were instructed to attend to and remember the array of pitch intervals. This suggests that at least some participants in the no-attending condition elected to attend to duration information alone during the initial learning phase, which resulted in less error than in the pitch-alone condition, in which attention was directed exclusively toward nontemporal information. Finally, the lowest level of performance occurred when participants were attempting to learn both pitch and duration information during the encoding phase. The fact that participants were confronted with conflicting arrays of information led to interference effects such that duration could not be encoded with the same facility as when it was the sole dimension attended to. In sum, then, this overall pattern of results converges with the results of other studies that assessed duration judgments in prospective versus retrospective circumstances (e.g., Brown, 1985; Brown & Stubbs, 1992; Zakay, 1989, 1993).

The novel findings of this study emerged with regard to the duration memory of highly coherent melodies. In contrast to the predictions of the resource allocation model, the durations of these melodies were well remembered regardless of which structural dimension was selectively attended to. The results of the no-attending condition are particularly interesting because even though participants could choose to attend to temporal or nontemporal information alone or in tandem, the degree of error in duration memory was nonetheless quite low. However, results also showed that performance improved when participants were aware that they would later be tested on the melodies' structure and could therefore engage in memory-aiding strategies. Moreover, attending to a melody's duration led to the same level of accuracy as attending to the array of pitch intervals, and performance did not decline when attending was shared between the processing of temporal and nontemporal information. As a set, these results indicate that the two structural dimensions were jointly encoded. Even when participants focused exclusively on nontemporal information, duration information was learned incidentally with the same degree of efficacy as intentional learning. This makes sense given that the temporal and nontemporal information of these events are inextricably entwined and structurally reinforce one another.

The second notable finding of Experiment 1 involves certain systematic biases that emerged as a function of the experimental manipulations. In general, the ratio scores for highly coherent events indicated that the probability of an overestimation was equal to that of an underestimation. Across all four conditions of the experiment, 54% of the participants overestimated coherent durations, whereas 46% underestimated these time spans. Bias was also absent when participants recalled incoherent durations under prospective circumstances. However, significant overestimations were observed when participants were primarily attending to the nontemporal information of incoherent events during the learning phase. For example, overestimations appeared in the no-attending condition for those participants who apparently elected to attend to pitch information, and overestimations further increased in the pitch-alone condition, where participants were focused exclusively on the melodies' pitch relations. This bias has been reported in previous studies (Boltz, 1995; McClain, 1983; Zakay, 1993) and is consistent with the use of an availability heuristic. Given that event duration has not been encoded into memory, participants rely on those characteristics of nontemporal information that are available, such as the number of chunks (Ornstein, 1969) or changes (Block, 1985), to infer duration estimates. Incoherent events will seem particularly long because their overall lack of an underlying organizational scheme creates an array of seemingly disjointed items. Similarly, an availability heuristic may also be responsible for the underestimation bias that emerged when participants were simultaneously attending to both the temporal and nontemporal information of incoherent melodies. As noted by Zakay (1989), divided attending not only disrupts any conscious

time-keeping strategies that may be invoked but also results in less nontemporal information's being encoded into memory. In either case, the result is shorter estimates of judged duration.

Experiment 2

The purpose of Experiment 2 was to more thoroughly examine the relation between temporal and nontemporal information as a function of variations in selective attending and event structure. In particular, two additional issues were raised concerning those processes involved in remembering. The first concerns whether other types of temporal information, beyond an event's total duration, can be learned incidentally and later retrieved from memory. With respect to music, one obvious candidate is a melody's rhythm. At least in highly coherent tunes, the pattern of temporal accentuation is assumed to guide attending to the underlying organization of melodic relationships and to provide cognitive referents for recapitulating the melody's total time span. If this sort of retrieval process does in fact occur, then participants should be able to remember a melody's pitch relations, total duration, and rhythmic patterning regardless of which structural dimension was selectively attended to during the encoding phase. All three types of information should be inextricably entwined in a melody's representation and retrieved as such during remembering. Incoherent events, however, display a very different type of structural arrangement. Given that temporal accents bear no meaningful relationship to melodic structure, they cannot be reliably used to recapitulate an event's duration. Hence, there is no reason to expect a melody's rhythm to be retrieved from memory unless it was selectively attended to during encoding. Furthermore, because rhythm is structurally incompatible with the accompanying melodic line, performance should decline when attending is divided between the processing of both temporal and nontemporal information. In sum, then, one goal of Experiment 2 was to determine whether rhythm is available in memory under different conditions of event structure and selective attending.

A second and related issue concerns whether there is a systematic relationship between the ability to remember an event's temporal and nontemporal information. According to Zakay's (1989) resource allocation theory, there is a trade-off: Attending to temporal or nontemporal information alone will decrease performance on the other dimension. This finding has in fact been reported in the previous literature. Brown and Stubbs (1992), for example, required participants to listen to a set of melodies while simultaneously performing a proofreading task. Results showed that in a prospective paradigm, time judgments of melodies were quite accurate but proofreading performance was poor. Conversely, the opposite pattern of results occurred in a retrospective paradigm: Although proofreading performance was high, time judgments were inaccurate.

In this latter study, it is important to note that the two tasks of music listening and proofreading were independent of one another (i.e., had no intrinsic interrelation) and therefore required separate processing resources. In many ways, this situation is analogous to the structure of incoherent melodies, in which temporal and melodic relations are also independent. Hence, if participants are asked to attend to a melody's rhythm, total duration, or the sequence of pitch relations, a similar trade-off in performance should be expected—although memory should be better on the dimension that was selectively attended to, performance on the other dimensions should be quite low. However, from the perspective of structural remembering theory, this overall pattern of results should not appear with the remembering of highly coherent melodies. Given that the temporal and nontemporal information of these events directly coincide, the remembering of pitch, of rhythm, and of total duration should all be quite high and comparable regardless of the dimension to which participants were instructed to attend. This should be particularly true when participants know that some sort of memory task will be forthcoming, which should encourage the use of memory-aiding strategies.

These issues were investigated by relying on essentially the same design as in Experiment 1. Once again, participants were presented with a set of coherent and incoherent melodies and asked to selectively attend to different structural dimensions while performing a series of perceptual rating tasks. However, two modifications were introduced. One involved the type of information participants were asked to attend to for a subsequent memory task. In addition to a group instructed to attend to total duration, a second group was asked to focus on the particular rhythm of each melody. As before, a divided attention condition was included, but here participants were informed that they would later be tested on pitch and rhythm. The two remaining conditions were identical to those of the previous study: Participants were either asked to attend to pitch information alone or were given no information at all about a subsequent memory task. After performing the set of rating tasks during the initial learning phase, all participants were then asked to perform three types of memory tasks: reproducing the melodies' total durations, reproducing their particular rhythms, and recognizing their sequences of notes.

Method

Only those methodological differences from Experiment 1 are described.

Participants and Design

The design was a 2 × 2 × 5 mixed factorial. All participants listened to the same set of folk tunes from Experiment 1, which varied in their degree of structural coherence (coherent or incoherent). The two between-subjects variables were melody set (A or B) and the attended-to dimension of a melody's structure (duration alone, rhythm alone, pitch alone, both pitch and rhythm, or no attending).

Eighty students, all nonmusicians, participated in the experiment for course credit. Eight were randomly assigned to each of the 10 between-subjects conditions.

Stimulus Materials

For the pitch recognition task, two new distractor melodies were created for each of the four melodies within a given set (A and B). These are shown in Table 3. Each distractor always preserved a melody's tonality and rhythm but displayed a series of new notes within one melodic phrase. In the two melodies containing six-note melodic phrases ("The Rabbit and the Turtle" and "In My Dreams"), one set of deviations was imposed on notes within the second melodic phrase (i.e., Notes 7 through 12) and a second was imposed on notes within the third melodic phrase (i.e., Notes 13 through 18). Similarly, melodies containing eight-note melodic phrases ("Lithuanian National Hymn" and "Fond Regret") were changed such that deviations occurred either within the second melodic phrase (i.e., Notes 9 through 16) or within the third melodic phrase (i.e., Notes 17 through 23). To the extent that it was possible, the set of deviated notes within a phrase preserved the melody's original contour (or pattern of ups and downs in pitch) and interval relationships: In most cases, notes were simply transposed up or down in pitch by one or two scale steps. There were three exceptions to these general rules. In two instances, the magnitude of a pitch change was either 0 or 3 scale steps, and in one instance, a contour change was reversed in direction. These exceptions, however, were necessary in order to render the new melodic phrases as plausible ones that could occur within a given melody.

In each melody set (A or B), melodies for the pitch recognition task were randomly arranged into a series of 16 trials such that there were an equal number of same–different trials: Each of the four tunes was presented twice in its original format and once in each of its deviated formats.

Procedure

For the set of perceptual rating tasks, participants were randomly assigned to one of the five attending conditions. Participants in the no-attending, pitch-alone, and duration-alone conditions received the same set of instructions as in Experiment 1. In the rhythm-alone condition, participants were told they would later be asked to remember each melody's rhythm and pattern of short, medium, and long notes. In the both condition, participants were told they would be asked to remember the particular rhythm and sequence of notes

for each of the four melodies they would be hearing. After completing the initial learning phase, all participants were given the duration recall task under the same procedure as before. Following this, participants were then asked to perform two additional tasks that were ordered in such a way as to minimize any potential carryover effects. In the first, participants were asked to recall the recurrent rhythm for each tune. In particular, they were told that each of four melodies they had heard earlier contained a total of 24 notes and that in two of these melodies, notes were arranged into four groups of six notes, whereas in the remaining melodies, notes were arranged into three groups of eight notes. Participants were also informed that all groupings within a given melody displayed a repetitive pattern of short, medium, and long notes that they should try to recall with the letters S, M, and L, respectively. Participants were given 5 min to complete this task on a provided sheet that contained the name of each melody they were to recall.

In the third and final memory task of the experiment, participants were asked to recognize each melody's sequence of notes. On each trial, a 1-s warning tone preceded a melody by 2 s. During the 7-s response period that followed, participants were instructed to circle "same" on their response sheet if the tune contained exactly the same sequence of notes as during the earlier rating task. Conversely, they were told to circle "different" if some notes appeared to have been changed. In addition to their same–different judgment, participants were asked to indicate the certainty of their response on a 7-point rating scale (7 = very sure).

Results

The data from the perceptual rating tasks are not presented because, overall, they revealed the same pattern of results as in Experiment 1. Findings from the three memory tasks are each discussed in turn. In all cases, results were collapsed over melody set (A or B) because this variable exerted a negligible effect on performance.

Duration Reproduction

Absolute errors scores. Column 1 of Table 4 displays the absolute error scores from the duration reproduction

Table 3
The Set of Deviated Notes for Melodies in the Pitch Recognition Task of Experiment 2

Song and deviation set	Notes																							
"The Rabbit and the Turtle"	E4	F4	A4	C5	B4	G4	D4	F4	A4	B4	G4	E4	A4	C5	A4	G4	G4	E4	G4	E4	D4	A3	B3	C4
Deviation Set 1								F4	A4	C5	D5	C5	G4											
Deviation Set 2													B4	D5	C5	F4	F4	G4						
"In My Dreams"	C4	F4	G4	A4	F4	C4	G4	A#4	D5	E5	C5	A4	D5	A#4	G4	E4	F4	A4	C5	A#4	D4	C4	E4	F4
Deviation Set 1								A#4	D5	F5	G5	E5	C5											
Deviation Set 2													F5	D5	A#4	G4	A4	C5						
"Lithuanian National Hymn"	G4	D4	G4	A4	B4	A4	G4	D4	A4	B4	C5	F#5	E5	F#5	G5	D5	F#5	D5	C5	E4	D4	E4	F#4	G4
Deviation Set 1									F#4	G4	A4	D5	C5	D5	E5	B4								
Deviation Set 2																	E5	C5	A4	F#4	E4	D4	E4	
"Fond Regrets"	C4	D4	E4	F4	D4	C4	B3	G3	D4	E4	F4	A4	F4	E4	D4	G4	A4	G4	F4	A4	G4	A4	B4	C5
Deviation Set 1									E4	F4	G4	B4	A4	F4	E4	G4								
Deviation Set 2																	B4	A4	G4	B4	A4	B4	D5	

Table 4
Mean Performance in the Duration Reproduction, Rhythm Recall, and Pitch Recognition Tasks of Experiment 2 as a Function of Structural Coherence and the Attended-to Dimension of Melody Structure

Attended-to dimension	Mean % absolute error in duration reproduction		Mean ratio of reproduced to actual duration		Mean % error in rhythm recall		Mean % hit rate in pitch recognition	
	M	SD	M	SD	M	SD	M	SD
Coherent melodies								
No attending	15	4.28	1.09	0.11	20	6.84	85	6.15
Duration alone	10	3.01	1.04	0.05	10	6.01	85	5.87
Pitch alone	9	3.58	1.07	0.09	13	5.89	98	2.00
Rhythm alone	8	2.97	1.05	0.07	8	4.15	87	5.18
Both pitch and rhythm	11	4.12	1.03	0.08	10	5.37	96	2.75
Incoherent melodies								
No attending	24	29.17	1.22	0.40	45	22.86	58	21.00
Duration alone	13	6.77	1.10	0.18	24	8.21	63	28.12
Pitch alone	37	28.33	1.35	0.34	49	23.03	84	12.75
Rhythm alone	10	7.24	1.09	0.22	15	7.64	60	24.09
Both pitch and rhythm	40	37.91	0.74	0.49	64	30.17	51	32.33

task, which, again, represent the mean percentage deviation between a melody's actual and reproduced durations. As indicated by an overall ANOVA, the pattern of performance was identical to that in Experiment 1. A main effect for structural coherence, $F(1, 70) = 62.17$, $MSE = 19.93$, $p < .001$, indicated that the durations of coherent melodies were recalled with less error ($M = 11\%$) than were those of incoherent melodies ($M = 25\%$). This effect was a relatively robust one that applied to 71 of the 80 participants in the experiment. In addition, there was also a significant interaction between structural coherence and the attended-to dimension, $F(4, 70) = 22.36$, $MSE = 25.44$, $p < .004$. In the case of coherent melodies, performance remained comparably high regardless of which structural dimension was attended to. Once again, there was no decrement when attention was divided between pitch and rhythm. The duration estimates of incoherent melodies, on the other hand, showed a very different pattern of results. Here, error was quite low ($M = 13\%$) when participants were expecting to be tested for event duration. In addition, error was comparably low when participants were attending to rhythm during encoding, which suggests that rhythm was successfully used as a retrieval scheme for a melody's total duration. Relative to these conditions, a set of Tukey post hoc comparisons indicated that duration estimates were significantly less accurate when participants were not aware of any upcoming memory task ($p < .05$) and even less accurate when attending was directed toward pitch information alone or was divided between pitch and rhythm ($p < .01$).

These effects on duration memory were also reflected in the variability of judgments. As can be seen in Table 4, the pattern of standard deviation scores directly corresponds to the pattern of accuracy scores.

Ratio scores. To determine whether participants' reproductions revealed any systematic pattern of bias, a set of ratio scores was obtained for each participant which evaluated reproduced duration relative to a melody's actual duration. These scores are shown in column 2 of Table 4.

The most important finding was a significant interaction between structural coherence and the attended-to dimension, $F(4, 70) = 19.01$, $MSE = 0.09$, $p < .01$. Across all conditions of attentional demand, a set of Tukey post hoc comparisons indicated that the ratio scores for coherent melodies did not significantly deviate from a value of 1—indicating that bias was absent in these conditions. Incoherent melodies, however, yielded a differential pattern of over- and underestimations. Although bias did not appear when participants were attending to either rhythm or total duration, significant overestimations occurred when participants were attending to a melody's nontemporal information during the encoding phase. This was most apparent when participants were focused exclusively on pitch information ($M = 35\%$), but overestimations also appeared in the no-attending condition, where reproduced durations were 22% longer than the melodies' actual durations ($p < .01$). This in turn suggests that some participants in this latter condition opted to attend to the melodies' pitch relations. Conversely, the divided-attention condition (both pitch and rhythm) produced significant underestimations on the order of 26%.

Rhythm Recall

In the scoring of the rhythm recall data, the pattern of short, medium, and long notes that repetitively recurred across all melodic phrases of a given tune was compared with a participant's reproduced protocol. Any durational value that inappropriately occurred at a given serial position was counted as an error. For each participant, the number of errors for a given melody was divided by 24 (the total number of note durations that could be misremembered) to yield a percentage error score. These mean values are depicted in the third column of Table 4.

In general, the overall pattern of results was similar to that of the duration accuracy (absolute error) data. As indicated by a main effect for structural coherence, $F(1, 70) = 33.46$, $MSE = 22.02$, $p < .005$, the degree of error in rhythm recall

was significantly lower for coherent melodies than for incoherent ones, an effect that applied to 72 of the 80 participants. In addition, the coherence of melodic structure exerted a differential influence on different attending conditions, $F(4, 70) = 29.11$, $MSE = 31.07$, $p < .01$. With coherent melodies, attending to a melody's pitch relations or total duration resulted in the same low level of error as did attending to rhythm itself during the encoding phase. More important, performance did not decline when attending was divided between the pitch and rhythmic structure of a tune. Finally, although error was significantly higher in the no-attending condition ($p < .05$), the relatively high level of performance here indicates that rhythm can be encoded into memory regardless of the structural dimension(s) to which participants choose to attend. In contrast, incoherent melodies once again displayed a different pattern of results. Error in rhythm recall was lowest when participants attended to a melody's rhythmic and duration information, but it increased in the pitch-alone condition ($p < .01$) and increased even more when attention was divided in the both condition ($p < .01$). As before, this overall pattern of results generalizes to the reliability of memory performance.

Pitch Recognition

The last column of Table 4 presents the mean percentage accuracy (i.e., hit rate) in the same–different pitch recognition task along with the variability of these responses. Overall, the pitch relations of coherent melodies were better remembered than were those of incoherent melodies, $F(1, 70) = 43.47$, $MSE = 17.87$, $p < .001$, and this was true for 75 of the 80 participants. Furthermore, the interaction between structural coherence and the attended-to dimension, $F(4, 70) = 18.74$, $MSE = 31.02$, $p < .01$, revealed a pattern of results analogous to that for the other memory tasks. In coherent melodies, although pitch was significantly better remembered when it was selectively attended to ($p < .05$), performance was still high ($M = 86\%$) when attending was directed toward the temporal dimensions of rhythm and duration. As before, there was no decrement when attending was shared with rhythm in the both condition. However, the pitch relations within incoherent melodies were not well remembered unless they were attended to during encoding. Relative to this condition, there was a marked decline when participants either were free to direct their selective attending (no attending) or focused on a melody's temporal information ($p < .01$). The level of accuracy further decreased when attention was divided between temporal and nontemporal information ($p < .05$).

Correlational Analyses

A final set of analyses was conducted in order to examine the extent to which participants' accuracy scores on the duration reproduction, rhythm recall, and pitch recognition tasks were intercorrelated. These intercorrelations are depicted in Table 5 and varied with both structural coherence and the attended-to dimension of a melody's structure. For coherent melodies, high positive values occurred in all

Table 5

Intercorrelations (rhos) Between Participants' Accuracy Scores on the Duration Reproduction, Rhythm Recall, and Pitch Recognition Tasks of Experiment 2 as a Function of Structural Coherence and the Attended-to Dimension of Melody Structure

Attended-to dimension	Duration and rhythm	Duration and pitch	Rhythm and pitch
Coherent melodies			
No attending	.76	.78	.80
Duration alone	.92	.80	.80
Pitch alone	.86	.84	.89
Rhythm alone	.89	.82	.83
Both pitch and rhythm	.84	.80	.91
Incoherent melodies			
No attending	.30	.24	.26
Duration alone	.87	.34	.30
Pitch alone	.33	.27	.24
Rhythm alone	.89	.23	.28
Both pitch and rhythm	.28	.31	.25

conditions ($p < .01$), which indicates that increased performance on one task was associated with increased performance on the other tasks regardless of which structural dimension was attended to during the encoding phase. Conversely, most of the intercorrelations for incoherent melodies were low, which suggests that performance on one task was independent of performance on another. Only two intercorrelations were significant, and these occurred in the duration reproduction and rhythm recall tasks, in which attending to one dimension of temporal structure led to a better remembering of the second dimension ($p < .01$).

Discussion

These overall results reinforce the idea that one's encoding strategy is dependent on event structure. The duration reproduction data directly converged with those of Experiment 1 and revealed the same pattern of accuracy and bias across the different experimental conditions. More important, however, the data from the rhythm recall and pitch recognition tasks provide further evidence that temporal and nontemporal information are jointly encoded into the cognitive system when event structure is highly coherent but independently encoded when events are less coherent.

For the incoherent melodies, results across all three memory tasks indicated that pitch, rhythm, and total duration could each be well remembered when these structural dimensions were solely attended to during the encoding phase. However, when attention was directed toward a melody's temporal characteristics, memory for the array of pitch intervals markedly declined. The reverse relationship was observed when participants attended to pitch and were later tested for the melodies' rhythms and total durations. Last, the lowest performance levels occurred when attending was divided between temporal and nontemporal information. These types of trade-offs in performance indicate that although both structural dimensions can be encoded into

memory, each must be attended to independently of the other.

In contrast, variations in selective attending exerted no influence on coherent melodies. In each of the three memory tasks, accuracy and reliability remained high regardless of which structural dimension was selectively attended to, and performance did not decline when attention was divided between the pitch and the rhythmic structure of the melodies. In addition, the set of correlational analyses indicated that performance on one task was strongly dependent on performance on other tasks such that an increased remembering of one dimension was associated with an increased remembering of the other dimensions. This in turn suggested that the temporal and nontemporal information of these types of events were jointly encoded into memory with no additional processing effort.

In sum, then, this overall pattern of results indicates that event structure is a primary determinant of the way that events are encoded in and later retrieved from memory. These effects are quite robust ones that generalized not only across the different memory tasks and to the majority of participants in the experiment but also to the pattern of accuracy and variability. It should also be noted that the participants in both Experiments 1 and 2 were all nonmusicians who had no formal training in musical structure. The drawback of this fact is that it mandated the use of a pitch recognition task in lieu of a pitch recall task (requiring music notation or musical performance abilities), which thereby made the mode of retrieval somewhat different from that used in the remembering of both rhythm and duration. Nonetheless, the recognition of pitch relations is a valid index of remembering, and given that the behavior of these nonmusicians systematically varied as a function of the experimental manipulations, this use of nonmusicians further ensures the overall reliability and generality of the results.

General Discussion

From a more global perspective, the results observed in the present set of experiments have implications for the areas of both music cognition and time estimation behavior. With respect to the latter, the results both modify and extend current theoretical perspectives on the remembering of event durations. Within the past literature, it has traditionally been assumed that the processing of temporal information is constrained by the relative allocation of attentional resources (Brown, 1985; Hicks et al., 1976; Thomas & Cantor, 1978; Zakay, 1989). Although duration judgments should be quite accurate in prospective situations where participants are consciously monitoring time's passage, duration memory is assumed to be highly inaccurate in retrospective circumstances in which participants do not know a time judgment is forthcoming. Because there is no apparent reason to attend to event duration, attentional resources are presumably directed toward nontemporal information alone. This particular structural dimension should therefore be well remembered, and given that duration has not been encoded into memory, it provides a basis for participants to infer an

event's total duration (i.e., via the availability heuristic). In general, the more chunks (Ornstein, 1969) or changes (Block, 1985) within the array of nontemporal information, the longer the event will seem. Last, conditions of divided attending are assumed to disrupt the remembering of both temporal and nontemporal information and to result in relative underestimations of judged duration. In these cases, insufficient resources are presumably available for processing both dimensions at once, and this insufficiency leads to less temporal and nontemporal information being encoded into memory.

This pattern of predicted results was in fact observed with the set of incoherent melodies used here. Although these findings are consistent with the notion of resource allocation, at a more fundamental level they can be attributed to the structure of the events themselves. Given the structural incompatibility of temporal and nontemporal information, any attempts to attend to both dimensions at once will result in interference effects. Hence, each dimension must be independently attended to in order to be successfully remembered.

The most important finding to emerge from the present data, however, is that these processes do not apply to all situations. One contribution of the structural remembering approach is the demonstration that the temporal and nontemporal information of many environmental events directly coincide and thereby facilitate cognitive processing. This joint array of information allows both structural dimensions to be encoded into memory with the same amount of effort required to input one dimension, which in turn yields a cognitive system that is both parsimonious and efficient. Overall, it ensures greater remembering because even though an individual may not necessarily attend to a given dimension at the time of encoding, this information will be learned nonetheless and represented in memory.

One condition that was of particular interest in the present set of studies was the no-attending condition, in which participants were not informed of any sort of upcoming memory task. This condition in fact is representative of the way that retrospective duration judgments are typically assessed in the time estimation literature. Although participants in this type of condition have traditionally been assumed to focus only on nontemporal information during the initial encoding phase, this is not necessarily true—participants are free to attend to any dimension(s) they choose. This in turn is important because different patterns of results are observed depending on the types of events that are used in the experimental context. When participants are presented with relatively coherent events, memory for temporal and nontemporal information will be quite accurate regardless of the dimension to which participants elect to attend, and as found here, no trade-offs in performance should be observed. However, if participants are presented with relatively incoherent events, then different levels of performance may be observed for different participants. Those participants who elect to attend to temporal information will yield fairly accurate estimates of judged duration but poor remembering of nontemporal information. Conversely, if participants elect to attend to nontemporal infor-

mation during encoding, then this dimension should be well remembered but temporal information should not. In either case, the present results indicate that memory performance in a no-attending condition will not be comparable to performance in conditions in which participants are instructed to intentionally learn one dimension or the other. At a more global level, these results have two implications. First, in terms of future research, an experimenter should not necessarily assume that all participants are attending to nontemporal information alone in retrospective research designs. In order to achieve this methodological goal, participants need to be explicitly instructed to attend to the event's nontemporal information. This will not only provide a more appropriate counterpart to prospective designs (in which participants are instructed to attend to event duration) but should also reduce the overall variability of participants' estimates. Second, the differential patterns of results that emerge for coherent and incoherent events may help to explain the conflicting nature of the remembered duration literature. When coherent events are used as experimental stimuli (e.g., Boltz, 1992, 1994, 1995), results reveal that, in contrast to those studies supporting the resource allocation theory, event duration can be accurately remembered in retrospective situations.

In addition to demonstrating that event duration can be incidentally encoded into memory under retrospective circumstances, the structural remembering approach makes a second contribution, which is to offer one explanatory account of the processes involved in remembering. In contrast to other theoretical frameworks, the structural remembering approach places a greater emphasis on the role of the environment and assumes that the same structural relations are used to guide both attending and remembering. Within coherent events, the array of regularly recurrent accents establishes an efficient course for attentional tracking that directs one toward the event's underlying organization of nested relations. Similarly, just as temporal accents can be used to guide attending, they can also be used as referential landmarks in recapitulating the event's total time span. This recapitulation includes both the event's temporal and nontemporal information because the two are inextricably entwined and jointly characterize the event itself. In Experiment 2, some support for these ideas was observed. Rhythm not only was incidentally learned (i.e., in the pitch-alone condition) with the same level of efficacy found in intentional learning, but it also facilitated performance in the pitch recognition task. Incoherent events, however, did not produce these types of effects on remembering because in incoherent tunes, rhythm was structurally dissociated from the pattern of pitch relations. Temporal accents, then, no longer guided attending toward the nested organization of nontemporal relations, nor did they serve as an effective retrieval scheme for pitch information. However, as found in Experiment 2, rhythm can aid duration memory when this latter dimension is solely attended to. Assuming that the pattern of temporal accentuation is a lawfully recurrent one, it provides an effective time-keeping strategy in prospective situations.

In sum, the approach being offered here is one that has

incorporated many of those ideas inherent in other models of time estimation behavior. At the same time, however, it has also extended these ideas by acknowledging the role of environmental structure in determining the way that event duration is encoded and later retrieved from memory. Together, these ideas about time estimation behavior and the role of the environment form a common framework that can address both the accuracy and bias of duration judgments under prospective and retrospective circumstances.

Although the variable of learning was not manipulated in the present set of experiments, it is important to acknowledge its role in the processing of event information. In the two experiments described here, participants were given sufficient learning experience (i.e., eight exposures to each melody) to ensure that they would have the opportunity to discern the underlying relations within events and their patterns of organization. This provision of learning experience is critical because at early stages of learning (one or two trials), the overall organization of a coherent event may not be readily apparent so that, in effect, it appears incoherently structured. Cognitively, this means that temporal and nontemporal information may be processed in an independent fashion. This prediction is one that needs to be empirically tested, and it could easily be done by essentially replicating the present set of studies and adding a manipulation of learning trials. In general, though, it is assumed that learning contributes to the degree of event coherence and that both are codeterminants of the type of encoding that occurs.

One issue that is worthwhile to consider is whether this overall conceptual framework generalizes to other types of naturalistic events. One might argue that music is in many ways unique because the lawful relationship between nontemporal (i.e., pitch) and temporal (i.e., rhythmic) information is intentionally composed as such for aesthetic purposes. This may be true, but it has also been demonstrated that the structure of music is remarkably similar to that of speech (Jackendoff, 1989; Jackendoff & Lerdahl, 1982; Martin, 1972, 1986) and of various types of motor behavior such as locomotion (Carlsoo, 1972; Pierson, 1976). Hence, one would expect the encoding of these events' temporal and nontemporal information to produce a pattern of results similar to the one observed here. In addition, although a rhythmic patterning of information may not be apparent in other domains, all events contain nontemporal information that unfolds with a given rate. If this dimension can also be incidentally encoded into memory with sufficient learning experience, then those underlying processes described here may in fact have ramifications for a wide range of natural events. This issue is one that needs to be addressed in future research.

Although the present findings have been discussed primarily in terms of their implications for time estimation behavior, they also have relevance for the area of music cognition. To some extent, they offer insight into the independence-interactive debate and the question of whether the influence of a melody's rhythm on perception and memory is independent of its accompanying pitch pattern. As shown here, the answer to this question, in part, depends

on the degree to which the two dimensions are structurally compatible or incompatible as well as on an individual's degree of musical experience. In Experiments 1 and 2, incoherent melodies displayed a lawful rhythm, but because their temporal accents failed to coincide with melodic phrase relations, both required independent processing resources. Thus, if one is listening to the music of a novice pianist, only the dimension that was selectively attended to should be encoded into memory. On the other hand, if accents coincide with melodic phrase boundaries, then a listener should pick up both dimensions regardless of which one was initially attended to. This may help to explain why one can often seem to remember the rhythm of many well-learned folk tunes (e.g., "Frere Jacques" or "Happy Birthday") even though selective attention was not necessarily directed toward this dimension at the time of learning. One would also expect a melody's rate or tempo to be incidentally learned, and there is some evidence that this is so. Halpern (1988), for example, asked participants to both produce and recognize the tempo of several well-known folk tunes and found their performance to be quite accurate. This issue in turn leads to a related question of whether one structural dimension is more easily encoded into the cognitive system than the other. That is, does the incidental learning of rhythm occur after a fewer number of trials than the incidental learning of pitch (or vice versa)? Although the present study cannot answer this question because of differences in the types of memory tasks used (recall vs. recognition), this issue could be investigated by essentially replicating the present experimental design with musicians and creating more comparable memory tasks. The results, in turn, could have implications for teaching students of music both to play a new instrument and to improve their listening abilities.

Last, at a more general level, this research also suggests a need to acknowledge the role of event structure and learning in cognitive research. As found in the present studies, these two factors may interact with one another to yield very different conclusions on the types of encoding and remembering strategies used in everyday behavior.

References

Berry, W. (1976). *Structural functions in music.* Englewood Cliffs, NJ: Prentice-Hall.

Block, R. A. (1985). Contextual coding in memory: Studies of remembered duration. In J. Michon & J. Jackson (Eds.), *Time, action, and behavior* (pp. 169–178). Heidelberg, Germany: Springer-Verlag.

Boltz, M. G. (1991). Some structural determinants of melody recall. *Memory & Cognition, 19,* 239–251.

Boltz, M. G. (1992). The remembering of auditory event durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 938–956.

Boltz, M. G. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception & Psychophysics, 53,* 585–600.

Boltz, M. G. (1994). Changes in internal tempo and effects on the learning and remembering of event durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 1154–1171.

Boltz, M. G. (1995). Effects of event structure on retrospective

duration judgments. *Perception & Psychophysics, 57,* 1080–1096.

Boltz, M. G., & Jones, M. R. (1986). Does rule recursion make melodies easier to reproduce? If not, what does? *Cognitive Psychology, 18,* 389–431.

Botsford, F. H. (Ed.). (1921). *Folk songs of many people* (Vol. 1). New York: Woman's Press.

Brown, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics, 38,* 115–124.

Brown, S. W., & Stubbs, D. A. (1992). Attention and interference in prospective and retrospective timing. *Perception, 21,* 545–557.

Carlsoo, S. (1972). *How man moves.* New York: Cran, Russak.

Cone, E. (1968). *Musical form and musical performance.* New York: Norton.

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech, 15,* 103–113.

Grosjean, F., Grosjean, L., & Lane, H. (1979). The patterns of silence: Performance structures in sentence production. *Cognitive Psychology, 11,* 58–81.

Halpern, A. R. (1988). Perceived and imagined tempos of familiar songs. *Music Perception, 6,* 193–202.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. (1976). Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology, 89,* 719–730.

Jackendoff, R. (1989). A comparison of rhythmic structures in music and language. In P. Kiparsky & G. Youmans (Eds.), *Phonetics and phonology: Vol. 1. Rhythm and meter* (pp. 15–44). New York: Academic Press.

Jackendoff, R., & Lerdahl, F. (1982). A grammatical parallel between music and language. In M. Clynes (Ed.), *Music, mind, and brain* (pp. 83–117). New York: Plenum Press.

Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review, 96,* 459–491.

Jones, M. R., Boltz, M. G., & Kidd, G. R. (1982). Controlled attending as a function of melodic and temporal context. *Perception & Psychophysics, 32,* 211–218.

Julesz, B., & Hirsh, I. J. (1972). Visual and auditory perception—An essay of comparison. In E. E. David & P. Denes (Eds.), *Human communication: A unified view* (pp. 283–349). New York: McGraw-Hill.

Kramer, J. D. (1982). Beginnings and endings in Western art music. *Canadian University Music Review, 3,* 1–14.

Kubovy, M. (1981). Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55–98). Hillsdale, NJ: Erlbaum.

Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Marks, L. E. (1978). *The unity of the senses.* New York: Academic Press.

Martin, J. G. (1972). Rhythmic (hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79,* 487–509.

Martin, J. G. (1986). Aspects of rhythmic structure in speech perception. In J. Evans & M. Clynes (Eds.), *Rhythm in psychological, linguistic, and musical processes* (pp. 79–98). Springfield, IL: Charles C Thomas.

McClain, L. (1983). Interval estimation: Effects of processing demands on prospective and retrospective reports. *Perception & Psychophysics, 34,* 185–189.

Michon, J. (1972). Processing of temporal information and the cognitive theory of time experience. In J. T. Fraser, F. C. Haber,

& G. W. Muller (Eds.), *The study of time* (pp. 224–258). New York: Springer-Verlag.

Miller, G. W., Hicks, R. E., & Willette, M. (1978). Effects of concurrent verbal rehearsal and temporal set upon judgments of temporal duration. *Acta Psychologica, 42*, 173–179.

Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception, 3*, 1–32.

Ornstein, R. E. (1969). *On the experience of time*. Baltimore, MD: Penguin.

Palmer, C., & Krumhansl, C. L. (1987). Independent temporal and pitch structures in determination of musical phrases. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 116–126.

Pierson, K. (1976). Control of walking. *Scientific American, 235*, 72–86.

Shields, J. L., McHugh, R., & Martin, J. G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology, 102*, 250–255.

Stetson, R. (1905). A motor theory of rhythm and discrete succession. *Psychological Review, 12*, 250–350.

Thomas, E. A. C., & Cantor, N. E. (1978). Interdependence between the processing of temporal and nontemporal information. In J. Raquin (Ed.), *Attention and performance VIII* (pp. 43–62). Hillsdale, NJ: Erlbaum.

Todd, N. (1985). A model of expressive timing in tonal music. *Music Perception, 3*, 33–58.

Todd, R., Boltz, M. G., & Jones, M. R. (1989). The Midilab auditory research system. *Psychomusicology, 8*, 83–96.

Zakay, D. (1989). Subjective time and attentional resource allocation: An integrated model of time estimation. In I. Levin & D. Zakay (Eds.), *Time and human cognition* (pp. 365–397). Amsterdam: North Holland.

Zakay, D. (1993). Relative and absolute duration judgments under prospective and retrospective paradigms. *Perception & Psychophysics, 54*, 656–664.