THE INTEGRATION OF PITCH AND TIME IN MUSIC PERCEPTION

by

Jonathan B. Prince

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy,
Graduate Department of Psychology,
in the University of Toronto

© Copyright by Jonathan B. Prince, 2009
THE INTEGRATION OF PITCH AND TIME IN MUSIC PERCEPTION

Jonathan B. Prince

Doctor of Philosophy, 2009

Department of Psychology, University of Toronto

ABSTRACT

Nine experiments were conducted to explore pitch-time integration in music. In Experiments 1-6, listeners heard a musical context followed by probe events that varied in pitch class and temporal position. When evaluating the goodness-of-fit of the probe (Experiment 1), listeners’ ratings showed more influence of pitch class than of temporal position. The tonal and metric hierarchies contributed additively to ratings. Listeners again rated goodness-of-fit in Experiment 2, but with instructions to ignore pitch. Temporal position dominated ratings, but an effect of pitch consistent with the tonal hierarchy remained. Again, these two factors contributed additively. A speeded classification task in Experiments 3 and 4 revealed asymmetric interference. When making a temporal judgment (Experiment 3), listeners exhibited a response bias consistent with the tonal hierarchy, but the metric hierarchy did not affect their pitch judgments (Experiment 4). Experiments 5 and 6 ruled out alternative explanations based on the presence of pitch classes and temporal positions in the context, unequal numbers of pitch classes and temporal positions in the probe events, and differential difficulty of pitch versus temporal classification. Experiments 7-9 examined the factors that modulate the effect of temporal variation on pitch judgments. In Experiment 7, a standard tone was
followed by a tonal context and then a comparison tone. Participants judged whether the comparison tone was in the key of the context or whether it was higher or lower than the standard tone. For both tasks, the comparison tone occurred early, on time, or late with respect to temporal expectancies established by the context. Temporal variation did not affect accuracy in either task. Experiment 8 used the pitch height comparison task, and had either a tonal or an atonal context. Temporal variation affected accuracy only for atonal contexts. Experiment 9 replicated these results and controlled for potential confounds. The findings imply that the tonal contexts found in typical Western music bias attention toward pitch, increasing the salience of this dimension at the expense of time. Pitch salience likely arises from long-term exposure to the statistical properties of Western music and is not linked to the relative discriminability of pitch and time.
ACKNOWLEDGMENTS

I am grateful for the instruction, guidance and support provided throughout my graduate studies by my supervisors, Professor Mark Schmuckler and Professor William Thompson. I appreciate the detailed help of Professor Sandra Trehub and Professor Bruce Schneider with my thesis and their generous devotion of time as committee members. Additionally, I credit Professor Lola Cuddy and Professor Eyal Reingold for their service as members of my examination committee. I thank my family for their support and encouragement.

Copyright © 2009 by the American Psychological Association. Reproduced with permission. The official citation that should be used in referencing this material is:


No further reproduction or distribution is permitted without written permission from the American Psychological Association.

Some of this material is published in:

# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TITLE PAGE</td>
<td>i</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Pitch and time in music perception</td>
<td>9</td>
</tr>
<tr>
<td>The combination of musical pitch and time</td>
<td>13</td>
</tr>
<tr>
<td>Reconciliation</td>
<td>18</td>
</tr>
<tr>
<td>EXPERIMENT 1: PITCH-TIME DIMENSIONAL RELATIONS IN PROBE EVENT RATINGS</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>23</td>
</tr>
<tr>
<td>Method</td>
<td>24</td>
</tr>
<tr>
<td>Results</td>
<td>27</td>
</tr>
<tr>
<td>Discussion</td>
<td>30</td>
</tr>
<tr>
<td>EXPERIMENT 2: PITCH-TIME DIMENSIONAL RELATIONS IN SELECTIVE ATTENTION TO TIME</td>
<td></td>
</tr>
<tr>
<td>Introduction</td>
<td>32</td>
</tr>
<tr>
<td>Method</td>
<td>32</td>
</tr>
<tr>
<td>Results</td>
<td>32</td>
</tr>
<tr>
<td>Discussion</td>
<td>35</td>
</tr>
</tbody>
</table>
## EXPERIMENT 3: PITCH-TIME DIMENSIONAL RELATIONS IN SPEEDED TEMPORAL CLASSIFICATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>37</td>
</tr>
<tr>
<td>Method</td>
<td>37</td>
</tr>
<tr>
<td>Results</td>
<td>39</td>
</tr>
<tr>
<td>Discussion</td>
<td>45</td>
</tr>
</tbody>
</table>

## EXPERIMENT 4: PITCH-TIME DIMENSIONAL RELATIONS IN SPEEDED PITCH CLASSIFICATION

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>48</td>
</tr>
<tr>
<td>Method</td>
<td>48</td>
</tr>
<tr>
<td>Results</td>
<td>49</td>
</tr>
<tr>
<td>Discussion</td>
<td>52</td>
</tr>
</tbody>
</table>

## EXPERIMENT 5: BASELINE PITCH-TIME DIMENSIONAL RELATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>57</td>
</tr>
<tr>
<td>Method</td>
<td>57</td>
</tr>
<tr>
<td>Results</td>
<td>59</td>
</tr>
<tr>
<td>Classification Tasks</td>
<td>59</td>
</tr>
<tr>
<td>Goodness-of-fit Ratings</td>
<td>60</td>
</tr>
<tr>
<td>Discussion</td>
<td>60</td>
</tr>
</tbody>
</table>

## EXPERIMENT 6: PITCH-TIME DIMENSIONAL RELATIONS IN EQUALIZED CLASSIFICATIONS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>62</td>
</tr>
</tbody>
</table>
EXPERIMENTS 7-9: PITCH-TIME DIMENSIONAL RELATIONS IN TEMPORAL EXPECTANCIES

EXPERIMENT 7: THE ROLE OF TASK IN PITCH-TIME DIMENSIONAL RELATIONS

EXPERIMENT 8: THE ROLE OF CONTEXT IN PITCH-TIME DIMENSIONAL RELATIONS: PART I

EXPERIMENT 9: THE ROLE OF CONTEXT IN PITCH-TIME DIMENSIONAL RELATIONS: PART II

Method 62
Results 63
Discussion 67

Method 77
Results 82
Discussion 84

Method 87
Results 90
Discussion 92

Method 95
Results 97
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>42</td>
</tr>
<tr>
<td>Table 2</td>
<td>43</td>
</tr>
<tr>
<td>Table 3</td>
<td>68</td>
</tr>
<tr>
<td>Table 4</td>
<td>80</td>
</tr>
<tr>
<td>Table 5</td>
<td>88</td>
</tr>
<tr>
<td>Table 6</td>
<td>89</td>
</tr>
</tbody>
</table>
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>12</td>
</tr>
<tr>
<td>Figure 2</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3</td>
<td>29</td>
</tr>
<tr>
<td>Figure 4</td>
<td>30</td>
</tr>
<tr>
<td>Figure 5</td>
<td>39</td>
</tr>
<tr>
<td>Figure 6</td>
<td>45</td>
</tr>
<tr>
<td>Figure 7</td>
<td>52</td>
</tr>
<tr>
<td>Figure 8</td>
<td>66</td>
</tr>
<tr>
<td>Figure 9</td>
<td>81</td>
</tr>
<tr>
<td>Figure 10</td>
<td>84</td>
</tr>
<tr>
<td>Figure 11</td>
<td>90</td>
</tr>
<tr>
<td>Figure 12</td>
<td>91</td>
</tr>
<tr>
<td>Figure 13</td>
<td>99</td>
</tr>
</tbody>
</table>
INTRODUCTION

A central goal of cognitive psychology is to understand the perception and processing of visual and auditory stimuli. Unfortunately, creating stable mental representations of such external stimuli is not a simple task, because it requires apprehending the structure of complex multidimensional visual and auditory events. These stimuli are defined not only by variations along one or more of the dimensions, but also the context in which they are experienced. Additionally, the processing of these stimuli must exhibit stability in the face of complex and often subtle variations that frequently occur along any number of physical dimensions simultaneously. For example, listeners can easily recognize a familiar melody even if its pitches are changed by presenting it in a different musical key and its timing varied by using a different tempo. However, more subtle variations in the frequencies or durations of the notes in the melody can dramatically alter its structure and musical context, consequently yielding a considerably different mental percept. Hence, understanding the perception of melodies requires knowledge of how the musical dimensions interact.

This thesis explores how the two primary dimensions of musical experience (pitch and time) contribute to the perception of musical events. There are several ways in which pitch and time could combine to yield the perception of a melody. For instance, the “stability” or “goodness” of a melody could reflect an additive and independent contribution of the values along the individual dimensions that define it. Conversely, variation along one dimension might change or be affected by the values along another dimension; in other words, they interact.
Similarly, Garner’s (1974) classic research on dimensional interactions in information processing classified dimensions as either separable or integral. Consider a simple bidimensional stimulus. If the two dimensions are integral, the perceiver is unable to separate the component dimensions, resulting in more holistic perception, which will be dependent on the combination of values along the individual dimensions. If a task requires selective attention to one dimension, variation in an irrelevant but integral dimension will affect performance. For perceptually separable dimensions, variation in the irrelevant dimension will not affect processing of the relevant dimension.

Integrality and separability of dimensions can be demonstrated in similarity evaluations between two bidimensional stimuli. If the psychological distance between the two stimuli equals the sum of the difference along each dimension, these dimensions are separable and additive. If, however, the psychological distance between the stimuli does not equal (i.e., is less or greater than) the total difference along each dimension, then the dimensions are integral. In other words, perceivers can clearly notice and/or process changes in one dimension independent of changes in the other for separable but not for integral dimensions.

Using similarity ratings to create a psychological map of distances between percepts is one technique for delineating separable from integral dimensions. Another method is to use speeded classification tasks, in which objective measures of accuracy and reaction time (RT) can reveal the degree of integrality or separability of the dimensions. In this case, the perceiver sorts bidimensional stimuli according to one dimension while the other, irrelevant, dimension is varied not at all (baseline), or is varied in a correlated manner, or is varied orthogonally (filtering task). If the two
dimensions are separable, then there should be no difference in performance regardless of 
what is happening in the irrelevant dimension. For integral dimensions, however, 
correlated variation of the two dimensions will improve performance, and orthogonal 
variation will degrade performance. Thus the presence of redundancy gain (correlated) 
and loss (orthogonal) is another sign of dimensional integrality. This approach to 
determining whether dimensions are separable or integral was pioneered by Garner 
(1974).

Garner and Felfoldy (1970) used a speeded classification task to test the 
integrality versus separability of several dimensions, including color brightness and 
saturation, horizontal and vertical dot position, as well as circle size and diameter angle. 
The participants were shown stimuli containing paired values along the two dimensions 
and were required to classify the stimuli along one of these dimensions as rapidly as 
possible. The irrelevant dimension in each pair did vary not at all, varied in a correlated 
manner, or varied in a manner that was opposite (orthogonal) to the variation in the 
relevant dimension. For integral dimensions, such as brightness and saturation, correlated 
variation resulted in the fastest classification times, and orthogonal variation resulted in 
the slowest classification times. The invariant, baseline condition was intermediate, 
demonstrating that the correlated and orthogonal variation caused redundancy gain and 
loss, respectively.

This question of dimensional separability versus integrality has been explored in 
several domains, including vision (Garner & Felfoldy, 1970; Handel & Imai, 1972), 
audition (Ben-Artzi & Marks, 1999; Krumhansl & Iverson, 1992; Melara & Marks, 
1990a, 1990b, 1990c; Tekman, 2002) and cross-modal perception (Amazeen & Turvey,
Garner and Felfoldy (1970) reported that the horizontal and vertical positions of a dot were integral, but circle size and diameter angle were separable. Handel and Imai (1972) used similarity ratings to show that brightness and saturation are integral dimensions, whereas lightness and object size are separable.

Garner interference is also relevant to auditory perception. Specifically, concepts expressed through spoken words and several other dimensions, including pitch (Melara & Marks, 1990c), spatial position (Ben-Artzi & Marks, 1999) and loudness (Melara & Marks, 1990a) commonly exhibit interference across dimensions. Melara and Marks (1990b) also reported dimensional interactions between any bidimensional combination of pitch, timbre or loudness. Interestingly, Krumhansl and Iverson (1992) found that pitch and timbre had interactive effects on speeded classification for isolated stimuli but not for stimuli presented in a musical context. Moreover, Tekman (2002) found mutual inference in the detection of loudness and timing changes in tone sequences.

Cross-modal studies have revealed Garner interference between pitch and color (Melara, 1989), pitch and brightness (Marks, 1987), pitch and form (Marks, 1987) and pitch and spatial position (Ben-Artzi & Marks, 1995; Melara & O'Brien, 1987). The size-weight illusion is another example of interactions between cross-modal dimensions (Amazeen & Turvey, 1996). In this illusion, an object that is larger than another object of identical weight is judged to be heavier than the smaller object.

However, the strict categorization of dimensions as either integral or separable does not always fit the data. Asymmetric integrality (or separability) occurs when one dimension interferes with another, but not vice versa (Garner, 1976). There are several
explanations for such situations – differences in discriminability, level of processing, or physical primacy of the dimensions. Differences in discriminability appear when the processing of dimension A alone is easier than that of dimension B alone. Thus, when classifying dimension B, there is interference from the irrelevant dimension A due to its lesser processing requirements (Garner & Felfoldy, 1970). The ease of processing dimension A means that it can occur automatically and/or faster than dimension B; the availability of this information then interferes with the more effortful and/or slower processing of dimension B. Accordingly, task demands can change how integral or separable dimensions appear to be. Asymmetric interference can also occur if extraction and processing of one dimension occurs at an earlier stage than another, as the earlier-processed dimension may interfere with the dimension not yet analyzed (Garner, 1974). Finally, when one dimension cannot exist without another (thus exerting physical primacy), asymmetric integrality often emerges, in favor of the physically primary dimension. For example, when classifying a speech syllable, in order to be an auditory signal, it must have a pitch. Thus pitch can asymmetrically interfere with syllable identification because it exhibits physical primacy (Wood, 1974).

There are reports of degrees of integrality – dimensions that are neither integral nor separable, but somewhere in between (L. B. Smith & Kemler, 1978) because of age-related changes in the ability to separate a stimulus into its component dimensions (L. B. Smith & Kemler, 1977). Yet even in adulthood, individual differences can emerge in the pattern of observed integrality or separability of dimensions (J. D. Smith & Baron, 1981). Moreover, in some cases, performance strategies may result in differing patterns of dimensional processing (Pomerantz, 1983).
Another approach, based on signal detection theory (Green & Swets, 1966), holds that dimensional interactions occur at two different levels, perceptual and decisional (Ashby & Townsend, 1986). The perceptual process refers to the phenomenal experience of the perceiver when exposed to a given stimulus, whereas the decisional process is the rule used to select a response based on the stimulus percept. Perceptual separability occurs when one dimension fails to elicit perceptual effects on another at early processing levels. Accordingly, decisional separability holds when an irrelevant dimension does not affect the decisional rule used by participants, occurring at later processing levels. Garner interference can happen as a result of failure of perceptual or decisional separability. Furthermore, failure of either type of separability does not imply failure of the other. This distinction between perceptual and decisional separability has been successfully applied in work on facial recognition (Thomas, 2001) and on the dissociation of identification and categorization processes (Maddox, 2001; Maddox & Dodd, 2003).

Taken together, such findings suggest that the classification of dimensions as purely integral or purely separable does not adequately capture the complex nature of processes for combining stimulus dimensions. Potts, Melara, and Marks (1998) question the utility of establishing a list of integral and separable dimensions along with numerous qualifications on the applicability of the rules. Of greater importance, perhaps, are the circumstances that do and do not give rise to joint processing of dimensions.

The physical dimensions of a stimulus may vary along a single unidimensional continuum, such as the wavelength of light reflected by an object. However this physical variation is often not perceived as continuous, and is instead encoded as variation along a psychological dimension that may have several component dimensions and complex
internal structure. For example, the psychological experience of color has at least three sub-dimensions (hue, saturation, and brightness), even though the external stimulus consists of variation along a single physical dimension (wavelength). Thus it is important to distinguish between physical dimensions and their corresponding psychological dimensions, because the structure and interrelations of the dimensions can accordingly be considerably different. This thesis deals mainly with the psychological (i.e., perceived) structure of the physical dimensions of pitch and time, and especially how interrelations between these structures affect the perception of musical events.

Distinct values along the psychological dimensions of an object may function as perceptual reference points, or prototypes, that aid in the formation of categories. In the case of color, these prototypes are referred to as focal colors, such as a “good” red, compared to an “off” red (Rosch, 1973). Encoding information relative to these reference points provides an organizational framework for classifying objects into categories. Indeed, humans naturally sort objects into semantic categories at optimal levels of abstraction chosen to maximize the informative value of the categories (Rosch, 1973, 1975; Rosch, Mervis, Gray, Johnson, & Boyesbraem, 1976). In the language domain, the most prototypical words in a category are the most informative reference points and are maximally different from other prototypes (Rosch et al., 1976). An object that exemplifies a category functions as the central prototype around which the category forms. Membership is then determined on the basis of defining features (necessary for membership) for the category or as a function of the similarity of the objects along their component dimensions (Tversky, 1977). Furthermore, an exemplar within the category is not defined by possession of a particular set of characteristic features (optional, but
preferred for membership), but as a better or worse example of a category according to its similarity to the prototype along a number of dimensions (Rosch & Mervis, 1975).

One unaddressed issue is whether the internal structure of psychological dimensions affects how they combine in perception. How does the structure of these component dimensions combine to form a percept? Are there conditions under which the internal structure of one dimension may affect that of another? If so, then a small change along one of an object’s physical dimensions may have larger perceptual consequences not limited to its corresponding psychological dimension alone. That is, the seemingly small physical change might alter the perceived internal structure of its psychological dimension. Furthermore, a structural variation could then affect the perceived structure of another physical dimension. In such a manner, a shift along one physical dimension of an object may cause a cascade of perceived changes along psychological dimensions, ultimately resulting in a much different percept. Thus understanding how the internal structure of the psychological dimensions of an object integrate and contribute to perception is critical to understanding how the physical and psychological attributes of perceptual objects are extracted and encoded. In this discussion, perceptual “objects” are not limited to those that are static and three-dimensional; they also include auditory objects and events such as words or musical notes.

There are several reasons why music is an ideal arena for investigating questions about the nature of dimensional interactions and the internal organization of stimuli in memory. First, it is a psychologically rich stimulus that provides several dimensions such as pitch, time, timbre, and loudness that may or may not interact. Unlike static objects that provide their own self-contained context, musical events (or auditory objects) unfold
over time. Thus the dimensions of musical events often require a preceding context to establish their internal structure.

Second, there are music theoretic models of stimulus structure that translate into general and quantitative psychological theories. In particular, musical pitch perception illustrates classic themes of categorization and stimulus prototypes (Krumhansl, 1990); research on musical time is relevant to the same issues (Jones, 1976). Third, within a musical context the primary dimensions of pitch and time also display a quantifiable internal hierarchical organization. Musical pitches and temporal positions are organized on the basis of their hierarchical position when a preceding context establishes this hierarchy, in accordance with the culturally learned internal structures of these dimensions (i.e., tonality and meter). As a result, there can be quantitative descriptions of the psychological structure of pitch and time in music, which make it possible to investigate how these dimensions combine in a musical context. In other words, questions about how and under what circumstances the organization of pitch structures affect temporal structures (and vice versa) can be addressed.

**Pitch and time in music perception**

Pitch and time are the primary dimensions of music. Accordingly, much research in music cognition has focused on these dimensions. There are rich theoretical models of the encoding of pitch and time within the hierarchical structures of tonality and meter.

Notable works on pitch perception abound, led, perhaps, by Helmholtz’s seminal volume (Helmholtz, 1875/1954). More relevant for the current research, however, are the studies of Carol Krumhansl and her colleagues on the perception of musical key or tonality (Krumhansl, 1990; Krumhansl & Kessler, 1982). When listeners perceive a
hierarchical ordering of stability of the twelve possible pitch classes in music (collapsing across octaves), the resultant perceptual structure is referred to as the musical key (Krumhansl, 1990). Krumhansl and Kessler (1982) used a probe tone technique that involves presenting listeners with a key-defining musical excerpt followed by a single tone, or probe. Listeners rate the “goodness-of-fit” of each of the 12 pitch classes in the chromatic scale. Examination of these ratings suggests three hierarchical levels of tonal stability (as evidenced by goodness-of-fit): tones within the tonic triad, those within the key, and those outside the key (see Figure 1 for a graphical representation of the tonal hierarchy). These levels correspond to category membership of exemplars – the within-tonic triad members are the most prototypical pitches of the key. Additionally, the remaining pitches fall into categories in accordance with their level of similarity to the prototypical pitches. The pattern of ratings for all of the pitch classes correlates near-perfectly with the frequency of occurrence of those tones in real Western music and in music-theoretic accounts, suggesting, perhaps, that the ratings merely reflect their frequency of occurrence in the musical context preceding the probe tone. However, randomizing the assignment of pitch classes to relative frequencies of occurrence fails to elicit any hierarchical perception of pitch (N. A. Smith & Schmuckler, 2004). In other words, if the preceding context uses a hierarchy based on frequency of occurrence that is anything other than the standard tonal hierarchy, listeners perceive no hierarchy at all. However, Oram and Cuddy (1995) did observe sensitivity to an artificial hierarchy induced by frequency of occurrence of pitch classes. Nevertheless, the frequency of occurrence of tones in the preceding musical context cannot solely account for the observed pattern of ratings in a typical Western musical context. Instead, listeners must
also be accessing a stored representation of the statistical properties of pitch as used in Western music (the tonal hierarchy) and bringing this experience to bear on their perception of the immediate context.
Figure 1. Tonal and metric hierarchies (Krumhansl & Kessler, 1982; Palmer & Krumhansl, 1990).
Similarly, there is a large body of research on the role of temporal patterns in music perception, resulting in perceptual accounts of time and meter (Jones, 1976; Palmer & Krumhansl, 1990). Sequences of durations related to each other by simple ratios (i.e., 1:2 or 1:3) often characterize temporal patterns in music. More complex ratios (such as 4:5) tend to become assimilated both in perception and production to simpler ratios (Povel, 1981). These durational ratios contribute to the perception of accented (strong) and unaccented (weak) points in time. This oscillation defines musical meter – a regular pattern of alternation between weak and strong temporal positions that serves as an organizational basis for musical information. Lerdahl and Jackendoff (1983) formalized a theory of meter based on nested hierarchies and well-formedness rules, derived from linguistic theory. Upon establishing a meter, or pattern of relative accents, the psychological stability of temporal positions varies according to their location in the metric hierarchy (Palmer & Krumhansl, 1990). Using a discrimination task and the probe tone technique with temporal patterns, Palmer and Krumhansl (1990) showed the psychological reality of these principles of musical meter. Subjective ratings and similarity judgments of musical rhythms demonstrate three components of the psychological experience of rhythm and meter: a cognitive/structural component that denotes form and salience, a motoric/motion component that represents physical components of motion, and affective/emotional component that reflects subjective feelings in response to the meter (Gabrielsson, 1973a, 1973b, 1973c).

\textbf{The combination of musical pitch and time}
The question of possible joint contributions of pitch and time to the perception of musical events has a long history. For the most part, there have been two opposing viewpoints: one that considers pitch and time as independent experiential dimensions; whereas the other argues that they are highly interactive dimensions. Numerous researchers have failed to find significant interaction effects of pitch and time in music perception with a variety of tasks, including judgments of melody completion, pleasantness and similarity (Makris & Mullet, 2003; Monahan & Carterette, 1985; Palmer & Krumhansl, 1987a, 1987b; Pitt & Monahan, 1987) as well as recall, pitch change detection, matching, well-formedness ratings, and modulation detection (Krumhansl, 1991; K. C. Smith & Cuddy, 1989; Thompson, 1993, 1994; Thompson, Hall, & Pressing, 2001). Hence these studies suggest that pitch and meter combine additively in musical experience.

Consider, for example, the work of Palmer and Krumhansl (1987a, 1987b). These investigators manipulated the form of a melody by holding the pitch constant in one condition, time in another, and using the original melody in another condition. A second experiment disrupted the phase of the pitches and durations, and further experiments used durations from a recording that had subtle variations in expressive timing. Musicians of varying skill level judged how complete the melody sounded. In all cases, both tonal and metrical hierarchies affected ratings in an additive and linear fashion. Similarly, Makris and Mullet (2003) found that pleasantness ratings of melodies were predicted by linear combinations of pitch, contour, time and timbre. When evaluating similarity of melodies combined with different polyrhythms (multiple temporal patterns occurring simultaneously), Pitt and Monahan (1987) proposed that pitch and temporal information
were processed separately, and perhaps combined at a late stage in perception. Another study reported a tradeoff in attention to pitch and temporal structures, suggesting a fixed capacity of attentional resources that can be directed to either dimension at will (Monahan & Carterette, 1985).

Testing recognition memory for musical excerpts, Krumhansl (1991) found that recombining different pitch and time patterns from an atonal piece (i.e., whose distribution of pitch classes did not conform to the tonal hierarchy) did not affect recognition ratings, suggesting that they were not encoded jointly in memory. A melodic dictation task requires the listener to notate a heard melody in written musical format. Detection of pitch changes in this task is more accurate in meters heard more often in typical music, but is not aided by the specific combination of pitches and durations (K. C. Smith & Cuddy, 1989). Other work on memory for tones and detection of changes in tone sequences found that listeners were sensitive to the particular combinations of pitch and duration only when actively attending (Thompson, 1994; Thompson et al., 2001).

Neuroimaging data and clinical evidence of double dissociations support the idea that separate cortical areas process pitch and duration (Fries & Swihart, 1990; Mavlov, 1980; Peretz, 1990, 1996; Peretz & Kolinsky, 1993; Peretz et al., 1994; Schön & Besson, 2002). In a music-reading task, musicians could selectively attend to either dimension with no interference from the irrelevant dimension, as measured by both behavioral results and ERP components (Schön & Besson, 2002). Peretz (1990) investigated melodic processing in four patients with unilateral damage to the temporal lobe. Half of the patients had damage on the right hemisphere, whereas the other half had lesions in the left hemisphere. Interestingly, the patients with right hemisphere damage exhibited
selective pitch processing deficits with preserved temporal processing whereas the left hemisphere patients exhibited the opposite pattern. Thus Peretz (1990) observed a double dissociation between pitch and temporal processing, suggesting that anatomically and functionally independent systems underlie the two dimensions.

On the other side of the debate, there is a large body of research supporting the concept of joint accent structure and dynamic attending (Jones, 1987; Jones & Boltz, 1989). Joint accent structure refers to the weighting of pitch and time accents to form a combined pattern of overall salience of each note in a musical passage. The hierarchical organization of these dimensions contributes to their respective salience, but pitch contour and intervals (between successive notes) are regarded as the prime components of pitch accent. Tempo, rhythm and meter are the corresponding mechanisms for time. According to Jones (1987), regular temporal accents provide a perceptual anchor upon which all other musical information is encoded, even when listeners are instructed to ignore temporal structure. Dynamic attending posits that a regular meter provided by a preceding sequence causes an oscillation in attentional energy that is synchronized with the rhythm of the sequence (Barnes & Jones, 2000; Jones, Johnston, & Puente, 2006; Jones, Moynihan, MacKenzie, & Puente, 2002; Large & Jones, 1999). More recent work expanded this theory to pitch/time entrainment in which listeners track successive events in a sequence, extracting the pattern structure in both pitch and time (Jones et al., 2006). Listeners extrapolate the pattern structure to form dynamic expectancies in both dimensions, anticipating where (in pitch space) and when (in temporal space) subsequent events will occur. This dynamic attending model is built upon the notion that the pitch and temporal patterning of a sequence jointly drive listeners’ attention to future events. In
other words, pitch and time are presumed to have interactive relations in the sense that they jointly determine expectancies with respect to both the pitch of, and the timing of events when operating in a musical context.

Kidd, Boltz, and Jones (1984) found that temporal context and uncertainty affected the ability to discriminate between two versions of a melody. Other research demonstrated that the dynamic shape of a melody, defined by the combination of durations and contour, determines the likelihood of confusing a target melody with a lure (Jones & Ralston, 1991; Jones, Summerell, & Marshburn, 1987). Hébert and Peretz (1997) report that melody is more important than temporal structure for recognition of familiar music, but that the combination of the two is best and that this integration happens late in perceptual processing.

In contrast to Palmer and Krumhansl (1987a, 1987b), Boltz (1989b) found joint effects of pitch and time in judgments of melody completion. Specifically, the decrease in completion ratings as a result of concomitant violations in temporal and melodic structure was greater than the sum of the decreases caused by violating each dimension separately. This finding points to joint or holistic contributions of pitch and time. Another study found that completion ratings were highest when tonal and rhythmic accents coincided (Boltz, 1989a). On the other hand, this interaction was only present when the tritone interval was present in the melody, reinforcing its tonality. When judging how well a chord fits with a preceding melody, both harmonic and rhythmic expectancies play a role (Boltz, 1993a; Schmuckler & Boltz, 1994).

Temporal predictability provided by a coherent temporal structure aids the detection of pitch changes in pitch sequences (Boltz, 1993a; Jones, Boltz, & Kidd, 1982;
Jones et al., 2006; Jones et al., 2002; Kelley & Brandt, 1984; Monahan, Kendall, & Carterette, 1987). Likewise, duration judgments and time estimation are more accurate when presented in the context of an organized pitch structure (Boltz, 1989c, 1993b, 1995; Crowder & Neath, 1995). Also, the number of contour reversals and the size of pitch intervals affect the perceived rate (tempo) of a melody (Boltz, 1998b).

Pitch notation tasks (instructing participants to ignore time) show reliable effects of temporal structure. Memory for contour and tonal information improves when accompanied by supporting temporal patterns, and also at phrase boundaries defined by the conjunction of both dimensions (Boltz, 1991; Deutsch, 1980). More recent work found that although temporal shifts had large effects on melodic tasks, the reverse was not true (Abe & Okada, 2004). Infants are more likely to detect changes to patterns of pitch and temporal combinations when they are presented within a metrical framework than within a non-metrical framework (Hannon & Johnson, 2005).

Some neuroimaging research points to a joint contribution of pitch and time, at least in musically untrained subjects. ERP data show that temporal disruptions interfere with pitch judgments (Nittoono, Bito, Hayashi, Sakata, & Hori, 2000). A PET study posited a distributed neural network for processing both dimensions (Griffiths, Johnsrude, Dean, & Green, 1999). Finally, Schellenberg, Krysciak, and Campbell (2000) report combined effects of pitch and time when judging perceived emotion in melodies, although pitch tends to dominate.

**Reconciliation**

There have been several attempts to reconcile these divergent findings on the contributions of pitch and time to the perception of music. One suggestion is that they
combine relatively late in processing, such that tasks using lower-level “early” processing preclude joint contributions of pitch and time (Hébert & Peretz, 1997; Peretz & Kolinsky, 1993; Pitt & Monahan, 1987; Thompson et al., 2001; Tillmann & Lebrun-Guillaud, 2006). Accordingly, higher-level tasks might allow top-down influences, such that joint expectations about pitch and time become relevant. If so, then pitch height comparison tasks that require no knowledge of musical structure (implicit or explicit) and, therefore, do not involve top-down processes should prevent temporal influences on the task. However, Jones et al. (2002) report the opposite when listeners compare the pitch height of two tones separated by randomly-pitched intervening tones.

Another explanation is that local judgments (i.e., focusing on a single and/or isolated event) foster the ability to focus selectively on one dimension, whereas global judgments (that require integration of material over longer time periods) make selective attention difficult (Tillmann & Lebrun-Guillaud, 2006). The local/global distinction differs from the notion of top-down/bottom-up processing in that both local and global judgments may involve top-down and/or bottom-up processes. As such, these explanations do not necessarily refer to the same type of processing, although depending on the task requirements; there may be overlap between these theories. The global/local idea aligns with work on melodic completion ratings that finds joint influences of the two dimensions (Boltz, 1989b). Yet recall that other melodic completion judgments failed to find this pattern (Palmer & Krumhansl, 1987a, 1987b) and, furthermore, that melodic similarity judgments exhibited an attentional tradeoff between pitch and temporal structures (Monahan & Carterette, 1985).
Schellenberg et al. (2000) suggested that “subjective tasks” involving ratings (i.e., no defined correct or incorrect answer) elicit joint contributions of pitch and time, whereas “objective” tasks (in which accuracy measures are possible) result in the opposite pattern. This reasoning was developed with regard to emotional judgments of a melody, but it may overstate the distinction of subjective and objective aspects of tasks.

Boltz (1998a, 1999) offers another explanation based on the inherent structure of the melody and the extent of participants’ experience with the stimuli. It is easy to encode pitch and temporal structure when they are coherent in a melody, thus allowing a joint encoding in a dynamic shape. Conversely, a melody with little musical coherence requires more effortful processing because there are no ordered relations between the dimensions that enable a joint encoding. As a result, selective attention becomes easier. However, there have been joint contributions of the dimensions when the tonal coherence of a melody was systematically varied (Deutsch, 1980). The other component of Boltz’s theory concerns learning, more specifically that the perceived coherence of a melody changes as the participant gains more experience with it (Boltz, 1999). Because most studies use a variety of melodies and do not explicitly test differences across presentations, it is difficult to know whether this explanation can account for the contradictory findings.

The role of expertise in music perception is an important theme, and has also been cited as a possible source of differences in how pitch and time contribute to music perception. A number of the previously mentioned studies included this issue as part of their experimental design, including those that support dimensional independence (Makris & Mullet, 2003; Palmer & Krumhansl, 1987a, 1987b; Pitt & Monahan, 1987; K.
C. Smith & Cuddy, 1989) and those that favor joint contributions (Boltz, 1989a; Hébert & Peretz, 1997; Lebrun-Guillaud & Tillmann, 2007; Tillmann & Lebrun-Guillaud, 2006). However, of the few studies that found a main effect of musical expertise, none reported interactions between musical expertise and the pattern of integration of pitch and time.

From this review, it should be clear that in some situations and for some tasks pitch and time combine additively, whereas in other situations they do not. Rather than arguing for one position versus the other on this issue, a more reasonable approach is to determine the factors and circumstances that affect whether pitch and time combine jointly or fail to do so. This approach can result in deeper insights on the perception of music. Moreover, understanding the effects of interrelations between the psychological structures of a stimulus’ physical dimensions on perception can facilitate understanding of more general issues such as object perception.

Unquestionably, pitch and temporal structures are correlated in Western music, such that stable pitches (i.e., those occupying higher positions in the tonal hierarchy) occur at stable points in time (higher positions in the metric hierarchy). Passive exposure to this connection between pitch and temporal structures in Western music likely creates expectations for these combinations. As a result, violations of this principle could impair judgments along either dimension. Indeed, Lebrun-Guillaud and Tillmann (2007) found that the accuracy of detecting a temporal shift improved when the tonal function of the note was strong. Yet specifying the circumstances that elicit such effects is still unclear.

The aim of the present research was to explore how pitch and time combine in the perception of musical events. The first six experiments used subjective ratings of goodness-of-fit as well as objective classifications of pitch-time events. These
experiments add a new component to the issue of pitch-time integration in music by testing how the organizational principles of these two dimensions (tonality and meter) contribute to this phenomenon. In addition, examining the contribution of the tonal and metric hierarchies provides an opportunity to explore dimensional relations within the context of complex hierarchically structured categories, an unexplored issue in research on dimensional organization. Within a musical context, pitches and temporal positions correspond to entries/locations/etc. within tonal and metric hierarchies and, therefore, assume particular levels of stability within their respective dimension. Therefore, the combination of values along these two dimensions (pitch-time events) may influence how they combine perceptually. For example, the tonal stability of a pitch may be affected by the metric stability of its temporal position. Moreover, the overall stability of a musical event may be jointly defined by the specific combination of the stability provided by both tonality and meter.

The last three experiments explored how the experimental task and the presence of pitch structure in the stimulus context affect pitch-time relations in music perception. Specifically, these experiments tested how varying task and context characteristics modulate the effect of violating temporal expectancies on pitch judgments.
EXPERIMENT 1: PITCH-TIME DIMENSIONAL RELATIONS IN PROBE EVENT RATINGS

The goal of Experiment 1 was to examine perceptual relations between pitch and time as well as their structural organization of tonal and metric hierarchies in the perception of complex musical passages. Much previous research involved goodness-of-fit ratings to assess the psychological stability of musical events. Such “probe tone” ratings have been examined for tones that vary in pitch (see Krumhansl, 1990, for a review) and temporal position (e.g., Palmer & Krumhansl, 1990), but there are no studies that included systematic and simultaneous variations of the tonal (pitch) and metric (temporal) stability of musical events. The combinations of values along these two dimensions (pitch-time events) may influence how they combine perceptually. For example, the metric stability of a note’s temporal position may affect its tonal stability. Moreover, the specific combination of the stability provided by both tonality and meter may jointly define the overall stability of a musical event. Experiment 1 used pitch-time probe events following a musical context to investigate how pitch and time contribute to judgments of events in a musical context, and to determine the role of the structural organization of these dimension (i.e., tonality and meter).

For all experiments in this thesis, participants were musically trained so as to ensure that they could complete the tasks, some of which require explicit knowledge of musical structure. However, explicit musical training does not imply that the participants perceive musical events in a qualitatively different way than listeners without such training. Indeed, implicit tasks of music perception reveal that listeners perceive musical events quite similarly regardless of the amount of explicit training (Bigand & Poulin-
Moreover, studies of pitch-time relations have not found qualitative differences in patterns of pitch-time integration between musicians and nonmusicians (Boltz, 1989a; Hébert & Peretz, 1997; Lebrun-Guillaud & Tillmann, 2007; Makris & Mullet, 2003; Palmer & Krumhansl, 1987a, 1987b; Pitt & Monahan, 1987; K. C. Smith & Cuddy, 1989; Tillmann & Lebrun-Guillaud, 2006). Thus all listeners who are encultured in the Western musical system (i.e., grew up listening primarily to Western tonal music) engage in musical behavior by virtue of their passive exposure to music, and therefore are musical despite their lack of explicit training.

Method

Participants

The participants were 19 adults (M = 22.1 years, SD = 3.9), who had at least 8 years of formal musical training (M = 12.8, SD = 4). Participants received credit in an introductory psychology class at the University of Toronto Mississauga or monetary compensation ($10). Recruitment occurred through an online experiment database and flyers posted on campus.

Stimuli

A tonal melody with conventional harmonic accompaniment (I – IV – V – I) in the key of C major was composed following the stylistic rules of Western tonal music (Aldwell & Schacter, 2002). Figure 2 shows this passage, followed by a sample probe event. For this musical context there were 4 beats per measure, with a tempo of 120 beats per minute, producing a single beat duration (i.e., a quarter note in musical terminology) of 500 ms. All stimuli were generated with a PC computer and Finale 2005 software and
featured a harmonically complex piano timbre. The stimuli were exported to .wav files with a sampling frequency of 44.1 KHz. The amplitude of the passage was set at a comfortable listening level (approximately 60 dBA), with all pitches of equal amplitude. The duration of the passage was 4 s. Simultaneous with the musical passage was a metronome click that occurred on every beat (i.e., 500 ms intervals). The metronome click continued for one complete measure (2 s) after the passage finished to maintain a constant metric framework while also providing a pause between the melody and probe event.

![Piano](image)

![Click](image)

*Figure 2.* Musical context and sample probe event rated for goodness-of-fit in Experiments 1 and 2.

After the measure of metronome clicks, listeners heard a 250 ms pitch-time probe event. These probe events consisted of one of the 12 pitch classes of the chromatic scale ranging from middle C (262 Hz) to the B above middle C (494 Hz), at one of 8 possible temporal positions in the following measure. The eight temporal positions corresponded to an eighth-note subdivision of the metric hierarchy established by the passage (and its corresponding metronome click), which means that these probe events could occur at any
250 ms interval within the final measure. Therefore, time between the final metronome click and the probe event could be as short as 500 ms or as long as 2250 ms. Figure 2 also presents a sample probe event. Crossing the 12 different pitch classes with the 8 possible temporal positions produced a total of 96 possible probe events.

Apparatus

Participants were seated in a double-wall sound-attenuating booth (IAC) facing a Macintosh G4 computer running OSX 10.3, with a Viewsonic VG175 monitor. The Experiment Creator software package (http://www.ccit.utoronto.ca/billt/BillThompson_files/experiment.html) controlled presentation of the stimuli. Participants heard the stimuli through Sennheiser HD 280 Pro headphones and responded by means of the computer keyboard.

Procedure

All participants were tested individually and completed a questionnaire about their musical experience prior to the experiment. On each trial, listeners heard a musical passage followed by the probe event. After the probe event, the computer prompted listeners to rate how well the probe fit with the preceding context on a Likert scale of 1 (fits poorly) to 7 (fits well). There were no explicit instructions to attend to or ignore either the pitch or temporal aspects of the probe event. Participants completed 7-10 practice trials (depending on participants’ level of comfort with the task) to familiarize them with the nature of the stimuli and task. Practice trials were randomly selected from the 96 experimental stimuli and no feedback was provided. Trials were self-paced. Altogether, listeners completed three blocks of 96 trials, for a total of 288 trials. The entire procedure took about one hour.
Results

To check the consistency of the data, inter-block correlations (mean $r(94) = .62$, $SD = .18$) and inter-subject correlations (mean $r(94) = .59$, $SD = .06$) were examined. For trials on which the probe event occurred on the downbeat of the test measure, correlations with the standard tonal hierarchy values reported by Krumhansl and Kessler (1982) were computed (mean $r(10) = .80$, $SD = .13$). If all three of these correlations were more than two standard deviations below the group mean for any participant, that participant’s data were excluded from further analyses. Two listeners met these exclusion criteria. Analyses that included their data produced the same pattern of results.

Ratings were analyzed using a three-way repeated measures Analysis of Variance (ANOVA), with the within-subjects factors of pitch class (C, C#, ..., B), temporal position (position 1 through 8) and block (1, 2 and 3). There were significant main effects of pitch class, $F(11,176) = 55.85$, $MSE = 13.4$, $p < .001$, $\eta^2_p = .78$, and temporal position, $F(7,112) = 4.43$, $MSE = 3.32$, $p < .001$, $\eta^2_p = .22$, but not of block, $F(2,32) < 1$, $MSE = 1.2$. Importantly, the interaction between pitch class and temporal position was not significant, $F(77,1232) = 1.17$, $MSE = 1.06$, $p = .15$, $\eta^2_p = .07$. Neither the remaining two-way interactions, nor the three-way interaction, were significant.

The preceding analysis demonstrates that listeners’ ratings varied systematically as a function of the pitch and temporal context of the probe event, and that these effects were independent. To explore the relative contributions of the structure of these dimensions (i.e., tonality and meter) on judgments, the standard tonal hierarchy values

---

1 Correlations with the metric hierarchy were much more variable (mean $r(6) = .17$, $SD = .5$), negating their value as an exclusion criterion.
reported by Krumhansl and Kessler (1982) and the metric hierarchy values reported by Palmer and Krumhansl (1990) were used in a regression to predict average ratings. In the first step of this regression, the individual hierarchies predicted averaged probe event ratings, as separate variables. The second step in the analysis employed an interactive term, created by multiplying the tonal and metric hierarchies. For completeness, there were three other interactive terms as well (a division, maximum, and minimum model). Step one produced a multiple $R$ of .95 ($r^2 = .89$), with both tonal and metric hierarchy factors contributing significantly ($\beta = .94$ and .09, respectively; $p < .001$ for both).

Interestingly, of the 89% of the variance accounted for by these two factors, the tonal hierarchy uniquely explained all but 1%. In contrast, none of the interactive terms included in step two added any explanatory variance ($\Delta r^2 = 0$ for all). This analysis reveals that both the tonal and metric hierarchies influenced listeners’ ratings of the pitch-time probe events, with the tonal hierarchy accounting for most of the variance in mean ratings.

Figure 3 shows the strength of the pitch effect by displaying the ratings for the 12 probe pitch class of the chromatic scale, averaged across temporal position. The pattern of ratings associated with the standard tonal hierarchy reported by Krumhansl and Kessler (1982) is clearly discernible. In comparison, Figure 4 shows the ratings for the eight temporal positions, averaged across pitch class. The metric hierarchy, as reported by Palmer and Krumhansl (1990), is also visible, although the pattern is more subtle. Furthermore, the longer wait between the final metronome click and the later temporal positions in the probe measure did not decrease the amplitude of the oscillation in ratings between on- and off-beat temporal positions. Overall, the analyses demonstrate that
goodness-of-fit ratings differed as a function of both pitch and time, consistent with the dimensional structures of tonality and meter, and the two dimensions made independent contributions to judgments in this task and context.

Figure 3. The size of the pitch effect in Experiments 1 and 2.
Discussion

Why did pitch exhibit such a strong effect in this experiment? Perhaps the greater number of potential pitch classes (12) than temporal positions (8), led to greater discriminability (Garner & Felfoldy, 1970; Melara & Mounts, 1994) and psychological accessibility for pitch, overshadowing differences in temporal position. Another possibility is that without specific instructions to attend to either pitch or time, listeners intuitively focus on pitch, magnifying the effect of pitch class relative to temporal position. The relative strength of pitch class variations raises questions about what is the

---

2 Experiments 5 and 6 addressed this issue more explicitly and in greater detail.
most appropriate context for exploring interactions between the two dimensions and between the tonal and metric hierarchies. Perhaps the relatively modest effect of temporal position on ratings (compared to pitch class) reduced the likelihood of interactions. It is therefore of interest to search for interactions of pitch and time when temporal variation plays a stronger role.
EXPERIMENT 2: PITCH-TIME DIMENSIONAL RELATIONS IN SELECTIVE ATTENTION TO TIME

In Experiment 1, pitch structure (tonality) dominated judgments of musical probe events that varied in both pitch class and temporal position. In an attempt to increase the impact of time on ratings, listeners in the present experiment were asked to attend selectively to temporal position and to ignore pitch. This task enabled evaluation of the degree to which listeners are able to classify events in terms of temporal position. It also provided another test of potential interactions between musical pitch and time and between tonal and metric hierarchies.

Method

Participants

There were 22 adults ($M = 21.9$ years, $SD = 4.2$) with musical training ($M = 13.1$, $SD = 4.3$) who participated in this experiment. They were recruited from an introductory psychology class or through flyers posted at the University of Toronto Mississauga. Participants received course credit or $10 for their time.

Stimuli, Apparatus, and Procedure

The stimuli, apparatus, and procedure were the same as in Experiment 1, except for the instructions given to listeners. Specifically, participants were directed to ignore pitch and to consider only the temporal aspects of the melody and the probe event when making goodness-of-fit ratings.

Results
Data were again checked by examining inter-block correlations (mean $r(94) = .49$, $SD = .25$), inter-subject correlations (mean $r(94) = .63$, $SD = .08$), and correlations with the metric hierarchy of Palmer and Krumhansl (1990) for trials whose probe event used the tonic pitch (the most psychologically stable pitch; mean $r(6) = .72$, $SD = .16$). Overall, inter-block correlations were more variable in this experiment, resulting in a high standard deviation for the group. Although no one participant’s inter-block correlation fell more than two standard deviations below the mean, the inter-subject and metric hierarchy correlations of some subjects were low enough to suggest that they were clearly not attending to the task. Therefore, the exclusion criteria were adjusted such that if both of these correlations of any given participant were more than two standard deviations below the mean, their data were excluded. Six participants met these criteria, and their data were removed from further analysis. However, analyses including the data from these excluded participants obtained the same pattern of results. Furthermore, using inter-subject correlations as the only exclusion criterion resulted in removal of data from the same six participants.

A three-way repeated-measures ANOVA tested for effects of pitch class (12 levels), temporal position (8 levels) and block (3 levels) on listeners’ ratings. Pitch class and temporal position both exerted significant effects, $F(11,165) = 7.79$, $MSE = 3.09$, $p < .001$, $\eta_p^2 = .34$, $F(7,105) = 41.05$, $MSE = 20.34$, $p = .001$, $\eta_p^2 = .73$, respectively. There was no effect of block, $F(2,30) = 1.41$, $MSE = 1.74$, $p = .26$, $\eta_p^2 = .09$, and there were no significant two-way or three-way interactions. Ratings clearly varied as a function of both pitch class and temporal position, even though listeners were instructed to ignore pitch variation.
As in Experiment 1, the relative contributions of tonality and meter were examined using a stepwise linear regression. Step one introduced the tonal and metric hierarchies, and their (multiplicative) interaction entered in step two. Step one revealed that both tonal and metric hierarchies had significant predictive power in the model ($\beta = .191$ and .836, respectively; $p < .001$ for both), resulting in a multiple $R$ of .86 ($r^2 = .74$). The interaction variable from step two added no additional predictive power ($\Delta r^2 = 0$). Thus the individual hierarchies explained 74% of the variance in ratings, with meter accounting for 70% and tonality the other 4%. Figures 3 and 4 show the variability in ratings as a function of pitch class and temporal position, respectively. Reversing the pattern observed in the previous study, variation based on the tonal hierarchy was attenuated, although still present, compared to the more pronounced metric hierarchy.

Comparison of Figures 3 and 4 reveals a clear effect of instructions, with pitch dominating ratings in Experiment 1 and time prevailing in Experiment 2. Even though the degree of differentiation (or range) along the two dimensions varied across experiments, the profiles (or general shape) of these ratings were the same across experiments. That is, the ratings from both experiments reflected both the tonal and metric hierarchies, even though the range of effects differed across experiments. As a further comparison across experiments, a four-way repeated measures ANOVA compared the ratings from Experiment 1 and 2, using within-subjects factors of pitch class, temporal position, and block, and a between-subjects factor of instructions (experiment). Ratings were higher overall in Experiment 2, producing a main effect of instructions, $F(1,31) = 5.76, MSE = 72.38, p < .05, \eta^2_p = .16$. Not surprisingly, both pitch class and temporal position showed significant main effects in this analysis, $F(11,341) = 59.27, MSE = 8.41, p < .001, \eta^2_p =$
.66; $F(7,217) = 47.17, MSE = 11.56, p < .001, \eta^2_p = .6$, respectively. Block was not significant, $F(2,62) < 1, MSE = 1.46$, although it did interact with temporal position, $F(14,434) = 1.96, MSE = 1.41, p < .05, \eta^2_p = .06$, reflecting greater differentiation between temporal positions as the experiment progressed. Confirming that the instructions given to listeners did indeed modify how pitch class and temporal position influenced their ratings, instruction interacted with both pitch class and temporal position, $F(11,341) = 29.95, MSE = 8.41, p < .001, \eta^2_p = .49; F(7,217) = 28.51, MSE = 11.56, p < .001, \eta^2_p = .48$, respectively. The significant three-way interaction between pitch class, temporal position, and instruction, $F(77,2387) = 1.31, MSE = 1.21, p < .05, \eta^2_p = .04$, accounted for little of the total variance and was not clearly interpretable.

Discussion

These results indicate that listeners were able to shift the relative priority of dimensions for their goodness-of-fit rating of probe events varying in pitch class and temporal position. The extent to which they were successful in this respect indicates the degree to which they could separate the dimensions of pitch and time in this task. Furthermore, linear and additive combinations of the hierarchical structures of the two dimensions (tonality and meter) predicted the ratings in both experiments. No interactive factors increased the predictive power of either model. These findings fit with the extant body of results on the combination of pitch and temporal structures (Palmer & Krumhansl, 1987a, 1987b) or melodic and harmonic information (Schmuckler, 1989).

However, the results revealed that listeners were unable to ignore the pitch information completely, despite instructions to do so. According to Garner (1974), an inability to attend selectively to one dimension while ignoring another dimension
suggests they are not independent. Without making such strong claims of integrality or separability, these results do suggest more holistic perception of musical pitch-time probe events. However, it is important to keep in mind the much-attenuated size of the pitch effect in this experiment. Using a convergent method to explore this issue further, a speeded classification task was used in Experiments 3 and 4.
EXPERIMENT 3: PITCH-TIME DIMENSIONAL RELATIONS IN SPEEDED TEMPORAL CLASSIFICATION

The present experiment used a speeded classification task to examine how musical pitch and time work together in the perception of pitch-time probe events. Previous research on dimensional relations in musical and non-musical contexts (Ben-Artzi & Marks, 1999; Dyson & Quinlan, 2002; Krumhansl & Iverson, 1992; Melara, Marks, & Lesko, 1992; Tillmann & Lebrun-Guillaud, 2006) also used speeded classification tasks. Specifically, Krumhansl and Iverson (1992) investigated the perceptual relations between pitch and timbre, both in isolation and in a musical context. When they used a memory paradigm in a musical context, they found the dimensions functioned independently. In isolation, however, and with a speeded classification task, the dimensions interfered with each other, as Melara and Marks (1990b) found. Thus examining accuracy and reaction time in speeded classification tasks can reveal dimensional interactions that are not apparent in goodness-of-fit ratings. In the present experiment, accuracy and reaction time were measured in a speeded classification task to determine whether this task yielded a different pattern of pitch-time dimensional relations than that observed with goodness-of-fit ratings. In this task, pitch variation was not only to be ignored, but it was also task-irrelevant. Presumably, pitch would be easier to ignore in a temporal classification task than in a goodness-of-fit rating task. The orthogonal (filtering) task was used, in which all possible combinations of pitch class and temporal position formed the range of musical probe events.

Method

Participants
The participants were 36 adults (\(M = 19\) years, \(SD = 2.54\)) with an average of 9.67 years of musical training (\(SD = 3.19\)). They were recruited via the introductory psychology class and through flyers posted at the University of Toronto Mississauga.

**Stimuli**

Three new melodies were used as stimulus contexts. The increase in the number of melodies arose from comments from participants in the previous experiments about the challenge of remaining engaged in a task with a single melody. These melodies (notated in Figure 5) had the same tonality, harmony and durational denominations as the melody in Experiments 1 and 2, and therefore were also conventional examples of Western musical style. Each block had a different melody, with the order of blocks counterbalanced across listeners.

**Apparatus and Procedure**

All aspects of the apparatus and experimental procedure were the same as in the previous experiments, except for the instructions. Specifically, listeners were instructed to indicate as quickly as possible whether the probe was on or off the beat (i.e., aligned in time with the metronome clicks, if they continued into the probe measure) after hearing the probe event. The “a” and “;” key were the two response keys, with the assignment of keys counterbalanced across listeners. Participants knew that the pitch of the probe would vary. They were instructed to disregard this variation, as it had no bearing on the temporal task they were to perform. The experiment took less than an hour.
Results

Mean accuracy in temporal classification was 80% ($SD = 14.7\%$) and was unrelated to years of formal training, $r(34) = .2$, $p = .25$. Accuracy and RT correlated

Figure 5. Musical contexts and sample probe event used in Experiments 3 and 4.
negatively, $r(3454) = -.18$, indicating that there was no speed-accuracy tradeoff. Analysis of RT data customarily entails removing incorrect trials and replacing them with mean values for each participant. Unfortunately, such a procedure would result in replacing an unacceptably high level of missing data (20%), making statistics performed on the modified data set misleading. Accordingly, analyses were restricted to accuracy data.

A three-way repeated measures ANOVA on the accuracy data, using pitch class, temporal position and block as within-subjects variables revealed main effects for both pitch class, $F(11,385) = 3.66, \text{MSE} = .21, p < .001, \eta_p^2 = .1$, and block, $F(2,70) = 14.89, \text{MSE} = .48, p < .001, \eta_p^2 = .3$, but no main effect of temporal position, $F(7,245) < 1, \text{MSE} = .2$. Averaging across temporal position, accuracy across pitch class did not show an interpretable pattern, and did not correlate with the tonal hierarchy, $r(10) = -.31, p = .6$.

The effect of block reflected decreasing accuracy over blocks, perhaps because of fatigue. The interaction between temporal position and block, $F(14,490) = 5.56, \text{MSE} = .36, p < .001, \eta_p^2 = .14$, reflected listeners’ more accurate performance at earlier temporal positions in the first block but not in the second or third block. Most importantly, there was an interaction between pitch class and temporal position, $F(77, 2695) = 1.48, \text{MSE} = .11, p < .01, \eta_p^2 = .04$, reflecting systematic variations in accuracy at each temporal position according to the pitch of the probe. No other interactions were significant.

It is difficult to understand the pitch-time interaction by inspecting all 96 conditions simultaneously (12 pitch classes by 8 temporal positions). However, correlations of accuracy across pitch classes with the tonal hierarchy at each temporal position reveal the nature of the pitch-time interaction (see Table 1), as do correlations of accuracy across temporal positions with the metric hierarchy at each pitch class (see
Table 2). As can be seen in Table 1, the correlations between accuracy and the tonal hierarchy oscillate with the metric stability of each temporal position. Correlations with the tonal hierarchy are positive for temporal positions with high metric stability (i.e., positions 1, 3, 5 and 7, which are on-beat; mean $r(10) = .47$), but they are negative for off-beat positions that are therefore low in metric stability (positions 2, 4, 6 and 8; mean $r(10) = -.48$). Table 2, in comparison, shows that the correlation between accuracy and the metric hierarchy is positive for all pitch classes. Interestingly, although the correlations in Table 2 are all positive (for Experiment 3), they are highest for pitch classes with the greatest tonal stability (tonic triad members C, E and G; mean $r(10) = .87$), and lower for the remaining pitch classes (mean $r(10) = .63$).
This pattern is interpretable as a congruity effect (Pomerantz & Garner, 1973) or a redundancy gain and loss (Garner, 1974) for the temporal classifications. Accuracy improved on metrically stable temporal positions when the probe was a stable pitch in the tonality of the passage (i.e., occupied a high position in the tonal hierarchy). Conversely, accuracy improved on metrically unstable temporal positions when the probe was tonally unstable. Therefore, congruity between pitch and temporal structure improved accuracy, and incongruity was detrimental.

### Table 1

Correlations of accuracy with the tonal hierarchy for each temporal position.

<table>
<thead>
<tr>
<th>Temporal Position</th>
<th>Metric Stability</th>
<th>Experiment 3</th>
<th>Experiment 4</th>
<th>Experiment 6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Temporal Condition</td>
<td>Pitch Condition</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>High</td>
<td>0.68*</td>
<td>0.77**</td>
<td>0.75**</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>-0.52</td>
<td>0.87**</td>
<td>-0.79**</td>
</tr>
<tr>
<td>3</td>
<td>Med</td>
<td>0.76**</td>
<td>0.57</td>
<td>0.46</td>
</tr>
<tr>
<td>4</td>
<td>Low</td>
<td>-0.68*</td>
<td>0.58**</td>
<td>-0.34</td>
</tr>
<tr>
<td>5</td>
<td>High</td>
<td>0.28</td>
<td>0.45</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>Low</td>
<td>-0.27</td>
<td>0.44</td>
<td>-0.59*</td>
</tr>
<tr>
<td>7</td>
<td>Med</td>
<td>0.18</td>
<td>0.46</td>
<td>-0.12</td>
</tr>
<tr>
<td>8</td>
<td>Low</td>
<td>-0.45</td>
<td>0.47</td>
<td>-0.45</td>
</tr>
</tbody>
</table>

Note: Positive correlations indicate agreement with the tonal hierarchy.

* p < .05  ** p < .01 (two-tailed).
To determine if the effect of pitch and its interaction with time arose from perceptual difficulty or a decision bias, signal detection theory (SDT, Green & Swets,

Table 2
Correlations of accuracy with the metric hierarchy for each pitch class.

<table>
<thead>
<tr>
<th>Pitch Class</th>
<th>Tonal Stability</th>
<th>Experiment 3</th>
<th>Experiment 4</th>
<th>Experiment 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>High</td>
<td>0.80*</td>
<td>0.36</td>
<td>0.56</td>
</tr>
<tr>
<td>C#</td>
<td>Low</td>
<td>0.62</td>
<td>0.37</td>
<td>-</td>
</tr>
<tr>
<td>D</td>
<td>Med</td>
<td>0.43</td>
<td>-0.59</td>
<td>0.10</td>
</tr>
<tr>
<td>D#</td>
<td>Low</td>
<td>0.68</td>
<td>0.36</td>
<td>-</td>
</tr>
<tr>
<td>E</td>
<td>High</td>
<td>0.90**</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>F</td>
<td>Med</td>
<td>0.68</td>
<td>-0.58</td>
<td>-0.52</td>
</tr>
<tr>
<td>F#</td>
<td>Low</td>
<td>0.73*</td>
<td>0.16</td>
<td>-</td>
</tr>
<tr>
<td>G</td>
<td>High</td>
<td>0.92**</td>
<td>0.08</td>
<td>0.36</td>
</tr>
<tr>
<td>G#</td>
<td>Low</td>
<td>0.79*</td>
<td>0.06</td>
<td>-</td>
</tr>
<tr>
<td>A</td>
<td>Med</td>
<td>0.71*</td>
<td>0.32</td>
<td>-0.48</td>
</tr>
<tr>
<td>A#</td>
<td>Low</td>
<td>0.44</td>
<td>-0.34</td>
<td>-</td>
</tr>
<tr>
<td>B</td>
<td>Med</td>
<td>0.55</td>
<td>-0.28</td>
<td>-0.82*</td>
</tr>
<tr>
<td>C’</td>
<td>High</td>
<td>-</td>
<td>-</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Note: Positive correlations indicate agreement with the metric hierarchy.

* p < .05. ** p < .01 (two-tailed).
1966; Macmillan & Creelman, 1991) was used to examine performance across pitch class. For each participant, a correct response of “on the beat” was a hit, whereas incorrectly responding “on the beat” was a false alarm. By collating the data in this manner, the data could not be analyzed across temporal position. To avoid infinite $d'$ measures, proportions of hit or false alarms of 0 or 1 were converted to $1/(2n)$ or $1-1/(2n)$ respectively, where n equals the number of trials included in the ratio (Macmillan & Creelman, 1991). This correction was also used for all subsequent experiments. A one-way repeated measures ANOVA on the $d'$ scores did not vary across the within-subjects factor of pitch class, $F(11,385) = 1, MSE = .27, p = .44, \eta^2_p = .03$, implying that probe events of one pitch class were no more distinguishable than those of other pitch classes. However, a separate one-way repeated measures ANOVA (again with pitch class as the within-subjects factor) testing response bias (c) revealed a difference across pitch class, $F(11,385) = 2.67, MSE = .09, p < .01, \eta^2_p = .07$. This finding indicates that the tonal stability of the probe event influenced the perceived metric stability of the probe event. In other words, listeners were more likely to report the probe event as on-beat when the pitch of the probe event had greater tonal stability. This pattern can be seen in Figure 6, which shows that response bias followed the tonal hierarchy, $r(10) = .87$. 
Discussion

There were two main findings in the present experiment. First, consistent with Experiment 2, pitch influenced speeded classification of temporal position even when listeners were instructed to attend selectively to temporal information. Although in principle pitch variation was irrelevant to the classification task, it was not irrelevant in practice. Second, pitch affected responses in a manner that was consistent with the tonal hierarchy. Accordingly, changing the task (classification instead of goodness-of-fit ratings) revealed a different form of interference of pitch on temporal judgments of musical probe events. The influence of the tonal hierarchy on response bias but not on discriminability provides insight into the level at which pitch and temporal structure were
integrated. Specifically, because discriminability did not differ across pitch class, there was no differential perceptual difficulty in classifying temporal position based on pitch class. Instead, the implication of the response bias data is that pitch affected listeners’ judgments at a higher cognitive (decisional) level, such that the tonal stability of the pitch biased the assessment of metric stability in a congruent manner.

However, the present data do not reveal whether this pattern was global or asymmetric. Although the pitch of the probe affected temporal judgment, the corresponding task of varying the temporal position of the probe event while classifying its pitch may or may not show a similar pattern of interference. This distinction is important for discerning the nature of dimensional processing. Garner and Felfoldy (1970) demonstrated that dimensions may appear to interact when differences in stimulus discriminability are causing the effect. If the irrelevant dimension is more discriminable than the attended one, then redundancy gains and losses will result whether or not the dimensions are integral. When systematically varying the relative discriminability of dimensions, Garner interference varies accordingly and practice does not affect this pattern (Melara & Mounts, 1993).

A more complete assessment of pitch-time interrelations requires the complementary condition in which listeners classify the probe on the basis of pitch in the context of varying temporal position. If increased pitch discriminability underlies the observed interference in temporal classifications, then the congruity effect should disappear for pitch classifications, because the less discriminable, distracter dimension (temporal position) would not interfere with classification of the more discriminable, attended dimension (pitch). Such a result would denote an asymmetric failure of selective
attention. If, however, joint contributions of pitch and time to musical-probe events remain evident after switching the attended dimension, that would reveal a global failure of selective attention (Melara & Mounts, 1993). This issue was explored in Experiment 4.
EXPERIMENT 4: PITCH-TIME DIMENSIONAL RELATIONS IN SPEEDED PITCH CLASSIFICATION

Listeners were unable to attend selectively to temporal position and to ignore pitch when rating the goodness-of-fit of probe events in Experiment 2. Similar interference was evident in Experiment 3 in speeded classification of temporal position; furthermore the tonal hierarchy biased responses via a congruity effect. The present experiment explored whether comparable interference between dimensions persisted in speeded classification of pitch.

Method

Participants

The participants were 38 adults (M = 19.8 years, SD = 2.54) with an average of 10.7 years of formal musical training (SD = 2.35). Participants were recruited from psychology and music classes at University of Toronto and from flyers posted on campus.

Stimuli, Apparatus, and Procedure

Stimuli and apparatus and were the same as in the previous experiments. The procedure was the same as in Experiment 3, except for the instructions. Participants were asked to indicate whether the pitch of the probe event was “in” or “out” of the key of the passage by pressing the “a” or “;” as quickly as possible. The assignment of keys to responses was counterbalanced across listeners. Although participants knew that the timing of the probe would vary, they were instructed to disregard the temporal dimension in both the melody and the probe.
Musical organization provides a convenient control of pitch height and key membership. Of the 12 chromatic pitches in Western music, 7 of these notes are “in” a given key, or tonality; the remaining 5 pitches are “out” of the key. Because of the organization of pitch classes within any key, judgments of in-versus out-of-key are not confounded with pitch height.

Results

Preliminary inspection of the data indicated that accuracy on the pitch classification task \( M = 74\%, SD = 14\% \) was lower than on the temporal classification task of Experiment 3 \( M = 80\%, SD = 14.7\% \). Furthermore, the mean RTs (1322 ms) for the present experiment were nearly twice as long as those in Experiment 3 (748 ms). The majority of errors occurred on the diatonic pitch classes that were not members of the tonic triad (the second, fourth, sixth, and seventh scale degrees). Mean accuracy for tonic triad pitch classes was 85%. For the remaining diatonic pitch classes not in the tonic triad (D, F, A and B in the key of C major), mean accuracy was 59%; for the non-diatonic pitch classes, accuracy was 80%. Diatonic pitch classes other than those of the tonic triad occupy an intermediate level in the tonal hierarchy. Thus, although technically part of the diatonic set (and classified accurately as in-key above chance, \( t(37) = 2.1, p < .05 \)), listeners may consider their key membership ambiguous, especially in speeded tasks. Given their intermediate status in the tonality of the context melody as well as the greater incidence of errors and longer RTs, these diatonic but non-tonic triad pitch classes were excluded from the remaining analyses. Instead, the focus was on comparisons between the two extreme levels of the tonal hierarchy, namely, tonic triad and non-diatonic pitch classes.
Removing the non-tonic triad, diatonic pitch classes (D, F, A and B) resulted in increased overall accuracy ($M = 82\%, SD = 7\%)$. As in Experiment 3, accuracy and RT correlated negatively, $r(3646) = -.23$, ruling out a speed-accuracy tradeoff. Moreover, years of formal training did not correlate significantly with accuracy, $r(36) = .2, p = .25$.

A three-way repeated measures ANOVA with the within-subjects factors of pitch class, temporal position, and block was used to analyze the accuracy data. The only significant main effect was that of pitch class, $F(7,259) = 3.27, MSE = .53, p < .01, \eta_p^2 = .08$, indicating that accuracy was higher for pitch classes that had high tonal stability. Therefore, accuracy across pitch class correlated significantly with the tonal hierarchy, $r(10) = .72, p < .05$. There was no main effect of temporal position, $F(7,259) = 1.23, MSE = .11, p = .29, \eta_p^2 = .03$, nor was there a main effect of block, $F(2,74) < 1, MSE = .22$.

There was an interaction between pitch class and temporal position, $F(49,1813) = 1.52, MSE = .1, p < .05, \eta_p^2 = .04$, but no other interactions were significant.

Tables 1 and 2 show the pattern of accuracy across temporal position and pitch class. As can be seen in Table 1, the correlation with the tonal hierarchy did not change across temporal position as it did in Experiment 3. Instead, all of the correlations of accuracy with the tonal hierarchy were positive. As can be seen in Table 2, correlations of accuracy with the metric hierarchy changed as a function of pitch class, but not in accordance with tonal stability. Accuracy for the tonic triad members (C, E and G) had a weak average correlation with the metric hierarchy (mean $r(6) = .18$); the average correlation for the nondiatonic pitch classes was essentially identical (mean $r(6) = .12$). Conversely, accuracy for the excluded diatonic pitch classes that were not members of the tonic triad had a much lower average correlation with the metric hierarchy (mean $r(6)$...
= -.28). Note that these correlations are not with overall accuracy, but rather how accuracy varied across temporal position within each pitch class.

Unlike Experiment 3, neither discriminability (d’) nor response bias measures (c) accounted for the interaction between pitch class and temporal position. For this analysis, hits were correct responses of in-the-key, and false alarms were incorrect responses of in-the-key. Neither discriminability, $F(7,259) = 1.26, MSE = .21, p = .27, \eta^2_p = .03$, nor response bias, $F(7,259) = 1.39, MSE = .05, p = .21, \eta^2_p = .04$, differed across temporal position. One possible explanation for the significant interaction of pitch class and temporal position in the omnibus ANOVA is the fact that accuracy on the tonic triad pitch classes was better than the nondiatonic pitch classes only for the first four temporal positions. Accuracy for tonic triad and nondiatonic pitch classes converged at the later temporal positions (see Figure 7).
Discussion

Although there were interactions between pitch and time in Experiments 3 and in the present experiment, the interactions differ in important ways. First, interference from the irrelevant dimension varied with task instructions. In Experiment 3, there was a strong main effect of the irrelevant dimension of pitch ($\eta_p^2 = .1$). In the present experiment, however, there was no main effect of the irrelevant dimension of time ($\eta_p^2 = .03$).

Second, the structure of the irrelevant dimension affected the pitch-time interaction in Experiment 3, but not in the present experiment. Specifically, the tonal

Figure 7. Accuracy for the tonic triad pitches and nondiatonic pitches across temporal position for Experiment 4.
hierarchy dictated the pattern of interference of pitch on temporal classifications (in the form of a response bias), but the reverse did not occur. Performance in the present experiment was unrelated to the metric hierarchy (see Table 1), with neither discriminability nor response bias differing across temporal position.

Third, the interaction in the present experiment can be interpreted as perceptual decay over time. In Experiment 3, participants always performed better when the probe occupied a more stable position in the hierarchy of the relevant dimension (i.e., on-beat probes better than off-beat probes). In the present experiment, accuracy was higher for in-key probes than for out-of-key probes, but this advantage decreased with increasing temporal distance from the melody (i.e., further along in the probe measure). As with any stimulus, details fade over time. Thus listeners’ perception of the difference between in-key and out-of-key probes decreased for those occurring later in the probe measure. Lastly, and as Experiment 6 will reveal, this pitch-time interaction in pitch classifications was not replicated in a later experiment, whereas it persisted for temporal classifications. Instead, there appears to be an asymmetry in selective attention to pitch versus temporal dimensions in these classification tasks. Musical pitch variation interfered with time judgments, but time did not interfere with pitch judgments.

Before considering the implications of this asymmetry, it is important to highlight the limitations of these data. It is possible, for example, that the present tasks emphasized the pitch dimension more than the temporal dimension. If so, then the observed asymmetry may not be a general process of musical probe event perception but rather one that is idiosyncratic to the experimental context. Which factors could have influenced the findings? First, some of the pitch classes used as probe events were not present in the
initial musical context. In contrast, all of the temporal positions employed as probe events occurred within the preceding context. The appearance of new pitch classes could have surprised the listener, resulting in accentuated attention to the pitch dimension relative to the temporal dimension.

A second, related factor derives from potentially greater emphasis on pitch arising from more pitch classes (12) than temporal positions (8) in the probe events. This inequality may have increased the relative importance of the pitch dimension and thus could have contributed to the strong interfering effect of pitch on temporal judgments in Experiment 3. Stimulus inequality could also offer an explanation for the unexpectedly poor performance on pitch classification in the present experiment. Recent work on Stroop effects offers a potential explanation for these findings. Melara and Algom (2003) propose that adding levels to one dimension (pitch in the present case) lowers the probability of occurrence of any particular level on a given trial and leads to two related phenomena. Specifically, more levels in the irrelevant dimension draw attention away from the relevant dimension with fewer levels. Consequently, the likelihood of interference from the irrelevant dimension increases, but accuracy is unaffected. Additionally, accuracy declines when classifying the dimension with more levels because the relevant dimension becomes less predictable, yet interference from the irrelevant dimension is less likely. The results of Experiments 3 and the present experiment are consistent with both of these phenomena – interference from the irrelevant dimension of pitch when classifying time (but accuracy unaffected) along with decreased accuracy when classifying all 12 pitch classes (but no interference from time). As a counter example, other research has demonstrated that in a non-musical context, increasing the
number of pitches relative to the number of loudness levels actually decreased (instead of increased) the degree to which pitch interfered with loudness judgments (Melara & Mounts, 1994). Nevertheless, it remains possible that the same pattern does not hold in musical contexts.

Differences in discriminability between dimensions comprise the most common explanation for asymmetric results such as those observed in the present research. Even for demonstrably independent dimensions, the more discriminable of two dimensions tends to influence the less discriminable dimension and not vice versa (Garner, 1974; Garner & Felfoldy, 1970; Melara & Algom, 2003; Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001). Melara and Mounts (1993) eliminated and even reversed the classic Stroop asymmetry (Stroop, 1935) by manipulating the relative discriminability of color and word dimensions. In each case, these authors found that the more discriminable of two dimensions consistently interfered with judgments of the less discriminable dimension. It is important to note, however, that such findings would actually predict findings opposite to those reported here. Given that performance was generally better for temporal than for pitch classification, the temporal dimension could be considered to be more discriminable and hence should have interfered with pitch classification. Accordingly, of the two explanations – unequal stimulus quantity and discriminability – the former factor is more likely to provide a satisfactory account for these results. Regardless, both factors may have contributed to these findings.

Presumably, issues of unequal stimulus quantity and dimensional discriminability can be reduced by using an equal number of pitch classes and temporal positions in baseline tasks in addition to an orthogonal (filtering) task. Baseline tasks can determine
whether discriminability is equal between dimensions by varying a single dimension while the other remains constant. Equal numbers of levels for both pitch and temporal dimensions avoid set-size effects that could unbalance efforts to establish equal discriminability in the baseline tasks (Melara & Mounts, 1994). Equal discriminability in the context of baseline pitch and time tasks that involve equal numbers of pitch classes and temporal positions enhance the validity and reliability of filtering tasks for investigating dimensional interference. The goal of Experiment 5 was to provide these baseline measures for such a subsequent filtering task (Experiment 6).
EXPERIMENT 5: BASELINE PITCH AND TIME DIMENSIONAL RELATIONS

In Experiment 3, listeners were instructed to ignore pitch while judging the temporal position of probes, with probe events encompassing 12 possible pitch classes and 8 possible temporal positions. Experiment 4 reversed the relevant and irrelevant dimensions of this task by asking listeners to judge pitch while ignoring time, and used the same probe events. The present experiment explored the impact of inequality in number of pitch classes and temporal positions in probe events by having listeners perform baseline classifications of pitch and time with eight pitch classes and temporal positions. Furthermore, all pitch classes and temporal positions used in probe events occurred in the preceding context. For the two pitch and temporal baseline tests (performed separately), one dimension changed while the other remained constant. If accuracy of classification is comparable for pitch classes and temporal positions, then unequal discriminability between dimensions would not confound subsequent tests of selective attention to pitch and time.

The present experiment also provided baseline goodness-of-fit ratings for the 12 pitch classes and 8 temporal positions used in the previous experiments. These ratings made it possible to ascertain whether the musical contexts were comparable in successfully invoking the tonal and metric hierarchies when only one dimension was varied in the probe event.

Method

Participants
The participants were 16 adults ($M = 20.1$ years, $SD = 3.4$), who received credit in an introductory psychology class or monetary compensation. Average musical training was 10.5 years ($SD = 2.8$).

**Stimuli**

The classification baselines used a small subset of probe events from the previous experiments, with one additional pitch. The pitch classification trials used the seven diatonic notes of C major as well as the C from the next octave, resulting in a probe set of increasing pitch height: C, D, E, F, G, A, B, and C'. Half of the pitches in this set belong to the C major tonic triad (C, E, G and C'), with the remaining half being diatonic pitch classes that are not in the tonic triad. All pitch classes were played on the downbeat at the start of the probe measure, resulting in eight possible probe events. For the temporal classification trials, probe events always had the same pitch (C) and had the same eight temporal positions as in the previous experiments (half on-beat, half off-beat), yielding eight unique probe events. Baseline goodness-of-fit ratings used the same stimuli as the classification baselines, adding the nondiatonic pitch classes to produce 12 pitch classes.

Trials were blocked by response type (classification or goodness-of-fit ratings) and dimension (pitch or time). All blocks had two repetitions of the melodies from Experiments 1 and 2 (Figure 2) and Experiments 3 and 4 (Figure 5), resulting in 64 trials in each of the classification blocks and 96 trials in each of the goodness-of-fit blocks. The entire experiment took approximately one hour.

**Apparatus and Procedure**

The apparatus and procedure were the same as in the previous studies. Listeners were instructed to indicate as quickly as possible whether the probe event was on or off.
the beat for temporal classification trials and in or out of the tonic triad for pitch classification trials. Listeners responded on temporal classification trials by pressing the “a” and “;” keys, with the assignment of keys counterbalanced across listeners. For goodness-of-fit ratings, listeners used a 7-point rating scale, with 1 denoting a poor fit with the context and 7 denoting a good fit.

Results

Classification Tasks

Accuracy was 80% (SD = 14%, RT = 1213 ms) for temporal classification and 87% (SD = 12%, RT = 1742 ms) for pitch classification. The correlation between accuracy and RT was negative for temporal classifications, \( r(126) = -.18 \), and pitch classifications, \( r(126) = -.08 \), ruling out speed-accuracy tradeoffs in both tasks. Years of training and accuracy did not correlate significantly, \( r(14) = .33, p = .1 \). Of greatest interest, a one-way repeated measures ANOVA on accuracy scores using dimension (temporal or pitch) as the within-subjects factor revealed no difference in accuracy between the temporal and pitch dimensions, \( F(1,15) = 2.8, MSE = .01, p = .12, \eta_p^2 = .16 \). When accuracy scores were converted to \( d' \) scores (discriminability), mean \( d' \) for temporal classification was 2.13 (SD = 1.3), and 2.69 for pitch classification (SD = 1.2). Performance did not differ between dimensions, \( F(1,15) = 2.41, MSE = 1.03, p = .14, \eta_p^2 = .14 \).

Further analyses tested if accuracy differed across temporal position and pitch class, and the respective roles of the metric and tonal hierarchies. Temporal classifications were analyzed with a one-way repeated measures ANOVA, with temporal position as a within-subjects variable. Accuracy differed significantly across temporal
position, $F(7,105) = 5.67, \text{MSE} = .03, p < .001, \eta_p^2 = .27$, with this profile reflecting the metric hierarchy, $r(6) = .78, p < .05$. Listeners were more accurate for on-beat than for off-beat positions. Accuracy of pitch classification differed significantly across pitch class, $F(7,105) = 6.78, \text{MSE} = .03, p < .001, \eta_p^2 = .31$, but this profile did not reflect the tonal hierarchy, $r(10) = .02, p = .96$. Instead, listeners tended to perform slightly worse on tonic triad members other than the tonic (i.e., E and G); this finding may reflect their lesser prototypicality as exemplars of the tonic triad.

**Goodness-of-Fit Ratings**

To determine whether the musical contexts established the tonal and metric hierarchies, goodness-of-fit ratings were analyzed with two separate one-way repeated measures ANOVAs, with either temporal position or pitch class as the within-subjects variable. Goodness-of-fit ratings varied significantly across the eight temporal positions, $F(7,105) = 11.32, \text{MSE} = 1.09, p < .001, \eta_p^2 = .43$, reflecting the metric hierarchy, $r(6) = .80, p < .05$. Similarly, goodness-of-fit ratings varied significantly across pitch class, $F(11,165) = 70.76, \text{MSE} = .58, p < .001, \eta_p^2 = .83$, reflecting the tonal hierarchy, $r(10) = .99, p < .001$.

**Discussion**

The baseline tests reveal that for the classification tests, the hierarchical pitch and temporal structure was equally discriminable in these musical contexts. Furthermore, the goodness-of-fit ratings suggest that the musical contexts were comparable in inducing the tonal and metric hierarchies. As such, it seems unlikely that the asymmetry in interference between pitch and time in the filtering tasks of the previous experiments can be explained by simple differences in discriminability between these dimensions.
However, it is important to determine whether similar (i.e., asymmetric) interference is evident when the filtering task involves equal numbers of pitch classes and temporal positions. Armed with the baselines from the present experiment, Experiment 6 used the filtering task to address this question.
EXPERIMENT 6: PITCH-TIME DIMENSIONAL RELATIONS IN EQUALIZED CLASSIFICATIONS

Experiments 3 and 4 revealed asymmetric interference between pitch and time. As noted, however, differences in the set size of pitch classes and temporal positions may have contributed to this finding. Although Experiment 5 found no difference in discriminability between pitch and temporal dimensions using a set of baseline tasks, the task intentionally employed stimuli with equal levels of pitch and temporal information (at least in the classification task). Therefore, it remains possible that the earlier asymmetry arose from unequal numbers of levels of these stimuli that then affected discriminability. The goal of the present experiment was to provide more definitive evidence of asymmetric interference in the filtering task using stimuli whose dimensions were of equivalent discriminability.

The present experiment used the filtering task of Experiments 3 and 4 with the matched stimuli of Experiment 5. Half of the listeners received instructions to ignore pitch while judging the timing of the probe, and half received the opposite instructions. If the influence of pitch on temporal classification, which was consistent with the tonal hierarchy, arose from unequal stimulus quantity or discriminability, then asymmetric interference should not be evident in the present conditions. In contrast, a finding of asymmetric interference of pitch on time would imply that the previous results cannot be attributed to unequal stimulus quantity or differences in discriminability.

Method

Participants
The participants were 36 adults ($M = 18.8$ years, $SD = 1.41$) who received credit in an introductory psychology class or monetary compensation. Average musical training was 10 years ($SD = 2.02$). Participants were randomly assigned to the task of judging temporal position ($n = 18$) or pitch class ($n = 18$).

**Stimuli**

The stimuli consisted of all possible combinations of the pitch classes and temporal positions used in the baseline classification tasks of Experiment 5. Specifically, there were 8 pitch classes (half tonic triad, half non-tonic triad) and 8 temporal positions (half on-beat, half off-beat), producing a total of 64 unique probe events. There were four blocks of trials, with a unique melody used in each block. The four melodies were the same as those in Experiment 5.

**Apparatus and Procedure**

The apparatus and procedure were the same as in the previous experiments. Listeners were instructed to indicate as quickly as possible whether the probe event was on or off the beat, or in or out of the tonic triad, depending on the assigned condition. There were 4 blocks of 64 trials, which they completed in just under an hour.

**Results**

Overall accuracy was 77% (RT $M = 1281$ ms) for temporal classification and 78% (RT $M = 1657$ ms) for pitch classification. Accuracy and years of formal training did not correlate significantly, $r(34) = .14, p = .41$. Accuracy and RT correlated negatively, $r(2302) = -.23, p = .17$, ruling out a speed-accuracy tradeoff.

A four-way repeated measures ANOVA, with pitch class, temporal position and block as within-subjects variables and instruction as a between-subjects variable,
revealed main effects of pitch class on accuracy, $F(7, 238) = 4.31, MSE = .56, p < .001$, $\eta^2_p = .11$, temporal position, $F(7, 238) = 6.54, MSE = .39, p < .001, \eta^2_p = .16$, and block, $F(3, 102) = 3.3, MSE = .37, p < .05, \eta^2_p = .09$. These main effects indicate that accuracy differed as a function of the pitch class and temporal position of the probe and that performance improved during the course of the experiment. There was no main effect of instruction, $F(1, 34) < 1, MSE = 4.03$. There were two-way interactions between pitch class and instruction, $F(7, 238) = 4.32, MSE = .56, p < .001, \eta^2_p = .11$, temporal position and instruction, $F(7, 238) = 4.57, MSE = .39, p < .001, \eta^2_p = .12$, and temporal position and block, $F(21, 714) = 1.81, MSE = .13, p < .05, \eta^2_p = .05$. These interactions reflected the larger effects of pitch and time when they were the relevant dimensions. The interaction of time and block reflected lesser differentiation between temporal positions in the first block compared to the remaining blocks. Of principal interest, however, was the significant three-way interaction of pitch class, temporal position, and instruction, $F(49, 1666) = 1.76, MSE = .13, p = .001, \eta^2_p = .05$. Inspection of this interaction suggested that pitch and time interacted in temporal classifications but not in pitch classifications. To verify this interpretation, two subsequent three-way repeated measures ANOVAs (one for each instruction type) tested the within-subjects variables of pitch, time, and block. These analyses are described separately, in turn.

For temporal classification, there were significant main effects of time, $F(7, 119) = 6.12, MSE = .68, p < .001, \eta^2_p = .27$, and block, $F(3, 51) = 3.44, MSE = .42, p < .05, \eta^2_p = .17$. These effects indicated improved accuracy for on-beat probes compared to off-beat probes and for later blocks compared to earlier blocks. Accuracy across temporal position followed the metric hierarchy, $r(6) = .75, p < .05$. There was no main effect of
pitch class, $F(7, 119) = 1.74, \text{MSE} = .11, p = .11, \eta^2_p = .09$. The interaction between time and block was significant, $F(21, 357) = 1.83, \text{MSE} = .15, p < .05, \eta^2_p = .1$. Most importantly, and replicating Experiment 3, there was a significant interaction between pitch and time, $F(49, 833) = 1.97, \text{MSE} = .14, p < .001, \eta^2_p = .1$. No other interactions were significant.

A signal detection analysis further explored the interaction between pitch and time. As in Experiment 3, $d'$ scores did not differ across pitch class, $F(7, 119) = 1.49, \text{MSE} = .17, p = .18, \eta^2_p = .08$, but response bias ($c$) did, $F(7, 119) = 3.47, \text{MSE} = .11, p < .01, \eta^2_p = .17$. Moreover, response bias again followed the tonal hierarchy, $r(6) = .69, p = .06$ (see Figure 8). It is likely that the marginal significance of this correlation coefficient stems from the removal of an entire level of the tonal hierarchy (nondiatonic pitch classes).
Table 1 displays correlations between accuracy as a function of pitch class and the relevant tonal hierarchy values for each temporal position. As in Experiment 3, there was a regular oscillation in these correlations, with positive correlations for on-beat temporal positions (positions 1, 3, 5 and 7; mean $r(6) = .38$) and negative correlations for off-beat positions (positions 2, 4, 6 and 8; mean $r(6) = -.54$). Table 2 shows correlations between accuracy at each pitch class and the metric hierarchy. As in Experiment 3, correlations with the metric hierarchy were higher for tonally stable pitches, mean $r(6) = .40$ and -.43 for tonic triad and non-tonic triad tones, respectively. In short, results of the temporal classifications replicate the congruity effect of Experiment 3.

*Figure 8.* Response bias to respond “on the beat” across pitch for Experiment 6, temporal condition.
For pitch classifications, there was a main effect of pitch class, $F(7, 119) = 4.58, \text{MSE} = 1.02, p < .001, \eta^2_p = .21$. However, this effect only weakly followed the tonal hierarchy, $r(6) = .4, p = .16$. Instead, this effect can be explained in terms of increased accuracy for probes with the tonic pitch (C) relative to the remaining pitch classes. No other effects or interactions were significant. Most importantly, there was no interaction between pitch and time, $F(49, 833) < 1, \text{MSE} = .12$. As in Experiment 4, signal detection analyses failed to reveal differences in discriminability, $F(7, 119) = 1.69, \text{MSE} = .14, p = .12, \eta^2_p = .09$, or response bias, $F(7, 119) < 1, \text{MSE} = .05$, as a function of temporal position. Table 1 shows that correlations with the tonal hierarchy did not vary across temporal position, and instead were uniformly positive, mean $r(6) = .35$ and .41 for on- and off-beat temporal positions, respectively. Likewise, Table 2 shows that correlations with the metric hierarchy were also positive across pitch class, mean $r(6) = .16$ and .29 for the tonic triad and non-tonic triad tones, respectively.

Discussion

Overall, these results replicate Experiments 3 and 4 in that a response bias driven by the tonal hierarchy influenced listeners’ temporal judgments, and no response bias based on the metric hierarchy affected pitch judgments. Again, temporal classification was characterized by a congruity effect in terms the impact of pitch on time. Finally, there was convergence between the current experiment and Experiments 3 and 4 in that discriminability did not differ across pitch class or temporal position. Instead, the difference across pitch class in temporal classification was specific to response bias. This finding reinforces the idea that a higher-level decisional bias rather than perceptual
difficulty underlies the observed effect of pitch on time. Table 3 summarizes the classification judgment data to illustrate the asymmetrical effects of pitch on time.

Table 3

Summary of experimental findings on classification judgments (Experiments 3-6).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Effect of Pitch</th>
<th>Effect of Temporal Position</th>
<th>Response Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (Classify Time)</td>
<td>Yes</td>
<td>No</td>
<td>Yes*</td>
</tr>
<tr>
<td>4 (Classify Pitch)</td>
<td>Yes*</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>6 (Classify Time)</td>
<td>No</td>
<td>Yes†</td>
<td>Yes*</td>
</tr>
<tr>
<td>6 (Classify Pitch)</td>
<td>Yes*</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

* Follows Tonal Hierarchy
† Follows Metric Hierarchy

These results also address the previously noted limitations of Experiments 3 and 4. Specifically, the present experiment used equal numbers of pitch classes and temporal positions (all of which were in the context prior to the probe), and still produced the same asymmetries observed in Experiments 3 and 4. Therefore, neither differences in stimulus quantity nor the presence or absence of probe pitches in the melodies used to establish the context can explain the asymmetry between pitch and time. Furthermore, performance did not differ between temporal and pitch instructions. Accordingly, differences in stimulus discriminability cannot explain the persistent asymmetric effects of pitch on time in these musical contexts.

In comparing goodness-of-fit ratings to classification judgments, one might consider the interference of pitch on time to be greater for classification judgments.
However, this characterization of the data is inappropriate. For classification judgments as well as goodness-of-fit ratings, the tonal hierarchy exerted robust interference but in a different form. The tonal hierarchy influenced ratings of temporal goodness-of-fit in accordance with the probe’s tonal stability. The tonal hierarchy affected judgments of metric classification by biasing listeners to respond in accordance with the probe’s tonal stability.

Why did the congruity effect occur only for classification judgments? In classification judgments, tonal-metric congruity (high tonal and metric stability, or low tonal and metric stability) led to increased accuracy. When classifying the metric stability of a probe that was off the beat, nondiatonic pitches probably aided accuracy because they reinforced and facilitated the judgment. For goodness-of-fit ratings, however, a congruity effect would be counterintuitive. Consider a probe that has both low tonal and metric stability (congruent). A congruity effect would require this probe to receive a higher rating than a probe with incongruent stability (high tonal and low metric stability, or low tonal and high metric stability), even though the incongruent probe was more stable overall.

Similarly, one may ask why tonally stable pitches were not invariably more accurate than unstable pitches in temporal classification judgments. When classifying metric stability, tonally stable pitches increase accuracy for on-beat probes, but tonally unstable pitches increase accuracy for off-beat probes (the congruity effect). Averaging across temporal position therefore results in fairly constant accuracy for all pitches. At the very least, differences across pitch class are unlikely to correspond to tonal stability. However, when rating temporal goodness of fit, tonally stable pitches always received
higher ratings regardless of temporal position, whereas tonally unstable pitches had the opposite effect. This consistent effect of tonal stability caused the tonal hierarchy to be more apparent when averaging across temporal position for goodness-of-fit ratings than for classification accuracy.

Ultimately, the nature of the instructions determined whether the interference of pitch on time (by way of the tonal hierarchy) appeared as a congruity effect for classification judgments or as an additive effect for goodness-of-fit ratings. In both cases the tonal hierarchy strongly influenced responses. Accordingly, it is misleading to suggest that the interference of pitch on time was stronger in one task than in the other.

Why did pitch predominate in Experiments 1-6? One explanation is that the salience of pitch in these typical musical contexts made it difficult for listeners to ignore pitch information even when it was irrelevant. Although stimulus inequality between pitch and time may have contributed to the results of Experiment 1, results of the other experiments imply a dominance of pitch that is independent of instructions. The greater salience of pitch than time is in line with previous findings on long-term memory for music (Hébert & Peretz, 1997) and the perception of metric structure (Dawe, Platt, & Racine, 1994).

Alternative explanations of and limitations to these findings merit consideration. First, perhaps time did not get a fair chance. A larger effect of time may require the presentation of additional levels of the temporal hierarchy. This argument implies that a wider range of temporal events in the preceding musical contexts is necessary to activate the metric hierarchy. This logic also suggests that subdivisions of the metric hierarchy do not occur without explicit presentation of events at those temporal locations in preceding
contexts. Moreover, this argument further implies a comparable effect for pitch, with the tonal hierarchy instantiated only if all pitches are presented in a tonal melody. Ultimately, this argument is unsupported based on the wealth of data in the music cognition literature indicating that impoverished musical contexts involving a few notes, a simple chord, or short rhythmic patterns activate both tonal (Cuddy & Badertscher, 1987; Krumhansl, 1990; Oram & Cuddy, 1995; N. A. Smith & Schmuckler, 2004) and metric hierarchies (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Desain & Honing, 2003; Large & Palmer, 2002; Palmer & Krumhansl, 1990; Povel & Okkerman, 1981; Stoffler, 1985; Tekman, 2001). For example, Palmer and Krumhansl (1990) provided only one level of the metric hierarchy (the tactus, or beat level) and still observed nuanced metric hierarchies of matching temporal locations. More directly to the point, the baseline goodness-of-fit ratings in Experiment 5 demonstrated that these contexts were sufficient to induce strong tonal and metrical hierarchies. Thus, there is little basis for the notion that perception of the metric hierarchy was weaker in these studies than was perception of the tonal hierarchy.

Perhaps the most obvious limitation derives from the structure of the melodic context provided prior to the probe event. In Western music, the tonal and metric hierarchies show a strong positive correlation, with metrically stable temporal positions typically containing tonally stable pitches. The context melodies used here were no exception. This correlation makes it difficult to ascertain the relative contribution of the tonal and metric hierarchies. However in tests of rhythm perception that put pitch and temporal factors in opposition, Dawe, Platt, and Racine (1993; 1994; 1995) found independent contributions of both, and even a dominance of pitch. But to what extent do
the current findings depend on the specific context used? Perhaps the rich pitch structure in these musical contexts magnified the relative importance of pitch compared to time. It is possible that other contexts with varied strength of pitch and temporal structure might elicit corresponding differences in the perception of pitch-time probe events. Consequently, these results cannot currently generalize to all forms of Western music.

While considering this limitation, it is important to remember that the statistical properties of a standard Western tonal context are not the exclusive determinant of listeners’ perceptions. As noted, not any hierarchical ordering of pitch classes will result in a corresponding tonal hierarchy, but only the one to which listeners have been exposed throughout life (see N. A. Smith & Schmuckler, 2004, but also see Oram & Cuddy, 1995). Therefore it is likely that the existence of stored representations of the statistical properties of music are activated by a suitable context, perhaps in an all-or-none fashion, much like retrieval cues in memory. If so, any musical context that successfully activates tonal and metric hierarchies should produce findings like those reported here. Nevertheless, testing relative dimensional dominance, while systematically varying the pitch and temporal structure in musical contexts, is a fruitful avenue for further research.

Given the possibility that variations in the structure of the melodic context may give rise to differential patterns of pitch-time integration, the following experiments test how varying pitch structure (specifically tonality) changes how pitch and time combine in music perception. The next three experiments examine the circumstances under which the violation of temporal expectancies affects the processing of a pitch event.
EXPERIMENTS 7-9: PITCH-TIME DIMENSIONAL RELATIONS IN TEMPORAL EXPECTANCIES

Experiments 1-6 revealed an asymmetric pattern of interference in judgments of probe events following a musical context. Specifically, pitch influenced time, but time did not influence pitch. These findings conflict with the theory of dynamic attending (Jones & Boltz, 1989) that describes how a regular rhythmic sequence establishes temporal expectancies for subsequent events. According to this theory, violating these temporal expectancies is detrimental to processing. For a probe event presented off the beat, as in Experiments 4 and 6, dynamic attending theory would predict a decline in the accuracy of key membership judgments. In fact, there were no variations in accuracy across temporal position for judgments of key membership.

Jones and her colleagues have provided a variety of findings that are consistent with the notion of dynamic attending, and with interactive processing of pitch and temporal events (Jones et al., 2006; Jones et al., 2002). For example, Jones et al. (2002), demonstrated that pitch judgments of rhythmically expected tones are more accurate than pitch judgments of unexpected tones. In this study, listeners heard a standard tone, followed by an isochronous sequence of eight random pitches (i.e., atonal), followed by a comparison tone. Then they judged whether the comparison tone was higher or lower in pitch than the standard tone. When the intervening sequence had a regular rhythmic (isochronous) structure, listeners had clear expectancies about the timing of the comparison tone. When the comparison tone occurred on time, accuracy of the pitch height judgments was better than when the comparison tone occurred early or late. In a subsequent study, (Jones et al., 2006) had participants listen for a pitch change in a
repeating nine-tone sequence. They found that the pitch content of the sequence and
target as well as the timing of the target tone affected the accuracy of detecting a pitch
change. In short, both sets of findings demonstrate the importance of joint temporal and
pitch expectancy formation, highlighting the idea that variation in the temporal structure
of a sequence can influence pitch processing.

How can the findings of Experiments 1-6 be reconciled with the theory of
dynamic attending as well the findings of Jones and her associates? Close comparison of
the present research with that of Jones reveals a number of methodological differences
that could account for the divergent results. In Jones et al. (2002), listeners judged the
relative pitch height of a standard tone and comparison tone that were separated by an
intervening sequence (i.e., short-term recognition memory). Experiments 1-6, by contras,
used a variant of the probe tone procedure (Krumhansl, 1990; Krumhansl & Kessler,
1982; Krumhansl & Shepard, 1979) in which listeners rate how well a probe event fits
with a preceding musical context. Although Krumhansl (1979) reported convergent
findings from the use of probe tone ratings and a memory task, there have been few direct
comparisons. As a result, it remains possible that task differences account for the
divergence in findings.

A related difference concerns the nature of the processing required in the
comparison/memory and rating tasks. Comparing the pitch height of an earlier and later
tone relies on memory of the earlier event and a comparison with the later event on the
basis of the simple attribute of pitch height. The goodness-of-fit ratings, however, require
the processing of a complex attribute, namely, key membership. Thus, differential
processing requirements could account for the divergent results.
There are also considerable differences in the stimuli used in the two bodies of research. In Jones et al. (2002), the intervening sequences between comparison tones consisted of a set of random pitches with isochronous timing. These sequences promoted temporal expectancies because of their regular timing, but they could not promote pitch expectancies because they were unpredictable and unrelated to tonal conventions. In contrast, the present experiments used tonal contexts that were harmonically and rhythmically diverse. As a result, the presence or absence of tonal structure and isochrony in the context could underlie the differences across studies. These issues are explored more fully in the general discussion.

The purpose of the remaining experiments was to explore sources of the divergent findings between Experiments 1-6 and Jones et al. (2002). Three experiments examined the consequence of violating temporal expectancies for pitch perception, which Jones et al. (2006; Jones et al., 2002) consider most indicative of pitch-time relations. These experiments employed a variant of the short-term memory task used by Jones et al. (2002), in which listeners heard a standard tone, an intervening sequence, and a comparison tone. These experiments differed, however, by the nature of the task (Experiment 7) and the pitch structure of the intervening context (Experiments 8 and 9). Specifically, Experiment 7 tested the impact of task variation on pitch-time relations. The presence of tonal information was held constant by using the simple tonal sequences from the previous experiments as the intervening sequence. Experiment 8 assessed the role of the pitch structure (tonal versus atonal) of the context. Tonal contexts established a key by means of a hierarchical distribution of pitch classes, whereas atonal contexts consisted of random pitches that did not establish a key. The task remained constant in Experiment
8 by using only pitch height judgments. Experiment 9 also tested how pitch structure affected temporal expectancies and controlled for various stimulus differences between the tonal and atonal contexts.
EXPERIMENT 7: THE ROLE OF TASK IN PITCH-TIME DIMENSIONAL RELATIONS

The purpose of the present experiment was to examine how the violation of temporal expectancies influences the accuracy of pitch judgments that follow tonal contexts. Type of processing was manipulated to examine its effects. To achieve this goal, this experiment included typical tonal contexts (from the previous experiments) that conformed to the rules of Western tonal music. The task required either a comparison of pitch height or a judgment of key membership. The pitch height comparison task was the same as that used by Jones et al. (2002) in which the listeners compared the pitch of a standard tone and comparison tone separated by a sequence of intervening notes. Because listeners are instructed to ignore the intervening tones, the pitch structure of the context is irrelevant, especially because the task involves a comparison of more basic perceptual attributes than tonality. The key membership judgment in this experiment used the same stimuli as the pitch height comparison, but now required listeners to determine whether the comparison tone that followed the context belonged in the key of that context. In other words, this classification required listeners to attend to the tonality of the context and evaluate how the final tone fit with this structure. This task, then, requires listeners to ignore the initial standard tone and involves the processing of higher-level, or more complex, information.

The stimuli consisted of a standard tone, followed by a tonal context, and by a final comparison tone whose timing was early, on-time, or late relative to the temporal framework induced by the context. This means of violating temporal expectancies by manipulating the timing of the comparison tone is borrowed from Jones et al. (2002).
When the simple tonal contexts from the previous experiments are used as intervening sequences between the standard and comparison tones, they present a problem for the pitch height comparison. Because the contexts induce an organized tonal framework, this organization could facilitate encoding and retention of the standard and comparison tones. Indeed, listeners commonly encode pitches presented in a musical context according to their position in the tonal hierarchy (Bharucha, 1984, 1996). Consequently, listeners might identify the standard and comparison tones within the tonal framework and then use their knowledge of relations within the tonal framework to determine which pitch was higher. For example, if the standard tone was the second scale degree and the comparison tone was the third scale degree, then listeners could determine that the comparison tone was higher on the basis of relative scale degree. Because musicians easily use such information in the presence of a tonal context (Deutsch, 1980), this strategy would circumvent the task by obviating the need to remember the pitch height of standard and comparison tones. Furthermore, use of this strategy would mean that listeners applied a strategy that changed the nature of the pitch height comparison, defeating the purpose of the task manipulation.

One way to prevent listeners from using this strategy is to eliminate the ability to encode the pitch of standard tone into a tonal framework. The most straightforward means of accomplishing this goal is to use standard tones that are not members of the pitch classes used in Western music. The comparison tones, however, and the tonal context itself, used pitches drawn from typical pitch classes in tonal music. Accordingly, standard tones in this experiment consisted of frequencies halfway between the frequencies of two neighboring members of the equal-tempered Western pitch set. These
pitches are called “quarter tones” in musical terminology. Table 6 lists the frequencies of some equal-tempered and quarter tones (indicated by a + or – sign).

If task type accounts for the divergent findings between Jones et al. (2002) and Experiments 1-6, then violating temporal expectancies should affect performance only for pitch height comparisons and not for key membership judgments. This prediction stems from the fact that the former experiment revealed an effect of time on pitch height comparison, whereas no such effect appeared in the latter experiments on key membership judgments. In contrast, if task type does not account for these differences, then temporal variation of the comparison tone should have comparable effects across both task types.

Method

Participants

Twelve adults (18.3 years, SD = 1.9) with at least eight years of formal training participated in this experiment for credit in an introductory psychology class at the University of Toronto Mississauga or payment of $10. Average musical training was 10.1 years (SD = 1.9) and the average age was 18.3 years (SD = .5). Half of the participants remained musically active, and none reported possessing absolute pitch.

Stimuli

Table 3 denotes the pitch height (frequency) of standard and comparison tones in this experiment. There were four possible comparison tones, which were drawn from the Western tonal pitch set: F#₄, G₄, G♯₄, and A₄. There were three possible standard tones, consisting of the quarter tones halfway between the four comparison tones in pitch height. These tones are designated G−, G+, and A−. Each standard tone was associated
with the two comparison tones whose pitch height was directly lower or higher than the frequencies of the standard tone. Therefore, for each standard tone there was one comparison tone that was lower in pitch, and one that was higher (also depicted in Table 4).

Table 4

Fundamental frequency of standard and comparison tones, correct judgments of pitch height, and key membership judgments, for Experiment 7.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Standard Pitch</th>
<th>Comparison Pitch</th>
<th>Pitch Height Judgment</th>
<th>Key Membership Judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>440</td>
<td>A4</td>
<td>Higher</td>
<td>In Key</td>
<td></td>
</tr>
<tr>
<td>428</td>
<td>A-4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>415</td>
<td>G#4</td>
<td>Higher/Lower</td>
<td>Out of Key</td>
<td></td>
</tr>
<tr>
<td>404</td>
<td>G+4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>392</td>
<td>G4</td>
<td>Higher/Lower</td>
<td>In Key</td>
<td></td>
</tr>
<tr>
<td>381</td>
<td>G-4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>F#4</td>
<td>Lower</td>
<td>Out of Key</td>
<td></td>
</tr>
</tbody>
</table>

Figure 9 shows a schematic representation of an experimental trial. Each trial began with one of the three standard tones, which had a duration of 250 ms. The standard tone was followed by a silent interval of 750 ms, and then the tonal context. There were four different tonal contexts, which were the same as those used in Experiments 5 and 6, consisting of a melody with harmonic accompaniment (see Figures 2 and 5). As before, a
A metronome click sounded on each beat (every 500 ms) throughout the context, and continued beyond the end of the context for another four beats (2 s). This click maintained the temporal framework of the tonal context, and provided a break between this context and the comparison tone.

Following the interval of metronome clicks, one of the four comparison tones was heard for 250 ms. The timing of this comparison tone was the principal (and critical) temporal manipulation, with the comparison tone occurring early, on-time, or late, relative to the temporal framework established by the context. Because the context had one beat every 500 ms (the time interval corresponding to the metronome clicks), an on-time comparison tone occurred exactly 500 ms after the final metronome click. In contrast, an early comparison tone occurred 437 ms after the final metronome click (63 ms early), whereas a late comparison tone occurred 563 ms after the final metronome click (63 ms late). These ± 63 ms shifts represent a 12.6% deviation from the on-time comparison tone. Jones et al. (2002) found that deviations of this magnitude (12.7%) showed the strongest effect of violations of temporal expectancy.

Crossing three standard tones (G-, G+, A-) with two comparison tones each (higher, lower) with three temporal positions (early, on-time or late) and four tonal

---

**Figure 9.** Schematic of a trial.
contexts resulted in 72 unique trials. All stimuli were created using a PC computer, Finale and SONAR software, and used a harmonically complex piano timbre. All stimuli were exported to .wav files with a sampling frequency of 44.1 KHz. The loudness of the stimuli was set to a comfortable listening level.

Apparatus

The apparatus were the same as in the previous experiments.

Procedure

Participants completed a questionnaire about their musical experience prior to beginning the experiment. At the start of each trial, listeners heard the experimental stimulus and then responded according to the task instructions. For the pitch height comparison, listeners were instructed to ignore the intervening context and to judge whether the comparison tone was higher or lower in pitch than the standard tone. For the key membership judgment, listeners were instructed to ignore the standard tone and to judge whether the comparison tone belonged in the key of the context. Participants used the “a” and “;” keys to enter their responses, with the assignment of keys (higher/lower, in/out of key) counterbalanced across listeners. Within a given block, order of trials was randomized for each listener. There were two blocks of trials for each task, with half of the listeners performing the pitch height comparison first followed by the key membership judgment, and the remaining listeners receiving the blocks in reverse order. Altogether there were 288 trials, which took about one hour.

Results

The principal analysis involved comparisons of accuracy (percent correct) as a function of task and temporal position. Percent correct for each participant was collapsed
across the three standard tones (G-, G+, A-) and their associated responses (higher/lower, in-key/out-of-key). Average percent correct was then analyzed with a five-way mixed model Analysis of Variance (ANOVA), with the within-subjects factors of Task (pitch height comparison versus key membership judgment), Temporal Position (early, on-time or late), Block (first versus second), and Context (four melodies, one for each context; see Figures 2 and 5). There was also a between-subjects factor of Task Order (which task the participants completed first). There were no significant main effects nor interactions. Most importantly, there was no effect of temporal position, $F(2,20) < 1$, $MSE = .01$, nor was there any interaction between task and temporal position, $F(2,20) < 1$, $MSE = .01$. Figure 10 displays percent correct as a function of task and temporal position.
The absence of temporal position effects on pitch judgments stands in stark contrast to reports by Jones and her associates (Jones et al., 2006; Jones et al., 2002) that...
violating temporal expectations (early or late) has adverse effects on pitch judgments. The pitch height comparison of the current experiment was similar to the task used by Jones et al. (2002). Both studies used a typical recognition-memory task requiring listeners to compare the pitch height of the standard and comparison tones. For Jones et al. (2002), temporal variation affected pitch judgments, but it did not in the current experiment.

The most obvious difference between the present pitch height comparison task and Jones et al. (2002) involves the intervening pitches between standard and comparison tones. Jones et al. (2002) used random pitches, in contrast to the simple tonal context of the present experiment. The implication is that tonal structure decreased listeners’ sensitivity to temporal expectancy violations. This issue was explored in Experiments 8 and 9.

The inclusion of key membership judgments extended previous research by examining a different type of processing, one that involved more complex judgments than comparisons of pitch height. As with the pitch height comparisons, violations of temporal expectancy failed to influence judgments of key membership. Temporal expectancy violations had similar consequences in both tasks, offering no support for the suggestion that pitch-time interactions vary as a function of type of processing. Together, these findings imply that neither different task demands (recognition memory or rating techniques) nor the type of task (key membership judgment versus pitch height comparison) were responsible for the divergent findings between Jones et al. (2006; Jones et al., 2002) and Experiments 1-6. Instead, the tonal structure of the context is likely to underlie the different outcomes.
Unfortunately, there was no manipulation of the presence or absence of tonal structure in the previous experiments. At present, evidence for the impact of tonal structure on pitch-time relations is indirect, arising largely from the null results obtained in the present experiment (i.e., no effect of temporal variation on pitch judgments). This interpretation would be more compelling if differences emerged from manipulations of tonal structure. If tonal structure reduces the influence of temporal variation on pitch judgments, then intervening tones that lack tonal structure should yield clear effects of temporal expectancy violations. Specifically, when the intervening contexts are atonal, accuracy of pitch judgments should be greater for pitches that are on-time rather than early or late. In contrast, when the intervening context has tonal information, then temporal expectancy violations should reduce or eliminate the influence of timing on pitch judgments. These predictions were examined in Experiment 8.
EXPERIMENT 8: THE ROLE OF CONTEXT IN PITCH-TIME DIMENSIONAL
RELATIONS: PART I

Experiment 7 revealed that temporal expectancy violations did not affect accuracy on pitch height comparisons or key membership judgments when the context involved tonal pitch structure. The present experiment used atonal as well as tonal contexts to examine whether the presence or absence of tonal structure affected the sensitivity of pitch judgments to temporal expectancy violations. The task was restricted pitch height comparisons, as in Jones et al. (2002).

Method

Participants

Twelve adults ($M = 19.1$ years, $SD = .9$) with at least eight years of formal musical training participated in this experiment. They received credit in an introductory psychology class at the University of Toronto Mississauga or payment ($10). Average musical training was $9.5$ years ($SD = 1.4$). Half of the participants remained musically active, and none reported possessing absolute pitch.

Stimuli, Apparatus, and Procedure

All stimuli were created with the same equipment as in Experiment 7. The format of the stimuli in this experiment was analogous to that of Experiment 7 — a standard tone, followed by a context, followed by a comparison tone. In fact, half of the trials used the same stimuli from Experiment 7. Because the context for these trials consisted of a simple tonal sequence, these stimuli were designated as tonal trials.

For the remaining trials, the intervening context between the standard and comparison tones consisted of random pitches that failed to establish a tonality. These
trials, designated the atonal stimuli, had four possible standard tones (F#\textsubscript{4}, G\textsubscript{4}, G#\textsubscript{4}, and A\textsubscript{4}) and six possible comparison tones (F\textsubscript{4}, F#\textsubscript{4}, G\textsubscript{4}, G#\textsubscript{4}, A\textsubscript{4}, A#\textsubscript{4}), all drawn from the equal-tempered Western scale. As with the tonal trials, each standard tone was associated with two comparison tones, one lower in pitch and one higher in pitch (see Table 5).

**Table 5**

Frequency of standard and comparison tones, and correct pitch height judgment for Experiments 8 and 9.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Standard Pitch</th>
<th>Comparison Pitch</th>
<th>Pitch Height Judgment</th>
</tr>
</thead>
<tbody>
<tr>
<td>466</td>
<td>A#\textsubscript{4}</td>
<td>A\textsubscript{4}</td>
<td>Higher</td>
</tr>
<tr>
<td>440</td>
<td>A\textsubscript{4}</td>
<td>A\textsubscript{4}</td>
<td>Higher</td>
</tr>
<tr>
<td>415</td>
<td>G#\textsubscript{4}</td>
<td>G#\textsubscript{4}</td>
<td>Higher/Lower</td>
</tr>
<tr>
<td>392</td>
<td>G\textsubscript{4}</td>
<td>G\textsubscript{4}</td>
<td>Higher/Lower</td>
</tr>
<tr>
<td>370</td>
<td>F#\textsubscript{4}</td>
<td>F#\textsubscript{4}</td>
<td>Lower</td>
</tr>
<tr>
<td>349</td>
<td>F\textsubscript{4}</td>
<td>F\textsubscript{4}</td>
<td>Lower</td>
</tr>
</tbody>
</table>

Context sequences in the atonal trials consisted of eight 250 ms tones, one every 500 ms, each presented simultaneously with a metronome click (the same metronome pulse in the tonal trials). The pitches in the context consisted of randomly chosen pitches (without replacement) ranging from E\textsubscript{4} (330 Hz) to B\textsubscript{4} (484 Hz) in quarter-tone increments (see Table 6). Accordingly, the atonal context consisted of a random set of 8 out of 15 unique pitches, and it did not include the standard or comparison tones for that specific trial. Crossing each of the 4 standard tones (F#\textsubscript{4}, G\textsubscript{4}, G#\textsubscript{4}, A\textsubscript{4}) with 2 comparison tones...
tones (lower, higher), with 3 temporal positions (early, on-time, late), and with 3 possible random contexts created a total of 72 trials per block. Figure 11 depicts a sample atonal context for a trial in which F#₄ was the standard tone and G₄ was the comparison tone.

Table 6

Frequency of all possible intervening pitches for Experiment 8.

<table>
<thead>
<tr>
<th>Frequency (Hz)</th>
<th>Pitch Names</th>
</tr>
</thead>
<tbody>
<tr>
<td>494</td>
<td>B₄</td>
</tr>
<tr>
<td>480</td>
<td>B-₄</td>
</tr>
<tr>
<td>466</td>
<td>A#₄</td>
</tr>
<tr>
<td>453</td>
<td>A+₄</td>
</tr>
<tr>
<td>440</td>
<td>A₄</td>
</tr>
<tr>
<td>428</td>
<td>A-₄</td>
</tr>
<tr>
<td>415</td>
<td>G#₄</td>
</tr>
<tr>
<td>404</td>
<td>G+₄</td>
</tr>
<tr>
<td>392</td>
<td>G₄</td>
</tr>
<tr>
<td>381</td>
<td>G-₄</td>
</tr>
<tr>
<td>370</td>
<td>F#₄</td>
</tr>
<tr>
<td>360</td>
<td>F+₄</td>
</tr>
<tr>
<td>349</td>
<td>F₄</td>
</tr>
<tr>
<td>339</td>
<td>F-₄</td>
</tr>
<tr>
<td>330</td>
<td>E₄</td>
</tr>
</tbody>
</table>
The apparatus and procedure were the same as those in Experiment 7. Listeners heard two blocks of randomly ordered tonal trials and two blocks of randomly ordered atonal trials. Half of the listeners received the tonal blocks first and the remaining listeners receiving the atonal blocks first. The entire procedure lasted approximately one hour.

Results

The principal analysis for this experiment used a four-way mixed model ANOVA with the within-subjects factors of Context Type (tonal versus atonal), Temporal Position (early, on-time, or late), and Block Repetition, (first versus second repetition). There was also a between-subjects factor of Context Order (whether participants did the tonal or atonal blocks first). For this analysis, all data were collapsed across the standard tones (G-, G+, A- for the tonal trials, and F#, G, G#, A for the atonal trials) and associated response type (lower, higher), and the context repetition (four different tonal contexts and 72 different atonal contexts). This analysis failed to reveal significant main effects for any of the main experimental factors. In fact, the only noteworthy result arising out of this analysis was a significant interaction between context type and temporal position,
$F(2,20) = 7.16, MSE = .002, p < .01, \eta^2_p = .42.$ Figure 12 depicts this interaction by showing pitch height accuracy as a function of temporal position for both tonal and atonal contexts.

Figure 12. Accuracy for tonal and atonal trials across temporal position in Experiment 8. Error bars represent standard error of the mean.
Two subsequent one-way ANOVAs compared pitch height judgments as a function of temporal position for each context type individually, with temporal position as the sole within-subjects factor. For the tonal contexts, there was no effect of temporal position, $F(2,22) = 1.02, MSE = .002, p = .38, \eta^2_p = .09$. For the atonal trials, however, there was a significant effect of temporal position, $F(2,22) = 4.33, MSE = .003, p < .05, \eta^2_p = .28$, with accuracy for on-time comparison tones exceeding those for early and late comparison tones. Supporting this pattern of results was a significant quadratic trend across temporal position, $F(1,11) = 3.53, MSE = .002, p < .05, \eta^2_p = .32$. This final analysis replicates the findings of Jones et al. (2006; Jones et al., 2002), who repeatedly observed more accurate pitch judgments for events occurring at predictable (i.e., expected) temporal locations than for events occurring at unexpected times (i.e., early, late).

**Discussion**

Overall, this experiment demonstrates that the temporal regularity of an event can influence judgments of relative pitch height, revealing a pitch-time interaction in musical processing. Nevertheless, this finding was qualified by whether the intervening context that induces temporal expectations also induces tonal expectations. If the intervening context was tonally structured, then violations of temporal expectancies did not influence pitch judgments. In contrast, if the intervening context was atonal, then the accuracy of pitch judgments was affected adversely by deviations from temporal regularity.

There are alternative explanations of these findings. The most notable of these involves the rhythmic structure of the tonal and atonal contexts. The tonal contexts consisted of rhythmically diverse events that varied in their relative durations, whereas
the atonal context consisted of isochronous events. Could the divergence in rhythmic structure have generated the observed difference in results between tonal and atonal conditions? If so, then the observed variation in rhythmic properties could affect listeners’ ability to induce a temporal structure. Weaker temporal structure would be less effective in generating temporal expectancies in listeners, resulting in weaker (or non-existent) effects of deviations from temporal regularity.

Although theoretically possible, this explanation assumes that the more rhythmically diverse and ecologically valid (tonal) context was less effective in inducing a temporal structure than the isochronous and less ecologically valid (atonal) context. This argument is inconsistent with theoretical analyses of rhythmic and metrical structure (Cooper & Meyer, 1960; Lerdahl & Jackendoff, 1983), formal modeling of the apprehension of metrical structure (Longuet-Higgins & Lee, 1982; Steedman, 1977; Temperley, 2001, 2008), and empirical results on the perception of metric information (Griffith & Todd, 1999; Palmer & Krumhansl, 1990; Povel, 1981). However, the difference in rhythmic diversity between the tonal and atonal contexts represents a difference with potentially important consequences.

A second difference is that there were 4 unique tonal contexts in contrast to 72 different sequences of random pitches. Repetition of the four tonal contexts may have increased the salience of melodic structure at the expense of temporal structure and expectancies. However, one can argue the converse, that repetition and predictability might make such patterns easier to ignore, enhancing the effect of the temporal structure and temporal expectancies in the tonal contexts. Alas, the results do not follow from the latter line of reasoning, but they are logically possible.
A third alternative explanation is that monophonic (involving a single melodic line) versus homophonic (involving a melody line with an underlying sequence of chords) contexts could differ in the extent to which they establish temporal structure or generate pitch-time interactions. In this case, the monophonic texture of the atonal contexts may have enabled stronger pitch-time interactions than the homophonic texture of the tonal contexts. However, the available research on pitch-time relations with monophonic and homophonic stimuli has not revealed systematic differences in their likelihood of engendering independent versus interactive processing (Palmer & Krumhansl, 1987a, 1987b).

Differences in the size of pitch intervals of the tonal and atonal contexts may have contributed to different outcomes. The randomly ordered sequences of the atonal contexts had large as well as small pitch intervals, in contrast to the tonal contexts, which had small pitch intervals that are characteristic of coherent melody lines. Again, there is no theoretical or empirical basis for arguing that this difference would alter the strength of temporal expectancies, but it remains a potential influence.

A final difference between the atonal and tonal contexts concerned the contrasting pitch sets: 7 pitch classes in tonal contexts and 15 separate pitches in atonal contexts (although only 8 were heard in a given context), including quarter tones. Again, there is no reason to expect the number of pitch classes to affect the consequences of temporal expectancy violations.

Although the aforementioned alternatives do not provide compelling explanations of the current data, they cannot be ruled out definitively on the basis on this experiment. Replicating the findings while controlling these factors was the goal of Experiment 9.
EXPERIMENT 9: THE ROLE OF CONTEXT IN PITCH-TIME DIMENSIONAL RELATIONS: PART II

The findings of the previous two experiments imply that the presence or absence of tonal pitch structure determined the existence of a temporal expectancy profile in pitch height comparisons. It is necessary, however, to consider several stimulus differences between tonal and atonal contexts that could have influenced the results. To investigate the role of these factors, the contexts in the present experiment consisted of random sequences of isochronous pitches, with the same number of distinct pitch classes that either established a key (tonal context) or did not (atonal context). As before, listeners had to judge relative pitch height. Their accuracy was evaluated as a function of the temporal expectancy of the comparison tone.

Method

Participants

Twelve adults ($M = 19.8$ years) who had at least 8 years of formal musical training ($M = 9.8$, $SD = 2.4$) participated in this experiment. Most of the participants ($n = 11$) remained musically active ($M = 3.8$ hours weekly), and none reported possessing absolute pitch. Participants received either credit in an introductory psychology class at the University of Toronto Mississauga or payment ($10$).

Stimuli, Apparatus, and Procedure

The stimuli were created with the same equipment as in previous experiments, and had the same format: a relative pitch height judgment of a standard and comparison tone separated by an intervening context. The standard and comparison tones were the same as in the atonal trials of Experiment 8 (see Table 5). In all trials, the standard tone
sounded for 250 ms, followed by a silent interval of 750 ms, and by an intervening context of 6 s consisting of 12 tones, each of which sounded for 500 ms. A metronome click accompanied the context and continued for 2 s after the context.

The trials had entirely new intervening contexts between standard and comparison tones, for both tonal and atonal conditions. The tonal contexts distributed the 12 isochronous tones among 7 diatonic pitch classes, such that there were 3 occurrences of the first and fifth scale degree, 2 occurrences of the third scale degree, and one occurrence of the remaining scale degrees (second, fourth, six, seventh). This distribution strongly correlated with the major tonal hierarchy (Krumhansl, 1990; Krumhansl & Kessler, 1982), \( r(10) = .96 \). The atonal contexts also distributed the 12 context tones between 7 pitches (with 2 pitches having 3 occurrences each, and one pitch having 2 occurrences), but in a pattern that did not correlate with the tonal hierarchy. Transposing this “atonal hierarchy” through all 12 possible iterations and correlating each with the major and minor tonal hierarchies produced unanimously low correlation coefficients (\(-.21 < r(10) < .21\) and \(-.21 < r(10) < .3\), respectively). In other words, the tonal contexts corresponded strongly to a major tonal hierarchy, whereas the atonal contexts did not correspond well to any major or minor key. Because tones were ordered randomly for tonal and atonal contexts, each contained a wide variety of pitch intervals between notes. Accordingly, the only recognizable distinction between the tonal and atonal contexts was the presence versus absence of tonal structure, respectively. All other factors, as noted, were equivalent.

As in the previous experiments, the manipulation of temporal position consisted of changing the timing of the comparison tone. Unlike the previous experiments, there
were five possible timings of the comparison tone, corresponding to very early (-126 ms), early (-63 ms), on-time, late (+63 ms) or very late (+126 ms). These timings represent 25.2% and 12.6% deviations from the on-time (i.e., expected) temporal position. The very early and very late temporal positions were introduced to expand the variety of timings to include larger than usual violations of temporal expectancies (i.e., beyond those typically tested by Jones and colleagues) and to evaluate their effects on accuracy.

There were 3 repetitions of each of the 4 standard tones (each of which was associated with 2 comparison tones), 5 timings and 2 contexts (tonal or atonal), resulting in 240 unique conditions. The apparatus and procedure were the same as in Experiment 8. Listeners were asked to judge whether the comparison tone was lower or higher in pitch than the standard tone. They were told to ignore the intervening context. Listeners heard one block of randomly ordered tonal trials, and a second block of randomly ordered atonal trials. Half of the listeners heard the tonal block first and the remaining listeners heard the atonal block first. The entire procedure lasted approximately one hour.

Results

Accuracy data were analyzed using a three-way mixed model ANOVA. The within-subjects factors were Context Type (tonal versus atonal) and Temporal Position (very early, early, on-time, late or very late). The between-subjects factor was Context Order (whether participants did the tonal or atonal blocks first). For this analysis all data were collapsed across the four standard tone pitches, their matching responses (lower, higher), and the three different repetitions. There was a main effect of temporal position, $F(4,40) = 3.58, MSE = .01, p < .05, \eta_p^2 = .26$. The only other significant effect was the interaction between context type and timing, $F(4,40) = 5.15, MSE = .01, p < .01, \eta_p^2 =$
This interaction was explored in two further one-way ANOVAs, with temporal position as the within-subjects factor.

For the tonal contexts, there was a main effect of timing, $F(4,44) = 2.88, MSE = .01, p < .05, \eta^2_p = .21$. However, this effect did not correspond to a conventional temporal expectancy profile. Instead, the early temporal position had the highest accuracy, followed by the very late and very early temporal positions. Four pairwise $t$-tests revealed that the on-time temporal position did not differ significantly from any of the other four temporal positions, $t(11) = 1.1, p = .29$ (very early – on-time), $t(11) = 2.1, p = .06$ (early – on-time), $t(11) = 1, p = .36$ (on-time – late), $t(11) = .95, p = .36$ (on-time – very late). Instead, there was a significant cubic trend, $F(1,11) = 5.6, MSE = .01, p < .05, \eta^2_p = .34$; Figure 13 displays these data.
The atonal contexts exhibited an effect of timing, $F(4,44) = 7.4$, $MSE = .004$, $p < .001$, $\eta_p^2 = .4$, but this effect also followed an unexpected pattern. There was a significant fourth order trend, $F(1,11) = 28.4$, $MSE = .004$, $p < .001$, $\eta_p^2 = .72$, with the on-time temporal position revealing better performance than early and late positions, but with the
very early and very late temporal positions revealing better performance than the early and late positions. Figure 13 depicts these data. Another four pairwise $t$-tests of temporal positions (the same as for the tonal contexts) revealed significantly higher accuracy at on-time positions than at early and late positions, $t(11) = -4.49, p < .001$ (early – on-time), $t(11) = 4.87, p < .001$ (on-time – late). However, accuracy at on-time positions did not exceed accuracy of very early and very late positions, $t(11) = -2, p = .07$ (very early – on-time), $t(11) = 1.68, p = .12$ (on-time – very late).

Discussion

At first glance, the effect of temporal position seems contrary to a typical temporal expectancy profile in which accuracy decreases as the timing of the events diverges from the expected temporal position. For the tonal contexts, the pattern of accuracy across temporal positions did not replicate the temporal expectancy profile found by Jones and colleagues. In fact, the significant effect of temporal position on accuracy for the tonal contexts is not readily interpretable. It may be due to the unexpected and unexplainable finding of higher accuracy on early comparison tones than on late comparison tones.

As for the atonal contexts, the results confirm the findings of Jones and colleagues (Jones et al., 2006; Jones et al., 2002) of increased accuracy for on-time comparison tones relative to early and late comparisons. In line with the findings of Experiment 8, this pattern appeared for comparison tones that followed atonal sequences but not tonal sequences.

The results for the atonal contexts also extend previous findings in a manner that is consistent with the dynamic attending hypothesis. Consider the underlying assumption
of the theory of dynamic attending. Attentional energy vacillates via linked oscillators that are synchronized with the rhythm of the presented stimulus (Large & Jones, 1999). On the basis of computational modeling of the perception of temporal patterns in music, Large and Palmer (2002) suggest that some of these linked oscillators function at subdivisions of the tactus beat level, according to the metric hierarchy. Therefore, the strength of temporal expectancies will vary as a function of the metric hierarchy. In the present experiment, the very early and very late comparison tones corresponded to a 25.2% shift from the on-time temporal position, an amount that corresponds to a quarter of the inter-onset interval (IOI) of 500 ms, whereas the 12.6% shift translates to an eighth of the IOI. In terms of the metric hierarchy (Palmer & Krumhansl, 1990), a quarter subdivision (very early and very late) of the tactus beat (500 ms IOI) is a more stable temporal position than an eighth subdivision (early and late). Because very early and very late comparison tones are more metrically stable than the early and late comparison tones, note occurrences at those locations would violate temporal expectancies to a lesser degree. Thus, the observed improvement in accuracy with the more extreme temporal shifts aligns well with dynamic attending theory.

By replicating the finding of a temporal expectancy profile in atonal but not tonal contexts, this experiment rules out the importance of potentially problematic factors in the stimuli that were noted in Experiment 8. Specifically, the present experiment used tonal and atonal contexts with the same rhythmic structure (isochronous pitches) and the same number of unique sequences (120). Furthermore, all of the context sequences were monophonic, consisting of seven unique pitch classes and randomly selected tones that resulted in a mixture of large and small adjacent pitch intervals. Accordingly, the only
remaining candidate to account for the differential impact of tonal and atonal contexts is the presence or absence of tonal pitch structure.
GENERAL DISCUSSION

The first six experiments explored the relation between musical pitch and time in the perception of a typical Western musical passage. Experiments 1 and 2 collected ratings of goodness of fit for probe events that combined different pitch classes and temporal positions. Experiments 3 and 4 used speeded classification of either pitch class or temporal position, accompanied by variation along the other (irrelevant) dimension. Experiments 5 and 6 ruled out several alternative explanations of the results of Experiments 3 and 4, providing baseline measures and replicating the primary findings of these studies with stimuli that controlled for an array of extraneous factors. The effects of dimensional structure (tonality and meter) on performance were evident in all experiments, but there were important variations in how these dimensional structures combined. Specifically, tonality and meter combined additively in goodness-of-fit ratings (Experiment 1), even when participants were instructed to ignore pitch (Experiment 2). In a metric speeded classification task (Experiments 3 and 6), the tonal hierarchy biased responses, producing a congruity effect between the pitch and time dimensions, with this effect attributable to differential response bias but not discriminability. In contrast, in a pitch speeded classification task (Experiments 4 and 6), the metric hierarchy failed to influence responses, with no corresponding impact on either response bias or discriminability.

Subsequent experiments (7-9) tested the circumstances under which temporal variations affected the accuracy of a pitch judgment. In Experiment 7, the accuracy of key membership and relative pitch height judgments of a comparison tone following a context that induced both tonal and metric hierarchies was not influenced by violations of
temporal expectancies of that comparison tone. Experiment 8 extended this finding by manipulating the presence or absence of tonal structure in the intervening context with the pitch height comparison task. This experiment obtained an effect of temporal expectancies on pitch judgments, but only when the intervening context lacked tonal structure. The pattern of findings replicated previously reported temporal influences on pitch perception (Jones et al., 2006; Jones et al., 2002), with listeners processing temporally expected events better than temporally unexpected events. Experiment 9 replicated the findings of Experiment 8 while controlling for several differences between the tonal and atonal stimuli of Experiment 8. Experiment 9 also revealed enhanced accuracy on very early and very late comparison tones relative to smaller temporal shifts in atonal trials. Although unexpected, this finding is consistent with the theory of dynamic attending because the more extreme temporal positions were more metrically stable than the less extreme temporal positions. Regardless, it appears the presence or absence of a tonal framework rather than the type of processing (key membership judgments or pitch height comparisons) made the major contribution to pitch-time interactions in the form of temporal expectancies.

The findings of these experiments are intriguing, and they have important implications. The congruity effect in Experiment 3 is one such finding, in which responses were more accurate to tonally stable events when they appeared in metrically stable positions. Not only did pitch affect temporal judgments, but the hierarchical structure (tonality) of the irrelevant dimension also affected the nature of this interference. This finding converges with research of Lebrun-Guillaud and Tillman (2007), who found that the tonal function of a note influenced the detection of changes in
the temporal regularity of a sequence. Their listeners, who heard three chords followed by three notes, all in isochronous rhythm, were asked to detect temporal deviations on the fifth event. Performance improved with increasing tonal stability of the target note, an effect congruent with results of the present research. By contrast, however, Lebrun-Guillaud and Tillman (2007) found effects on discriminability ($d'$) and response bias ($c$), whereas the effect was limited to response bias in the present research. Accordingly, the experiments suggest later, decisional origins of pitch interference in contrast to pitch interference at both perceptual and decisional levels for Lebrun-Guillaud and Tillman (2007).

The findings of Experiments 7-9 shed light on the divergent results reported by Jones et al. (2006; Jones et al., 2002) and those reported in Experiments 1-6. Based on the theory of dynamic attending, Jones et al. (2006; Jones et al., 2002) argue that temporal and pitch pattern structure drive listeners’ attention by producing dynamic temporal and pitch expectancies for when (in temporal space) and where (in pitch space) subsequent events will occur. This model assumes an inherent interaction between the processing of pitch and temporal information, with deviations in one dimension influencing the apprehension of information in the other dimension. In contrast, Experiments 3, 4 and 6 found an asymmetry in pitch-time relations, with pitch variation influencing temporal processing, but temporal variation failing to influence pitch processing for Western tonal music. However, the presence versus absence of a tonal framework determined the role of temporal expectancies in Experiments 8-9, suggesting that the pitch structure of tonality automatically shifts listeners’ attention toward the dimension of pitch at the expense of time. Conceived more generally, it is likely that the musicality of the task
influences participants’ listening strategy. By virtue of exhibiting typical Western pitch structure, the stimuli of Experiments 1-8 are more musical, and therefore may invoke listening strategies that involve favoring pitch over time. In comparison, the isochronous and atonal stimuli of Experiment 9, and Jones’ work (Jones et al., 2006; Jones et al., 2002) do not reflect typical Western conventions, and as a result are less musical. As the stimuli and task become less musical, they are more likely to prompt listeners to use different listening strategies that do not resemble those used when listening to music.

Is interpretation of the present findings limited somewhat by the fact that listeners had 8 or more years of musical training, in contrast to participants in Jones et al. (2002), who had less than 6 years of musical training? It is possible that musicians’ greater understanding of music (in particular, tonal structure) enhanced their focus on pitch. It is also possible that listeners’ musical skill allowed them to maintain high levels of attention to all events, not only those occurring on the beat. In terms of dynamic attending theory (Barnes & Jones, 2000; Jones, 1987; Jones & Boltz, 1989; Large & Jones, 1999), musicians may have wider peaks of attentional energy oscillations that are synchronized with the beat. Untrained listeners, by contrast, may have difficulty processing events that do not occur at expected temporal location. If this explanation is correct, then nonmusicians tested on the same tasks could show a somewhat different pattern of findings than those reported here. However, given the similarity of musical processing across levels of formal training (Bigand & Poulin-Charronnat, 2006; Koelsch et al., 2000), it is more likely that preferential attention to pitch arises primarily from passive exposure to music. Musicians could still have broader peaks of temporal attention, like those suggested above, which could lead to differences between musicians
and nonmusicians on some tasks. The available evidence suggests, however, that these differences are quantitative (e.g., level of accuracy) rather than qualitative (Bigand & Poulin-Charronnat, 2006).

How do the present findings fit with other notions in the music cognition literature on the relation between musical pitch and time? Although the data from the present experiments cannot specify how musical pitch and time integrate in every task, the results implicate the relative salience of perceptual dimensions. The salience of pitch structure may play a critical role in a wide variety of findings in the music cognition literature. For example, Thompson (1994) tested listeners’ ability to detect a change in the pitch or duration of a repeated two-note musical sequence. In the “switch” condition, the pitch of the first note was combined with the duration of the second, and the pitch of the second note now fit with the duration of the first. In this case, accurate performance depended on sensitivity to combinations of pitch and duration, thus requiring integration of the pitch and temporal information. Increasing the pitch difference between the two notes decreased this sensitivity, perhaps because the improved discriminability of the pitch dimension diverted attention from its combination with a specific durational value. However, making the durational values more distinctive did not interfere with integration of the two dimensions, which means that pitch still attracted a high level of attention.

Because the present experiments were not designed explicitly to test them, the present findings have limited applicability to existing theories on pitch-time integration in music, including distinctions between local versus global processing (Tillmann & Lebrun-Guillaud, 2006), early versus late processing (Hébert & Peretz, 1997; Peretz & Kolinsky, 1993; Pitt & Monahan, 1987; Thompson et al., 2001), and coherence (Boltz,
However, there are some implications of the present findings for these theories that are worth mentioning. The global-local theory posits that pitch and time are processed independently for local processing and more interactively for global tasks. The local focus (i.e., the processing of a single probe event) of the present experiments on pitch and time contrasts with investigations of more global musical features such as melody and rhythm. There are few explicit tests of the local-global distinction with regard to pitch-time processing. Tillman and Lebrun-Guillaud (2006) found that when listeners classified the timbre of a chord (a local task), there was no interaction between temporal dimensions (ending time) and pitch dimensions (harmonic relatedness). In contrast, they found an interaction between temporal and pitch dimensions for global judgments, specifically, how well the chord completed a musical sequence. In the present study, the lack of mutual interference between pitch and time may stem from tasks that involved local processing. Although this global/local distinction may account for the absence of time effects on local pitch judgments, it makes the pervasive effect of pitch on local temporal judgments seem all the more remarkable. If pitch affected time on tasks that required local processing, as in the present study, one would expect greater effects on tasks that necessitate global processing.

It is notable that the present study did not provide an explicit test of stage theories of pitch-time integration, in which pitch and time function independently at earlier stages of processing and combine at later stages (Hébert & Peretz, 1997; Peretz & Kolinsky, 1993; Pitt & Monahan, 1987; Thompson et al., 2001; Tillmann & Lebrun-Guillaud, 2006). Specifically, the present experiments used high-level judgments (goodness of fit, location in the tonal or metric hierarchy), which presumably involved “late” processing.
of musical events. Even the pitch height comparison, which seems like a relatively lower-level judgment, required higher-level processing to remember the pitch of the standard tone, ignore the intervening pitch sequence, and evaluate the relative pitch height of the comparison tone. Moreover, the presence or absence of tonal pitch structure was the factor that modulated the pattern of pitch-time integration in Experiments 8 and 9. Classification judgments produced a joint contribution of pitch and time, whereas goodness-of-fit judgments did not (both presumably late processing tasks), which cannot be reconciled on the basis of early versus late processing. In short, this distinction cannot explain the observed variation in pitch-time relations in the present study.

Boltz (1998a, 1999) proposed that the independence or interaction between musical pitch and time varies as a function of melodic coherence, with coherence referring to the well-formedness of pitch and temporal structures in a musical passage. According to Boltz, the extent of listeners’ experience with a melody also affects its coherence (more experience leads to greater coherence). Although the present study did not manipulate melodic coherence in the context passages (as done by Deutsch, 1980), it can shed light on the impact of exposure to the sequences on the interaction of musical pitch and time. Specifically, Experiments 1 and 2 presented three repetitions of the same combinations of stimulus context and probe event, providing an experiential manipulation. As noted, repeated blocks failed to influence ratings of probe events in either experiment, providing little evidence to support the notion that the increased experience with a musical passage increased pitch-time interactions.

Overall, the most parsimonious explanation of the present findings is that tonality attracts attention to the pitch dimension at the expense of time. The asymmetry of
interference between pitch and time in Experiments 1-6 implies that pitch information was more salient to listeners in the typical musical contexts of these experiments. As a result, listeners were unable to ignore pitch even when it was irrelevant to the task.

Indeed, Experiments 7-9 provide converging evidence for this notion. The asymmetric results of Experiments 1-6 in which pitch variation influenced temporal classifications in tonal music (but not vice versa) are consistent with this hypothesis of pitch salience. Those experiments provided only partial support for a pitch-focus hypothesis. The subsequent experiments provided more direct evidence. Specifically, manipulation of the tonality of the stimuli affected the focus on pitch information, as indicated by corresponding variation in the influence of temporal expectancies on pitch judgments.

Overall, the present research reveals pitch influences on temporal judgments along with temporal influences on pitch judgments.

There is overwhelming acknowledgement of the importance of pitch structure in music perception. Undoubtedly, temporal regularity is also important for the processing of auditory events, as revealed in the present study and many others (Garner, 1974; Handel & Lawson, 1983; Jones, 1976; Jones, Kidd, & Wetzel, 1981; Longuet-Higgins & Lee, 1982; Steedman, 1977). In this regard, the current results can be reconciled with Jones’ theory of dynamic attention (Boltz, 1993a; Jones & Boltz, 1989; Jones et al., 2006; Jones et al., 2002; Large & Jones, 1999). Moreover, the present findings provide an important qualification to this theory by revealing task characteristics that modulate the role of temporal factors in music perception. The present research can be considered to extend existing work on dynamic attending by elaborating the role of pitch structure in temporal expectancy profiles.
The current findings do not suggest that the critical factor in pitch-time relations is the presence or absence of any pitch structure in a musical context. Rather, these results suggest that the nature of the pitch structure is crucial. In Jones et al. (2006), the sequences contained hierarchical pitch structure, although this structure was along the lines of classic serial pattern structure (Deutsch & Feroe, 1981; Jones, 1987; Restle, 1970, 1972; Simon, 1972). For Jones et al. (2006), the serial pattern structure of the sequence, relative to the pattern in which the target tone was embedded, influenced change detection of the target. Most importantly, these authors found that the temporal predictability of the target tone simultaneously influenced detection of a pitch change in a target sequence. It would be a misrepresentation to suggest, then, that there is something special about pitch structure per se that underlies the divergence between the current findings and Jones’ work.

This point leads to two related questions. First, what is it about tonal structure that distinguishes it from other forms of pitch structure (such as serial pattern structure)? Second, what other forms of pitch structure might have comparable effects on pitch-time interactions? With reference to the first issue, no definitive answer is available, but it is likely that the amount of listeners’ experience with tonal structure is relevant. Listeners are sensitive to pitch structure and can learn other pitch structures, but tonality is by far the most prevalent pitch structure in Western music. Perhaps a similar level of exposure to some other form of pitch structure might have similar effects. Another possibility is that tonality consists of a multi-leveled hierarchical structure of pitch events, with varying degrees of relatedness between the pitches of the chromatic set and the induced tonal center. In contrast, although serial patterns can contain multi-leveled hierarchical
structure (Deutsch & Feroe, 1981), the actual sequences used by Jones et al. (2006) were not multi-leveled, consisting instead of a much simpler hierarchical serial pattern. If this explanation is correct, then the use of intervening sequences with a richer hierarchical serial pattern could reduce the impact of temporal expectancy violations on pitch height judgments, in line with the findings of Experiments 7-9.

This prediction could provide a partial answer to the second question as well – whether forms of pitch structure other than tonality might have a similar effect on pitch-time interactions. In fact, there are various possibilities along these lines. In addition to using more complex hierarchical serial patterns, another possibility involves examining the impact of structural aspects such as the relations between notes in the intervening context and the standard and comparison tones. Indeed, early work found that recognition memory of a standard tone followed by an intervening sequence and a subsequent comparison tone was influenced by whether or not the intervening sequence repeated the standard tone (Deutsch, 1972, 1975). More generally, evidence that including such a critical tone in the context influences pitch height comparisons and the role of temporal expectancies would provide compelling evidence for the idea that the pitch structure of musical contexts influences pitch-time interactions.

An explanation of the present results based on pitch dominance begs the question of why pitch was more salient in these experiments. One possibility is that listeners’ greater focus on pitch information arises from their exposure to Western music. In most genres of Western tonal music, the complexity of pitch structures dwarfs the complexity of temporal structures. For example, Western music is overwhelmingly in binary or ternary meters and typically involves a handful of different notated duration values. In
contrast, the pitch structures in Western music use many different pitch sets in melody and harmony. Indeed, in Western music in general, the possible unique combinations of pitch are greater than those of duration. There is comparatively little variability in duration, often using only two or three duration values (e.g., quarter and eighth notes). In contrast, the same music typically contains considerably more variability and less predictability in pitch, typically with five to seven pitches of the scale represented. More importantly, the possible set of pitches that can define a tonal context is larger than the set of durations that (when organized properly) can create a metric framework. Therefore, any set of pitches will tend to have a lower probability of occurrence than any set of durations. Events with a low probability of occurrence should require more processing resources than events with a high probability of occurrence. Put differently, events with a high probability of occurrence should be more predictable and therefore easier to process. Over years of exposure to typical Western tonal music that generally contains more elaborated and compelling pitch variation relative to rhythmic/temporal variation, listeners come to allocate more mental effort to pitch than to temporal structures. Thus, this asymmetry in experience is likely to result in an involuntary bias to process pitch preferentially in typical Western music, making pitch more salient in the perception of music.

There is an analogous situation for picture-word interference. Low-frequency words are more distracting than high-frequency words during picture naming (Miozzo & Caramazza, 2003). Because low-frequency words are more distinctive, inhibiting their lexical entry takes longer, with adverse effects on picture naming. Similarly, greater distinctiveness of the pitch dimension relative to the timing dimension results in greater
attention to the pitch dimension and greater difficulty ignoring variations along this
dimension. This line of reasoning would predict a reverse pattern of asymmetric
interference for listeners familiar with musical systems that involve more complex
temporal than pitch patterning (e.g., Australian Aboriginal, African, South Asian). For
example, listeners from musical cultures that have rich temporal structure are better at
detecting violations to complex temporal structures in music than are Western listeners
(Hannon & Trehub, 2005). As a result, they might be less able to ignore variations along
the temporal dimension. However, it is unlikely that pitch or time dominance is a
dichotomous phenomenon. Almost certainly, dimensional salience functions along a
continuum, such that the degree of asymmetric interference of one dimension will vary in
accordance with the relative amount of structure in each dimension.

Another possible account for these findings arises from previous demonstrations
of asymmetric interference in the absence of discriminability differences. Indeed, there
are precedents for similar effects in both speech (Mullennix & Pisoni, 1990; Tong,
Francis, & Gandour, 2008) and face perception (Atkinson, Tipples, Burt, & Young, 2005;
Haxby, Hoffman, & Gobbini, 2000; Schweinberger, Burton, & Kelly, 1999). In short,
face perception work proposes that invariant characteristics will display asymmetric
interference on changeable dimensions because an invariant characteristic acts as a better
referent (Haxby et al., 2000). In keeping with the current findings, some authors suggest
that the source of interference is not perceptual but occurs at a later, decisional stage
(Atkinson et al., 2005). These authors allow for enculturation, similar to the culturally
learned favoring of pitch in typical Western music. Haxby et al. (2000) even specify
distinct neurobiological pathways for different dimensions; neuropsychological work on
music perception offers a parallel proposal for pitch and time (Peretz & Coltheart, 2003; Peretz & Kolinsky, 1993; Peretz et al., 1994).

Although pitch is not necessarily an invariant characteristic, it may function as a better perceptual reference point than other dimensions (e.g., timing, timbre, etc.) when a task is strongly pitch-based, thus acting as an invariant attribute. In the tasks of the present study, a pitch defined the presence of a musical event regardless of when it occurred, whereas any given temporal position was not an event unless a pitch occurred then. Therefore the pitch of a probe event or comparison tone may have provided a better reference point than its temporal position; as a result participants had difficulty ignoring pitch even when it was irrelevant to the task. Consequently, pitch would interfere in the temporal task while time would not interfere with the pitch task. Conversely, in a tapping task (e.g., Snyder & Krumhansl, 2001), the temporal position (the timing of the tap) is the main attribute of interest, whereas a pitch need not be present. In the latter case, timing may function as a better referent and reverse the pattern of pitch-time dominance observed here. This concept resembles and extends the concept of physical primacy (Garner, 1974) discussed in the introduction.

Indeed, presenting a musical context (especially if tonal) followed by a pitched probe event in these experiments means that these tasks were strongly pitch-based. Other tasks that use different methods may observe alternate patterns of dimensional salience. For example, research paradigms in which participants tap in synchrony with a rhythmic sequence (and thus are inherently time-based) can show dominance of temporal factors over pitch (Pfordresher, 2003; Snyder & Krumhansl, 2001; Vos, Vandijk, & Schomaker, 1993). Perhaps a given dimension is more likely to interfere with another dimension if
the task focuses attention towards it. It is difficult to quantify the extent to which a task favors one dimension over another. Therefore, this question is an important area for future research on dimensional relations.

A more general implication of these findings for research on dimensional relations is that they reveal an important distinction between dimensional discriminability and salience. That is, the more salient dimension in a stimulus may or may not be more discriminable than a less salient dimension. As a result, asymmetric dimensional interactions may result from inequalities in discriminability (Garner, 1974; Garner & Felfoldy, 1970; Melara & Algom, 2003; Melara & Mounts, 1993; Sabri et al., 2001), or inequalities in salience. Therefore, investigations aimed at definitively establishing dimensional interaction versus independence should ensure not only that the dimensions are equally discriminable, but also that one dimension does not dominate the other in terms of its salience or importance to perceivers. Normally, one would expect a correlation between salience and discriminability, such that the more discriminable dimension is also more salient to perceivers. That is, our perceptual systems tend to maximize processing of the most biologically relevant stimuli (e.g., faces, motion, voices), which may increase the salience of their stimulus dimensions. At times, however, the most salient dimension is difficult to process because it is noisy, degraded, or masked. In this case the most salient dimension is not the most discriminable. The findings demonstrate that under appropriate conditions, dimensional discriminability and salience can be dissociable.

The current research provides insight into relations between the processing of musical pitch and time. It reveals how some aspects of the structure and hierarchical
organization of these dimensions affect their perceptual integration. It also differentiates the role of discriminability and salience in perceptual integration. The findings point the way towards fruitful investigations not only within the domain of music perception but also in visual and other auditory domains.
REFERENCES


*Brain, 117*, 1283-1301.


*Perception and Psychophysics, 41*(6), 534-546.


segmental and suprasegmental features in Mandarin Chinese. *Language and


23*, 965-976.