THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
i*First*, 1–28

Psychology Press
Taylor & Francis Group

The integration of stimulus dimensions in the perception of music

Jon B. Prince

School of Psychology, Murdoch University, Murdoch, WA, Australia

A central aim of cognitive psychology is to explain how we integrate stimulus dimensions into a unified percept, but how the dimensions of pitch and time combine in the perception of music remains a largely unresolved issue. The goal of this study was to test the effect of varying the degree of conformity to dimensional structure in pitch and time (specifically, tonality and metre) on goodness ratings and classifications of melodies. The pitches and durations of melodies were either presented in their original order, as a reordered sequence, or replaced with random elements. Musically trained and untrained participants (24 each) rated melodic goodness, attending selectively to the dimensions of pitch, time, or both. Also, 24 trained participants classified whether or not the melodies were tonal, metric, or both. Pitch and temporal manipulations always influenced responses, but participants successfully emphasized either dimension in accordance with instructions. Effects of pitch and time were mostly independent for selective attention conditions, but more interactive when evaluating both dimensions. When interactions occurred, the effect of either dimension increased as the other dimension conformed more to its original structure. Relative main effect sizes (| pitch η^2 − time η^2 |) predicted the strength of pitch–time interactions (pitch × time η^2); interactions were stronger when main effect sizes were more evenly matched. These results have implications for dimensional integration in several domains. Relative main effect size could serve as an indicator of dimensional salience, such that interactions are more likely when dimensions are equally salient.

Keywords: Pitch; Time; Music; Dimensions; Integration.

The cognitive mechanisms responsible for the perceptual integration of stimulus dimensions are complex and multifaceted. Nonetheless, perception of multidimensional objects in the natural environment proceeds automatically and without conscious effort. Understanding how these dimensions combine in order to form a unified and coherent mental percept is a primary goal of cognitive psychology (Treisman & Gelade, 1980). Therefore much research has concentrated on the central issue of how stimulus dimensions combine in perceptual processing.

Correspondence should be addressed to Jon B. Prince, School of Psychology, Murdoch University, Murdoch, WA 6150, Australia. E-mail: j.prince@murdoch.edu.au

Grants from the Natural Sciences and Engineering Council of Canada to Mark A. Schmuckler and William F. Thompson supported part of this research, in addition to grants from the National Science Foundation of the USA awarded to Peter Q. Pfordresher. Portions of this work were presented at the Auditory Perception, Cognition, and Action meeting, Boston, MA, November 2009. The author would like to thank Peter Q. Pfordresher for helpful comments and discussions.

DOI:10.1080/17470218.2011.573080

Seminal work by Garner (1974) introduced the concept of integral and separable dimensions, specifying that integral dimensions cannot be isolated in perceptual processing, whereas separable dimensions can be processed independently. When attending a stimulus comprised of integral dimensions, changes along one dimension will affect judgements along the other dimension; for separable dimensions no such interfering effect arises. However, the integrality or separability of dimensions is influenced by their relative discriminability, which corresponds to the baseline difficulty of perceiving a given dimension (Garner & Felfoldy, 1970) and can be measured using reaction times, accuracy, or d'. A more discriminable dimension will interfere with a less discriminable dimension, even if they have demonstrated separability under conditions of equal discriminability. Yet many dimensions defy a neat classification into integral and separable dimensions. For example, asymmetric interactions may occur even when dimensions have been matched for discriminability, such that one dimension interferes with another, but not vice versa (Garner, 1976).

Further research refined the original framework of integral and separable dimensions into a more detailed and nuanced view of dimensional integration. One such development was the idea of decisional and perceptual separability (Ashby & Townsend, 1986), which proposes that a dimension may interfere with another at either of these processing levels when responding to a multidimensional stimulus. Other authors have proposed the concepts of dimensional imbalance and dimensional uncertainty (Melara & Algom, 2003), theorizing that the amount of information present in a given dimension influences how the perceiver attends to the stimulus. Additionally, some stimulus dimensions appear to be privileged and can exert dominance (i.e., asymmetric interference) over other dimensions despite careful matching of other factors (Atkinson, Tipples, Burt, & Young, 2005; Haxby, Hoffman, & Gobbini, 2000; Mullennix & Pisoni, 1990; Prince, Thompson, & Schmuckler, 2009; Schweinberger, Burton, & Kelly, 1999; Tong, Francis, & Gandour, 2008).

Despite considerable advances in the field of multidimensional visual object perception, less is known about integration of multidimensional properties in patterned auditory stimuli such as speech and music. Rather than representing a static object, auditory sequences unfold over time. Accordingly, integrating the dimensions of such sequences has an additional layer of complexity by virtue of their dynamic nature. In the case of music, the main stimulus dimensions of pitch and time can be manipulated independently, and there is a rich background on these two dimensions from the perspectives of music theory (Aldwell & Schacter, 2002; Schenker, 1935/1979) and cognitive science (Jones & Boltz, 1989; Krumhansl, 1990). Nevertheless, there is still little agreement on how exactly pitch and time contribute to, and combine in, our internal representation of music (this process is referred to hereafter as pitch–time integration). Responses to any stimulus will depend on the internal representation of its features, and music is no exception to this principle. Thus, understanding how pitch and time integrate is vital to research on musical behaviour (listening, performing, and composing) and can contribute to the larger literature on dimensional integration in other sensory modalities.

Much research on pitch–time integration has taken a position supporting either the idea that pitch and time are independent, or that they are interactive (as reviewed in Ellis & Jones, 2009; Krumhansl, 2000; and Prince, Thompson, et al., 2009). Palmer and Krumhansl (1987a) provided one of the landmark demonstrations of independence, where an additive combination of pitch and temporal information predicted judgements of melodic phrase goodness, suggesting that pitch and time were processed separately. Other research proposed instead that pitch and temporal information is processed jointly, because variations in rhythmic structure influenced detection of perturbations to pitch patterns (Jones, Boltz, & Kidd, 1982). Since then, the numerous reports of independence and interaction of these dimensions have led to much confusion over the nature of pitch–time integration. It is more likely, therefore, that pitch and time function neither as purely independent nor as interactive dimensions, but that pitch–time integration varies on the basis of a number of stimulus, task, and cultural factors. In real-life situations, the characteristics of the music that a listener hears (i.e., the stimulus) may highlight one dimension, such as emphasizing the pitch dimension by using a repeating pattern, or the time dimension by accelerating the tempo. Similarly, listeners may choose to focus on different aspects of a musical sequence, such as listening to a particular instrument, a rhythmic figure, or the timbral features of the music. Based on what aspect is of interest, the task of the listener will involve isolating different dimensions of the music. Finally, listeners encultured in the Western musical style will probably hear music of another culture differently from a native listener (e.g., Castellano, Krumhansl, & Bharucha, 1984).

Recent work has examined the conditions under which musical pitch and time combine independently or interactively. Among the existing theories on variations in pitch–time integration patterns, the most well established are an early- versus late-stage processing model (Peretz & Kolinsky, 1993; Pitt & Monahan, 1987; Thompson, Hall, & Pressing, 2001) and local versus global task characteristics (Bigand, Madurell, Tillmann, & Pineau, 1999; Jones & Boltz, 1989; Tillmann & Lebrun-Guillaud, 2006).

Work on stages of processing proposes that pitch and time in music are initially processed independently and integrate at later stages (Peretz & Coltheart, 2003). Accordingly, dependent measures that rely on early processing (such as perceptual judgements of pitch height or tempo) are likely to demonstrate independence because the dimensions are still being processed separately. Conversely, interactions are more likely to be observed when testing later stages of processing (such as melodic completion or rhythmic similarity), when the dimensions have been integrated.

The distinction between local and global task characteristics also shows promise for reconciling the variety of findings for pitch–time integration. This account (Tillmann & Lebrun-Guillaud, 2006) contends that tasks that favour processing of individual (local) stimulus elements foster observations of independence; conversely, interactions are more common when attending to characteristics of a stimulus on a larger timescale (global tasks). Because attending to larger, more global structural relations will require combining a number of preceding events, such tasks are likely to require processing of the stimulus in an integrated form. This integrated internal representation would include both pitch and temporal information; thus these stimulus dimensions are more likely to interact in perceptual processing. Conversely, attending locally to an isolated event can proceed without reference to more global properties of a stimulus, not depending on a joint (integrated) representation, and therefore fostering independent relations. Although typical listening situations are likely to involve attending to the global characteristics of a musical excerpt (e.g., the contour of a melody, or the musical tension and relaxation of a sequence), there are scenarios in which the listener may focus locally on a single event, such as a particularly striking musical occurrence, or listening for a favourite chord in a familiar piece.

However, not all findings concur with these preceding theories. Prince, Thompson, et al. (2009) reported independent and interactive contributions of pitch and time when responding to different aspects of a probe event (an isolated single tone) following a melody. In one task, listeners rated how well the probe event fitted with the melody; pitch and time contributed to goodness ratings with additive (independent) contributions. In another task, listeners classified whether the probe event was “on the beat” (if it was consistent with the rhythmic framework of the melody), and the pitch of the probe event influenced their responses. However, in a homologous task where the listeners instead indicated whether the probe event used a pitch class that was important in the melody, the timing of the probe event did not affect their responses, thus demonstrating an asymmetric interaction between pitch and time. Both of these tasks required relatively late stages of processing and attention to global characteristics of the stimulus, yet responses indicated independence in some tasks and interaction in others.

Prince, Thompson, et al. (2009) proposed dimensional salience as a factor to reconcile the variety of findings in the literature. The concept of dimensional salience refers to the tendency for one perceptual dimension to dominate another, even when both dimensions are equally difficult to process (i.e., equated discriminability). Instead, interactions may be due to stimulus and task properties that magnify the importance of one dimension at the expense of another. Stimulus dimensions that feature greater informative value are likely to dominate (cf. Melara & Algom, 2003; Prince, Schmuckler, & Thompson, 2009). Also, task design may attract greater attention to (and thus enhance the salience of) a dimension; for example, tapping tasks may highlight temporal information at the expense of pitch (Pfordresher, 2003; Snyder & Krumhansl, 2001).

Typical Western music emphasizes pitch over time, thus enhancing pitch salience through the properties of the stimulus. Indeed, both music theorists and psychological researchers attest to the centrality of pitch in Western music. For example, according to Schenker (1935/1979 p. 15), "all rhythm in music comes from counterpoint and only from counterpoint", and Everett (2000 p. 111) states that "pitch relationships are of central importance, forming the core of the structure, the identity, and even many of the expressive capabilities of pop-rock music". Empirical work demonstrates that pitch is the fundamental organizing principal for musical memory (Dowling, 1978; Hébert & Peretz, 1997), recognition (Dowling & Fujitani, 1971; White, 1960), similarity judgements (Cousineau, Demany, & Pressnitzer, 2009; Eerola, Järvinen, Louhivuori, & Toiviainen, 2001), segmentation (Dawe, Platt, & Racine, 1993, 1994, 1995), perceived stability (Bigand, 1997), and the overall formation of a mental representation of music (Krumhansl, 2000).

But why is pitch so important in Western music? One of the reasons for the dominance of pitch in typical Western music may be its structural complexity. There are many forms of structure in Western music, loosely categorized as either surface-based structures or more abstract organizational principles. Some examples of surface-based pitch structure include the pattern of ups and downs in pitch height within a melody known as pitch contour (Dowling, 1978), serial patterns (Boltz, Marshburn, Jones, & Johnson, 1985), and grouping/stream segregation (Bregman, 1990). Serial patterns and grouping also exist in the temporal domain (Garner, 1974; Jones, 1987).

Perhaps the most pervasive abstract organizational principle of pitch in Western music is tonality, or musical key. Tonality refers to the hierarchical organization of the 12 pitch classes (per octave) used in Western music around a central reference pitch, or tonic, that represents the key of a musical excerpt (Aldwell & Schacter, 2002; Krumhansl, 1990; Lerdahl, 2001; Schmuckler, 2004). The remaining pitch classes vary in terms of their psychological stability (and frequency of occurrence), in a hierarchical profile commonly referred to as the tonal hierarchy (Krumhansl & Kessler, 1982). For example, in the key of C major, C is the tonic and most stable pitch. The pitches E and G (used in conjunction with C to form the tonic triad of C major) are the next most stable, followed by the 4 remaining pitch classes used in C major (D, F, A, B). The 5 pitches that are seldom used in that key ($C^{\#}$, $D^{\#}$, $F^{\#}$, $G^{\#}$, $A^{\#}$) are the least stable. Music that does not conform to the tonal hierarchy (thus atonal) is perceived as disorganized and unpleasant by listeners regardless of training (Dibben, 1994; Gagnon & Peretz, 2000) and accordingly is heard extremely rarely (Everett, 2000).

There are abstract organizational structures in time as well, most notably metre. Metre refers to the hierarchical pattern of alternation between strong and weak moments in time (Hannon, Snyder, Eerola, & Krumhansl, 2004). Notes occur more often on strong beats, and such positions in time are more psychologically stable than weak beats, forming a metric hierarchy in a similar manner to that of the tonal hierarchy (Palmer & Krumhansl, 1990). Rhythmic figures, or repeating patterns of durations within a musical sequence, are a primary factor in establishing the accented (i.e., strong) moments in time that contribute to the construction of metre (Lerdahl & Jackendoff, 1983). Metre in Western music is overwhelmingly in binary (strong–weak pattern) or ternary (strong–weak–weak) metre, but may have nested levels of both to create a compound metre (e.g., strongest–weak–weak–strong–weak–weak). These strong/weak cycles repeat in units called measures, such that the earliest possible temporal position within the measure is the strongest metrical position (the downbeat). Rhythm and metre are indispensible components of music; however, in terms of sheer complexity, tonality has by far the most informative value. Whereas tonality involves the multilevel hierarchical organization of 12 pitch classes, typical Western music uses only 2–3 unique durations (Fraisse, 1982) and a metrical framework that rarely uses more than two levels (Temperley, 2001).

Stimuli that resemble typical Western music (i.e., tonal) should invoke pitch salience, whereas atonal sequences (not representing typical Western music) should be less likely to invoke pitch salience, due to the lack of conformity to tonal pitch structure. Prince, Schmuckler, et al. (2009) tested this possibility by varying the degree to which the stimuli were tonal and consequently observed changes in dimensional salience. Participants compared the pitch height of a standard and comparison tone (which pitch was higher?), but had to ignore a sequence of tones that intervened between the standard and comparison. This sequence was either tonal or atonal. The temporal interval between the onset of each tone was identical except for the interval between the final tone of the sequence and the subsequent comparison tone. That is, the comparison tone could be early, on time, or late by virtue of altering or preserving that final temporal interval. When the intervening sequence was tonal, this timing variation had no effect on accuracy; however, when the sequence was atonal, participants performed better when the comparison tone was on time. In other words, the timing information was irrelevant when the sequence was tonal, presumably because the dimension of pitch was sufficiently salient as to overwhelm the influences of time. Conversely, when the sequences were atonal (lacking tonal pitch structure), pitch was no longer more salient and enabled effects of time to emerge.

Dimensional salience has the potential to reconcile many of the seemingly contradictory findings on pitch–time integration. But how might dimensional salience be manipulated? Given that the presence of tonality in a sequence affected pitch salience (Prince, Schmuckler, et al., 2009), varying the degree of conformity to a predefined structure (surface based or abstract) in a given stimulus dimension might be an effective method of manipulating dimensional salience. For example, changing the presence of metre may influence the salience of time (e.g., Jones, Moynihan, MacKenzie, & Puente, 2002). However, the existing work on dimensional salience occurred only in the context of responses to single tones (Prince, Schmuckler, et al., 2009; Prince, Thompson, et al., 2009). Listeners do not hear only single notes in a musical experience, but rather more complex musical units such as phrases and entire melodies. On this larger scale, nuanced patterns and structure emerge in both the dimension of pitch (e.g., melody, contour, and tonality) and that of time (e.g., rhythm and metre). As the existing work on dimensional salience has focused on the role of abstract structural principles in music (tonality and metre), the experiments in the present paper continue along that line and aim to determine whether findings comparable to those found in single-tone rating studies can result from responses to tone sequences. However, future work may focus on exploring this possibility in other forms of structure (both surface based and abstract) and in other domains such as language and vision. Because the present research focuses on the abstract organizational principles of tonality and metre, in the context of this paper the terms "pitch" and "time" refer to the sequences of pitch and temporal elements that form these larger emergent patterns.

The primary goal of these experiments is to test whether systematically altering the degree of conformity to structure within the pitch and temporal dimensions (specifically, tonality and metre) affects their relative dimensional salience. Experiments 1 and 2 consist of goodness ratings of sequences that vary in the degree of conformity to tonality and metre; Experiment 3 has dichotomous conformity judgements (i.e., classification) of whether a sequence adheres to the tonal and/or metric structure (using the same sequences as those in Experiments 1 and 2). All experiments use selective attention conditions (respond based only on pitch, or time) and a more holistic condition (respond based on both pitch and time). In terms of the stages of processing and local/global theories, these tasks can be characterized as requiring a fairly late stage of processing and attention to global characteristics of the stimuli. Therefore, both of these theories predict interaction between pitch and time regardless of the degree of conformity to tonality and metre. From the perspective of dimensional salience, the pattern of pitch–time integration should vary if the manipulations of stimulus structure and task influence the relative salience of both dimensions. If one dimension is more salient than another, then asymmetric interactions are likely, whereas a more even matching of salience may result in more global interactions.

EXPERIMENT 1

The rationale of Experiment 1 was to investigate how ratings of pitch and temporal goodness varied as a function of the degree to which a sequence adhered to a coherent tonal and metric structure. Goodness ratings are not a new task. For example, Palmer and Krumhansl (1987a, 1987b) collected ratings of complete musical phrases, and also for versions of these phrases that separated the pitch or temporal patterns, yielding either equitonal rhythmic sequences or isochronous pitch sequences. The ratings of these separated pitch and temporal versions combined linearly to predict goodness ratings of the original melody.

Experiment 1 used a new approach in both the stimuli and task instructions. First, the stimulus sequences varied in their degree of conformity to tonality and metre (these manipulations are detailed in the Method section). Second, participants' instructions were to make judgements of "goodness" while attending selectively to pitch information (ignoring time), temporal information (ignoring pitch), or both pitch and time simultaneously. The two selective attention instructions resemble more recent work on judgements of single isolated notes following a musical context (Prince, Thompson, et al., 2009). In their study, judgements of goodness of fit of a probe event preceded by a melody also showed independent contributions of both the tonal and metric hierarchy to ratings, even under instructions to ignore the pitch dimension. However, the present study focuses on larger musical sequences, and not single notes isolated from a preceding context. In turn, this task requires attention to more global aspects of the stimulus than individual notes. Another difference concerns the selective attention manipulation. In the present study, instructions to attend to pitch and/or time was a within-subjects variable instead of a between-subjects variable. There is also the issue of musical training; both Prince, Thompson, et al. (2009) and Prince, Schmuckler, et al. (2009) used trained performers as participants. To preserve continuity with these earlier findings, only trained performers participated in Experiment 1. However, to address potential issues of external validity in generalizing results from trained performers to all listeners, untrained participants were tested in Experiment 2.

Method

Participants

There were 24 participants in Experiment 1, with an average age of 19.4 years ($SD = 1.0$); all of them had at least 8 years of musical training in private lessons ($M = 10.5$, $SD = 2.4$). Participants were recruited using a flyer posted on campus and an online experiment database for an introductory psychology course. Recruitment took place at the University of Toronto Mississauga ($n = 17$) and the University at Buffalo ($n = 7$). Compensation was either course credit or $10.

Stimuli

The stimuli used in Experiment 1 were based on 48 sightsinging melodies (from Berkowitz, Fontrier, & Kraft, 1997; and Ottman, 1986) that epitomized the principles of typical Western music (including establishing both tonality and metre). These melodies functioned as starting points, or "seed melodies" for variation and factorial recombination of tonal and metric conformity. Their average length was 34.2 notes ($SD = 12.7$), and average duration was 9.1 seconds ($SD = 1.8$). For each melody, four levels of structural conformity in both pitch and time were constructed. The first level was the sequence of elements (pitch classes or durations) found in the original melody. The second level consisted of a randomly reordered sequence of the original elements, which nonetheless preserved the existing tonal or metric structural conformity. The third level was a randomly reordered sequence of the original elements, but which perturbed the tonal or metric structural conformity. The fourth level was a randomized sequence that included elements not found in the original sequence (pitch classes or durations). These manipulations occurred separately in pitch and time, such that there were 16 factorial recombinations (hereafter variants) for each seed melody. Figure 1 depicts 4 of the 16 variants for a single seed melody. Starting from 48 seed melodies, creating all 16 variants yielded 768 stimuli.

To construct all 16 variants of each seed melody, the original sequence was loaded into MATLAB (Mathworks, 2004) as a midi file, and then the pitch and duration sequences were randomized separately. A script programmed in MATLAB randomly permuted the order of the pitches and durations until they met the specified criteria, detailed forthwith. After randomization, the script recombined the sequences to form the variants and save them as midi files. Further scripts converted the midi files to .wav files.
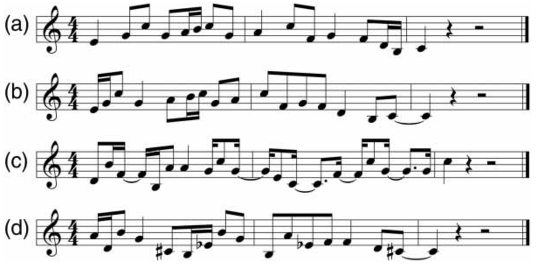
Figure 1. *Example variants of a single seed melody: (a) original seed melody (Pitch Variant Level 1, Time Variant Level 1); (b) reordered duration (Pitch Variant Level 1, Time Variant Level 2); (c) reordered pitch and duration (Pitch Variant Level 2, Time Variant Level 3); (d) random pitch, reordered duration (Pitch Variant Level 4, Time Variant Level 2).*

Pitch variant levels

Because Level 1 was the original pitch sequence, it required no modifications; however, the extent to which it adhered to the tonal hierarchy was assessed for comparison with the other levels. The Krumhansl and Schmuckler key-finding algorithm (Krumhansl & Schmuckler, 1986; described in Krumhansl, 1990) was used to find the best fit key of the original sequence. This algorithm also provides a maximum key-profile correlation (MKC)—the correlation coefficient of the tonal hierarchy of the best fit key—which was stored for comparison with other variant levels. The average MKC was .88 ($SD = .06$), indicating that the melodies were strongly tonal.

The order of the pitches in the original sequence was randomized to construct Level 2 pitch variants. The specified criterion of the second pitch variant level was that the MKC of the reordered sequence had to remain within $\pm.05$ of the MKC of the Level 1 (original) sequence. This criterion ensured that the new sequences were decidedly tonal, while allowing slight variation as a result of reordering the original sequence. Average MKC was .87 ($SD = .07$). Because this variant level consisted of reordering the existing elements, the frequency of occurrence of each pitch did not change.

For Pitch Variant Level 3, the original pitches were again reordered (as in Level 2), but had to reduce the MKC by at least .2 (average MKC was .64, *SD* = .07). The size of the numerical criterion was chosen on the basis that it would create a statistically significant reduction in MKC. Reordering the pitches reduced the MKC by altering the relative cumulative durations of each pitch class, subsequently leading changes in the duration-weighted MKC. Given that MKC is calculated by using cumulative durations of pitch class, this manipulation of tonal strength was therefore ultimately based on a temporal change (albeit based on cumulative sum, rather than sequential timings). The implications of this limitation are addressed in the Discussion section.

The sequences in Pitch Variant Level 4 were atonal, meaning that they did not adhere to tonal pitch structure. Construction of this variant level had two main constraints. First, 7 pitch classes were selected from the 12 found in Western music (the same number as that found in the other pitch variant levels), but comprised an artificial scale that did not correspond to any scale found in Western music (thus atonal). Starting on C, the pitch classes were C, C$^{\#}$, D$^{\#}$, F, G, A, and B; however, this set was transposed to start on a variety of different pitches (e.g., Figure 1d starts this scale on D). To address issues of ecological validity in the choice of pitch classes, this artificial scale preserved a near-universal feature of scales in all musical cultures in that the pitch intervals between adjacent scale members were of variable size (Dowling, 1978; Kessler, Hansen, & Shepard, 1984; Sloboda, 1985), in this case either one or two semitones. Such unequal-step scales are associated with processing advantages in infancy and adulthood (Trehub, Schellenberg, & Kamenetsky, 1999). Accordingly, the fourth level of pitch variant constitutes only a removal of tonality and is not confounded with the number of pitch classes. The second constraint on the fourth-level pitch variants was that the pitch range could be no larger than the corresponding original sequence. The average MKC was .54 (*SD* = .11).

Time variant levels

The method for creating time variant levels was equivalent to that used for creating the pitch variant levels. Level 1 was the original sequence of durations and thus involved no modifications. By correlating the number of onsets at each temporal location with the metric hierarchy as measured by Palmer and Krumhansl (1990), it was possible to calculate the extent to which the melody instantiated a duple metre (specifically, 4/4). This metric hierarchy correlation (MHC) for the original sequences was .76 (*SD* = .17), demonstrating that the sequences were distinctly metric.

Time Variant Level 2 consisted of a reordering of the original durations. The order of the durations was randomized, with the constraint that MHC (as calculated for Level 1) could not vary more than ±.05 of the MHC value for the original sequence. Pitch changes did not affect these correlations because MHCs are not weighted according to pitch class. The average MHC was .75 (*SD* = .16).

Another randomization of the order of durations was used for Time Variant Level 3, but in this case the MHC was reduced by .2 (average MHC was .26, *SD* = .32). The size of the required reduction was chosen to ensure a statistically significant reduction in MHC.

The sequences of Time Variant Level 4 had entirely random timings, not limited to standard note duration denominations (e.g., eighth note, half note, etc.). The number of temporal events (durations) remained the same as that found in the original sequence. Furthermore, no durations were shorter or longer than the minimum or maximum (respectively) durations found in the original sequence. Lastly, the cumulative (total) duration of each sequence could not vary more than ±10% from the cumulative duration of the original sequence. Given that this variant level was by design not quantized to standard metric positions, it was not possible to calculate the MHC, nor to depict the variants accurately in musical notation (thus Figure 1 does not have an example of Time Variant Level 4).

Apparatus

A Macintosh G4 running OSX 10.3 was used to run the experiment at University of Toronto Mississauga. The experimental environment was created in MATLAB (Mathworks, 2004), using the Psychophysics Toolbox extensions (Brainard, 1997). Participants wore Sennheiser HD 280 Pro headphones during the experiment and were in a double-walled soundproof booth (IAC). The participants from University at Buffalo used a Macintosh G5 running OSX 10.4 and were in a quiet room during the experiment. All other aspects of the apparatus were identical across location.

Procedure

All participants were tested individually. They provided informed consent and completed a background questionnaire on musical experience. Participants received instructions to listen to each melody and evaluate how "good" it sounded. The experimenter stressed that participants were not to indicate how much they personally liked the melody, but rather how well formed it was—for example, if it sounded like a normal, typical melody, conversely if it sounded like there was something wrong with the melody, and so on. The experimenter also explained that there were three instructional conditions—rate pitch (ignore time), rate time (ignore pitch), and rate both. The selective attention instructions were that participants were to ignore all aspects of the melody's timing and rhythm when the screen displayed the message: "Please rate this melody based ONLY on its PITCH characteristics." This explanation was repeated verbally with the opposite instructions (ignore the pitches) when the screen showed: "Please rate this melody based ONLY on its TEMPORAL characteristics." For the third instructional condition, participants were told to form their rating based on both dimensions simultaneously, when the screen read: "Please rate this melody based on BOTH its PITCH and TEMPORAL characteristics." These instructions remained on the screen before, during, and after each trial. An advantage of varying instruction as a within-subject variable is that potential individual differences to attend preferentially to pitch or time will not affect unequally the pattern of results across instructional conditions. A complication of this design is the possibility of contamination between blocks as a result of changing instruction. The analyses in the Results section address this issue. After the melody had finished, participants indicated their rating of melodic goodness on a scale of 1–7, and instructions on the screen reminded them on each trial that a rating of 1 meant very bad, and 7 meant very good. Participants were encouraged to use the full rating scale during the course of the experiment.

Participants completed eight practice trials, with the instructional condition changing every two trials. Each trial began when the participant pressed the space bar to indicate their readiness. Prior to beginning a trial when the instructional condition changed, the program displayed the following message: "Now the experiment changes! You will now rate the melodies based on a different dimension. PLEASE pay attention to the change (please please please!). Press the 'y' key if you understand, or get the experimenter if you need help." The experimenter was present for the practice trials and would indicate verbally at this point that the experiment was changing and that now they would rate the melodies based on a different dimension. Once the participants indicated their understanding by pressing "y", a standard trial message appeared ("Please rate this melody based . . .") with instructions describing which dimension to attend. The participant then pressed the space bar to start the trial. Upon completing eight practice trials, participants had experienced this instructional change three times. The participants were informed that the instructional condition would change only twice during the experimental trials: once after 64 trials and again after the second 64 trials. The experimenter ensured that the participant understood what the instructions meant, answered any remaining questions, then began the experimental trials and left the room. There were 64 trials in each of the three instructional condition blocks, yielding 192 total trials, and the entire procedure took approximately one hour.

The order of blocks (instructional condition) was counterbalanced across participants. Also, the assignment of variants of a given melody to participants was determined using a counterbalanced Latin square design, such that all participants experienced all 16 variant conditions in all blocks, but only heard 4 variants of a given seed melody. Which 4 of 16 variants of a given melody each participant heard was counterbalanced across participants, to complete the Latin square. The 48 total seed melodies were divided such that each participant heard 4 variants from 16 unique seed melodies in each of the three blocks ($4 \times 16 \times 3 = 192$ trials). Therefore, each participant heard 4 variants of every seed melody—a quarter of the 768 total stimuli—such that 4 participants were required to complete the design for all variants of all melodies. Coupled with 6 possible instructional block orders (3 factorial), 24 participants were needed to counterbalance block order and rate all variants of all melodies.

Results

Goodness ratings were averaged across melody, giving 48 data points per participant (16 variants in each of three instructional conditions). The mean intersubject correlation of the ratings was .64 ($SD = .07$), indicating substantial agreement across participants. Initial inspection of the data revealed that goodness ratings for Pitch Variant Levels 2 and 3 overlapped, as well as Time Variant Levels 2 and 3. Bonferroni-corrected pairwise comparisons verified this pattern, showing no significant difference between the second and third pitch variant levels for the instruction conditions of "rate pitch", $M = .23$, 95% CI [−.01, .47], $p = .07$, "rate time", $M = .20$, 95% CI [−.02, .43], $p = .10$, and "rate both", $M = .14$, 95% CI [−.16, .43], $p = 1$. The second and third time variant levels also did not differ for any instructional condition, $M = .12$, 95% CI [−.06, .29], $p = .45$, $M = .24$, 95% CI [−.03, .51], $p = .10$, $M = .09$, 95% CI [−.14, .31], $p = 1$, respectively. Therefore, in all subsequent analyses, these two conditions were averaged (within a single dimension), giving three levels of both pitch and time dimensions (nine factorial combinations).

The effects of pitch, time, and block order on goodness ratings were tested with three mixed analyses of variance (ANOVAs; one ANOVA for each instructional condition), each using a 3 (pitch: Level 1, Level 2−3, Level 4) × 3 (time: Level 1, Level 2−3, Level 4) × 3 (block order) design. For these analyses, pitch and time variant levels were within-subjects variables, and the between-subjects variable was block order—that is, the block number in which the participant rated that dimension. For the "rate pitch" instruction analysis, the block order variable indicated the block in which each participant rated pitch (1, 2, or 3); for the "rate time" instruction analysis, the block order variable denoted in which block each participant rated time, and so on for the "rate both" instruction. These block order variables were included because the within-subjects manipulation of instruction creates a potential complication in the results. Specifically, any interference of the ignored dimension observed in the selective attention conditions could be due to contamination between blocks as a result of switching the attended dimension, rather than a true failure of selective attention. By including block order as a between-subjects variable, these ANOVAs tested whether the effects of pitch and time (and any interaction between them) changed based on when the participants completed each instruction.

Rate pitch instruction

The "rate pitch" instruction 3 (pitch) × 3 (time) × 3 (block order) ANOVA (as described above) revealed a main effect of pitch, $F(2, 42) = 208.22$, $MSE = 1.00$, $p < .001$, $\eta^2 = .732$, and a main effect of time, $F(2, 42) = 35.10$, $MSE = .63$, $p < .001$, $\eta^2 = .078$ on goodness ratings. Block order (the block number in which the participants rated pitch) was not significant as a main effect, $F(2, 21) < 1$, $MSE = 2.73$, $p = .48$, $\eta^2 = .001$, but there was an interaction between pitch and block order, $F(4, 42) = 3.43$, $MSE = 1.00$, $p = .02$, $\eta^2 = .024$. This interaction was due to the fact that the effect size of pitch was slightly larger for the participants who rated pitch in their second block, $F(2, 14) = 179.60$, $MSE = .59$, $p < .001$, $\eta^2 = .889$, than those participants who rated it in their first block, $F(2, 14) = 37.42$, $MSE = 1.66$, $p < .001$, $\eta^2 = .675$, or last block, $F(2, 14) = 61.87$, $MSE = .73$, $p < .001$, $\eta^2 = .637$. Block order did not interact with time, $F(4, 42) = 1.22$, $MSE = .63$, $p = .32$, $\eta^2 = .005$. There was no interaction between pitch and time, $F(4, 84) = 2.37$, $MSE = .23$, $p = .06$, $\eta^2 = .004$, nor was there between pitch, time, and block order, $F(8, 84) < 1$, $MSE = .23$, $p = .59$, $\eta^2 = .003$. Figure 2a displays the goodness ratings as a function of pitch and time variant level.

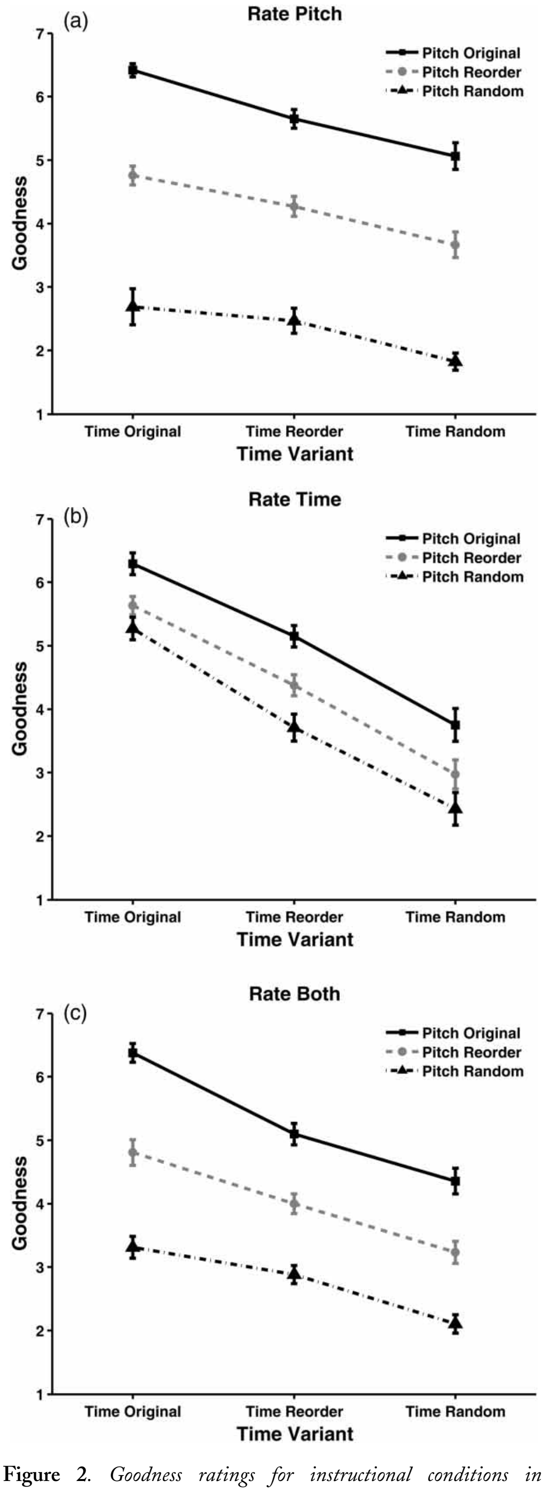

Figure 2. *Goodness ratings for instructional conditions in Experiment 1 (musically trained performers): (a) "rate pitch" condition; (b) "rate time" condition; (c) "rate both" condition. Error bars represent standard error of the mean.*

Rate time instruction

For the "rate time" instruction, the 3 (pitch) × 3 (time) × 3 (block order) mixed ANOVA yielded main effects of pitch, $F(2, 42) = 54.44$, $MSE = .53$, $p < .001$, $\eta^2 = .128$, and time, $F(2, 42) = 76.32$, $MSE = 1.70$, $p < .001$, $\eta^2 = .573$, but not block order (the block number in which the participants rated time), $F(2, 21) < 1$, $MSE = 3.21$, $p = .65$, $\eta^2 = .001$. None of the interactions was significant: pitch and block order, $F(4, 42) = 1.07$, $MSE = .53$, $p = .38$, $\eta^2 = .005$; time and block order, $F(4, 42) = 1.83$, $MSE = 1.70$, $p = .14$, $\eta^2 = .027$; pitch and time, $F(4, 84) = 1.06$, $MSE = .29$, $p = .38$, $\eta^2 = .003$; pitch, time, and block order, $F(8, 84) < 1$, $MSE = .29$, $p = .66$, $\eta^2 = .004$. Figure 2b exhibits these data.

Accordingly, for both selective attention conditions, both the dimensions of pitch and time contributed to melodic goodness ratings, despite instructions to ignore either dimension (i.e., the lines of Figure 2a are not flat, and the lines of Figure 2b are not superimposed). Also, pitch and time combined in a linear (additive) fashion, with no interaction between them. The change in relative effect sizes of pitch and time across instructional conditions signifies that these listeners were able to emphasize either dimension in accordance with the instructions, but not eliminate completely the contribution of the other dimension.

Rate both instruction

The results of the 3 (pitch) × 3 (time) × 3 (block order) mixed ANOVA were more complex for the "rate both" instruction, as shown in Figure 2c. The only significant main effects were, again, pitch, $F(2,\ 42) = 249.81$, $MSE = .45$, $p < .001$, $\eta^2 = .583$, and time, $F(2, 42) = 77.35$, $MSE = .60$, $p < .001$, $\eta^2 = .237$, whereas block order (the block number in which the participants rated both dimensions) was not significant, $F(2, 21) < 1$, $MSE = 3.55$, $p = .63$, $\eta^2 = .001$. The effect sizes denote that pitch contributed more strongly to ratings than time. More importantly, however, there was an interaction between pitch and time, $F(4,\ 84) = 7.33$, $MSE = .19$, $p < .001$, $\eta^2 = .014$, showing that the contribution of pitch and time in goodness ratings was not purely linear, but instead varied based on the particular combination of pitch and time variant levels. Block order did not interact with any variables: pitch and block order, $F(4, 42) < 1$, $MSE = .45$, $p = .68$, $\eta^2 = .003$; time and block order, $F(4, 42) = 1.11$, $MSE = .60$, $p = .36$, $\eta^2 = .007$; pitch, time, and block order $F(8, 84) < 1$, $MSE = .19$, $p = .90$, $\eta^2 = .002$.

Discussion

The first and most obvious result in these data is that, reassuringly, ratings of goodness varied in accordance with the conformity to pitch and temporal structure (tonality and metre). Interestingly, goodness ratings did not change between the second and third variant levels. Both of these levels used the same pitches or durations found in the original, but reordered them to preserve or perturb (respectively) the tonal or metric hierarchy. One possible explanation for this finding is that the goodness ratings were not linked to the numerical degree of conformity to musical structure (i.e., the tonal and metric hierarchy correlations), but instead were affected only by upsetting the original sequential order of elements found in the first variant level. Certainly in the case of pitch, reordering the pitches destroys the contour of an original melody, including its primarily stepwise nature (most melodies move up or down by small intervals). This reordering therefore made the pitch sequences atypical and consequently not as well formed, or as good, as the originals. However, the fourth variant level also disturbed the contour of the original melody but added a violation of the tonal hierarchy by using pitches that did not belong in any one key; this variant level received ratings significantly lower than the other levels. Accordingly, the operation of randomizing the order of elements cannot be solely responsible for the effect of pitch on goodness ratings.

In fact, the ineffectiveness of the manipulation between the second and third pitch variant levels suggests a different conclusion. Specifically, a musical context can establish a tonality in an all-or-none fashion simply by presenting the collection of pitch classes that belong in that key, rather than having to arrange the cumulative relative durations of these pitch classes in close accordance with the tonal hierarchy. In turn, listeners' experience of tonality cannot be only a reflection of its immediate pitch distribution, but instead is largely formed by invoking a stored internal representation of the tonal hierarchy activated by a sufficiently similar context (i.e., collection of pitch classes) in the stimulus they hear. Additionally, this lack of effect renders moot the potential limitation of differences between Pitch Variant Levels 2 and 3 being due to a manipulation of cumulative duration (as mentioned in the Method section).

It is more difficult to explain why there was no difference between the second and third variant levels that altered the metric conformity, compared to the original sequence. Reordering the order of durations inevitably makes a musical sequence sound more syncopated. Perhaps an all-or-none effect (similar to that found described in the pitch variants) of syncopation exists with the perception of metre and temporal goodness, such that the degree of syncopation is not the operative factor, but its presence or absence. Another possibility is that the manipulation of correlation with the metric hierarchy may not have been sufficiently obvious so as to cause listeners to differentiate them in terms of a goodness rating.

For the "rate pitch" condition, there was an interaction between pitch variant level and the block in which participants rated pitch. The effect size of pitch was larger for those listeners who rated pitch selectively in their second block than for those who completed the "rate pitch" instruction at the beginning or end of the experiment (yet the effect was strong in all blocks). It is unclear why this interaction arose. There were no other interactions involving block order (neither with pitch nor time) in any instructional condition. Furthermore, the critical three-way interaction between pitch, time, and block order was not significant in the "rate pitch" instruction (or any other instruction), showing that block order did not affect the pattern of pitch–time integration.

Another interesting finding in these data is that both pitch and time contributed to goodness ratings regardless of instruction. Listeners were unable to ignore pitch entirely when rating temporal goodness (and time when rating pitch goodness), instead demonstrating a failure of selective attention (Figures 2a and 2b). Nevertheless, they did emphasize the attended dimension in their ratings, such that the effect size of pitch and time varied according to the instructions. The effect of pitch ($\eta^2 = .732$) was substantially larger than that of time ($\eta^2 = .078$) in the "rate pitch" instruction, and the effect of time ($\eta^2 = .573$) easily exceeded that of pitch ($\eta^2 = .128$) when participants rated time. The ability to emphasize selectively either dimension in accordance with the instructions suggests that participants were largely (although not entirely) able to process these dimensions independently. If instead the effect sizes of pitch and time remained constant for both selective attention conditions (i.e., listeners could not accentuate either dimension as per the instructions), then such failures of selective attention would suggest instead that listeners processed the dimensions in a more interactive manner. Furthermore, in these selective attention conditions, the effects of pitch and time combined linearly to predict goodness ratings, with no statistical interaction between them. Taken together, the additive manner in which these dimensions combined and the ability to weight one dimension as instructed suggests that on the continuum of independent to interactive pitch–time relations, listeners processed pitch and time more independently than interactively in this task.

However, a slightly different picture emerges from the "rate both" condition. When instructed to incorporate both dimensions into their goodness rating (i.e., no selective attention instructions), pitch and time showed a more interactive relation. Thus a simple summation of the contributions of these dimensions was not sufficient to predict goodness ratings when listeners attended to the sequence as a whole instead of selectively attending one dimension. What was the nature of this interaction? Judging from Figure 2c, the effect of one dimension increased when the other dimension adhered more to its original structure—the slope of the lines (indicating the effect of time) was greater as the degree of tonal conformity more closely approximated that of the original variant level. Another way of summarizing the same result is to note that the separation between the lines (indicating the effect of pitch) was greatest for the original time variant.

Indeed, overall these findings support well the notion that pitch was more salient than time in this experiment. Pitch accounted for over twice the variance that time did when participants attended both dimensions ($\eta^2 = .583$ vs. $\eta^2 = .237$, respectively). Furthermore, the advantage of pitch over time in the "rate pitch" condition (pitch η^2 − time $\eta^2 = .654$) was considerably larger than the corresponding advantage of time for the "rate time" condition (time η^2 − pitch $\eta^2 = .445$). Nevertheless, both dimensions contributed significantly to predict ratings in all three instructional conditions.

Musical training imparts the ability to examine music analytically, which could potentially help listeners to separate the dimensions of pitch and time. However, the failures of selective attention suggest that these musically trained participants were not entirely successful in this attempt. Nevertheless, musical expertise may have influenced the results in other unpredictable ways.

Thus to ensure the validity of generalizing these findings to other populations (i.e., not only those with formal musical training), untrained listeners participated in Experiment 2.

EXPERIMENT 2

Neither rating the goodness of a melody, nor understanding the instruction to attend only to the pitch or timing of a musical sequence, requires explicit musical training. Therefore Experiment 2 had the exact same stimuli and design as those in Experiment 1, but the participants were untrained listeners instead of trained performers. If passive exposure to typical Western music (as opposed to formal training) is sufficient to be sensitive to its statistical properties (e.g., pitch salience), then the same pattern of results should occur regardless of the degree of formal musical training. However, the familiarity with music afforded by formal training does tend to improve the overall reliability and accuracy of responses to musical stimuli, especially in tasks that require explicit knowledge of musical structure (cf. Bigand, 2003; Bigand & Poulin-Charronnat, 2006), so some quantitative (not qualitative) differences may be expected.

Method

Participants

There were 24 musically untrained participants in Experiment 2; none had more than 5 years of private music lessons ($M = 1.19$, $SD = 1.70$). The average age of the participants was 23.13 years ($SD = 6.00$). These participants were recruited at Murdoch University using flyers and an online experiment database, and they were compensated with either course credit or $10.

Stimuli

All stimuli were the same as those in Experiment 1.

Apparatus

A PC running Windows XP was used to run the experiment in a quiet room, but all other aspects of the apparatus were the same as those in Experiment 1.

Procedure

The procedure for Experiment 2 was identical to that for Experiment 1.

Results

Data were analysed the same way as in Experiment 1. Goodness ratings were averaged across melody, giving 48 data points per participant; the mean intersubject correlation of the ratings was .58 ($SD = .09$). As in Experiment 1, initial examination of the data revealed that goodness ratings for Pitch Variant Levels 2 and 3 overlapped, as well as Time Variant Levels 2 and 3. Bonferroni-corrected pairwise comparisons showed a significant difference between the second and third pitch variant levels only for the "rate pitch" instruction, $M = .24$, 95% CI [.06 .42], $p = .01$, but not for "rate time", $M = -.03$, 95% CI [−.34, .29], $p = 1$, nor for "rate both", $M = .17$, 95% CI [−.08, .42], $p = .37$. Time Variant Levels 2 and 3 did not differ for any instructional condition, for "rate pitch", $M = .14$, 95% CI [−.20, .47], $p = 1$, "rate time", $M = .09$, 95% CI [−.16, .34], $p = 1$, and "rate both", $M = .08$, 95% CI [−.18, .35], $p = 1$. Because only one of these differences reached significance (and only in one instructional condition), and to maintain continuity with Experiment 1, these two variant levels were averaged (within a single dimension), giving three levels of both pitch and time dimensions for all subsequent analyses.

Next, three mixed 3 (pitch: Level 1, Level 2−3, Level 4) × 3 (time: Level 1, Level 2−3, Level 4) × 3 (block order) ANOVAs tested the effects of pitch, time, and block order on goodness ratings, separately for each instructional condition. Pitch and time variant levels were within-subjects variables for these analyses, and block order was the between-subjects variable.

Rate pitch instruction

For the "rate pitch" instruction there was a main effect of pitch, $F(2, 42) = 154.22$, $MSE = .71$,

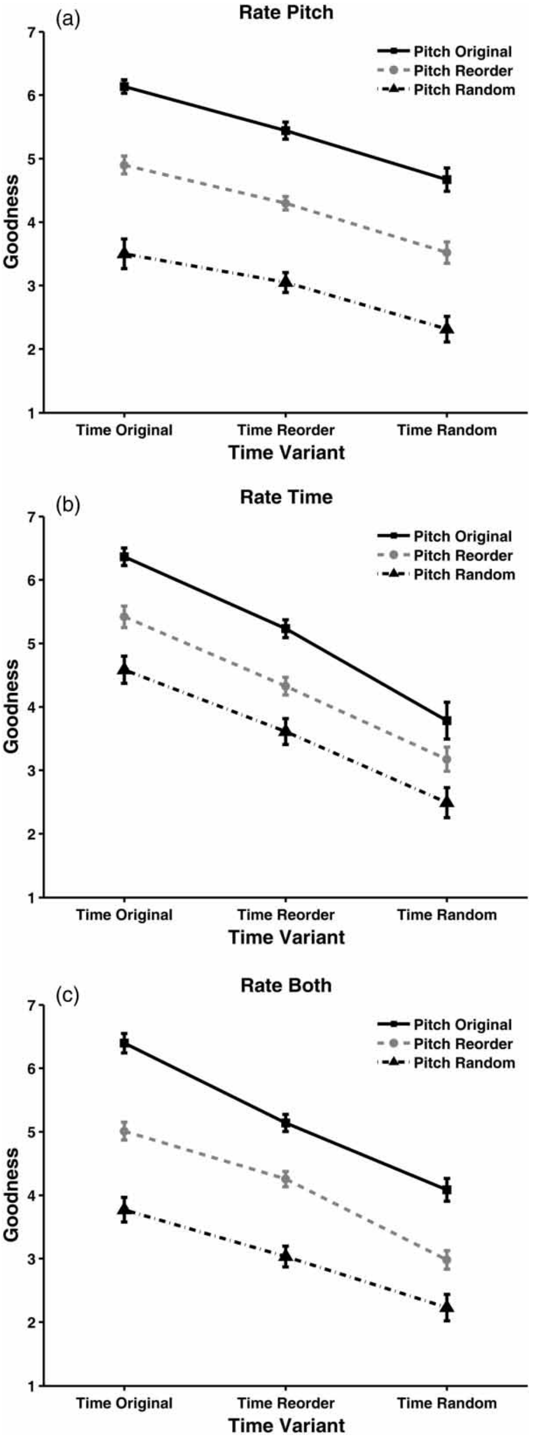

Figure 3. *Goodness ratings for instructional conditions in Experiment 2 (musically untrained listeners): (a) "rate pitch" condition; (b) "rate time" condition; (c) "rate both" condition. Error bars represent standard error of the mean.*

$p < .001$, $\eta^2 = .586$, and a main effect of time, $F(2,\ 42) = 59.49$, $MSE = .55$, $p < .001$, $\eta^2 = .176$ on goodness ratings. There was no main effect of block order (the block number in which the participants rated pitch), $F(2,\ 21) = 1.17$, $MSE = 1.89$, $p = .20$, $\eta^2 = .001$, but there was an interaction between time and block order, $F(4,\ 42) = 3.23$, $MSE = .55$, $p = .02$, $\eta^2 = .019$. This interaction was caused by a larger effect size of time for the participants who rated pitch in their second block, $F(2,\ 14) = 85.87$, $MSE = .26$, $p < .001$, $\eta^2 = .349$, than for those participants who rated it in their first block, $F(2,\ 14) = 17.44$, $MSE = .45$, $p < .001$, $\eta^2 = .111$, or last block, $F(2,\ 14) = 6.94$, $MSE = .95$, $p = .01$, $\eta^2 = .125$. Block order did not interact with pitch, $F(4,\ 42) = 1.42$, $MSE = .71$, $p = .24$, $\eta^2 = .011$. There was also no interaction between pitch and time, $F(4,\ 84) < 1$, $MSE = .27$, $p = .70$, $\eta^2 = .002$, nor was there between pitch, time, and block order, $F(8,\ 84) < 1$, $MSE = .27$, $p = .56$, $\eta^2 = .005$. Figure 3a displays the goodness ratings as a function of pitch and time variant level.

Rate time instruction

The same main effects emerged in the "rate time" instruction: pitch, $F(2,\ 42) = 70.92$, $MSE = .62$, $p < .001$, $\eta^2 = .217$, and time, $F(2,\ 42) = 81.62$, $MSE = 1.18$, $p < .001$, $\eta^2 = .474$, but not block order, $F(2,\ 21) < 1$, $MSE = 3.27$, $p = .92$, $\eta^2 = .000$. None of the interactions was significant: pitch and block order, $F(4,\ 42) < 1$, $MSE = .62$, $p = .50$, $\eta^2 = .005$; time and block order, $F(4,\ 42) = 2.02$, $MSE = 1.18$, $p = .11$, $\eta^2 = .023$; pitch and time, $F(4,\ 84) < 1$, $MSE = .41$, $p = .41$, $\eta^2 = .004$; pitch, time, and block order, $F(8,\ 84) < 1$, $MSE = .41$, $p = .79$, $\eta^2 = .005$. Figure 3b exhibits these data.

As in Experiment 1, both the dimensions of pitch and time contributed to melodic goodness ratings for both selective attention conditions, despite instructions to ignore either dimension. Also, pitch and time combined in a linear (additive) fashion, with no interaction between them. The change in relative effect sizes of pitch and time across instructional conditions signifies that these listeners were able to emphasize either dimension in accordance with the instructions, but not completely eliminate the contribution of the other dimension.

Rate both instruction

The "rate both" instruction results also mirrored Experiment 1, as shown in Figure 3c. Pitch and time were the only significant main effects, $F(2, 42) = 117.39$, $MSE = .74$, $p < .001$, $\eta^2 = .420$, $F(2, 42) = 105.61$, $MSE = .66$, $p < .001$, $\eta^2 = .335$, respectively, but not block order, $F(2, 21) < 1$, $MSE = 1.62$, $p = .80$, $\eta^2 = .000$. Again, the effect sizes denote that pitch contributed more strongly to ratings than time. More importantly, however, there was an interaction between pitch and time, $F(4, 84) = 3.66$, $MSE = .32$, $p = .01$, $\eta^2 = .011$, showing that the contribution of pitch and time in goodness ratings was not purely linear, but instead varied based on the particular combination of pitch and time variant levels. Block order did not interact with pitch, $F(4, 42) < 1$, $MSE = .74$, $p = .78$, $\eta^2 = .003$, with time, $F(4, 42) = 2.10$, $MSE = .66$, $p = .10$, $\eta^2 = .013$, or with the pitch–time interaction, $F(8, 84) = 1.59$, $MSE = .32$, $p = .14$, $\eta^2 = .010$.

Discussion

The goodness ratings of melodies varying in their degree of conformity to the tonal and metric hierarchies provided by musically untrained listeners (Experiment 2) showed nearly exactly the same patterns as those of trained performers (Experiment 1). Namely, neither group successfully eliminated the influence of pitch or time in selective attention conditions; however, in both cases the two dimensions made independent contributions to goodness ratings. Additionally, pitch and time demonstrated interactive relations in the "rate both" condition, for both participant groups. Finally, as in Experiment 1, pitch appeared to be more salient than time for Experiment 2, as measured by relative main effect sizes. In the "rate pitch" condition, pitch enjoyed a substantial advantage in effect size over time (pitch η^2 − time $\eta^2 = .410$), but time did not experience a similarly sized advantage in the "rate time" condition (time η^2 − pitch $\eta^2 = .256$). In the "rate both" condition the two dimensions were more equally matched (pitch η^2 − time $\eta^2 = .084$), but again tipped in favour of pitch.

In fact, the main divergence between Experiments 1 and 2 is that all effects (main and interaction) were quantitatively weaker for untrained listeners. This finding may be due to motivation—musically trained individuals are likely to engage more in musical tasks because they have invested deeply in this activity. Therefore, they tend to give cleaner and more robust effects than untrained participants who find it more difficult to remain engaged in a musical task. However, the smaller relative main effect sizes (pitch η^2 − time η^2) do not suggest that untrained listeners showed more evidence of interactive pitch–time processing than trained performers, as the interaction effect size (pitch × time η^2) was also weaker and only present in the "rate both" condition for both experiments. Indeed, the two populations showed the same qualitative pattern of effects, suggesting that the underlying implicit cognitive mechanisms responsible for pitch–time integration are unchanged across levels of expertise. This common pattern demonstrates selective attention failure, pitch salience, and more independent contributions in selective attention tasks yet increasingly interactive contributions when attending to both dimensions.

There were other differences between the two groups, however. For musically trained performers, pitch interacted with block order in the "rate pitch" instructional condition, whereas untrained listeners displayed an interaction of time with block order in the "rate pitch" instructional condition. These findings were due to variations in the main effect size of one dimension (curiously always largest in the second block) and did not affect any interactions between pitch and time. No obvious theoretical interpretation of these effects is forthcoming. Also, the difference between Pitch Variant Levels 2 and 3 reached significance only in the untrained listeners. The fact that this effect was small, was only present in one instructional condition, and did not occur in the trained performers (who, if anyone, would presumably be more sensitive to such manipulations, not less) renders this finding uninterpretable.

Overall, the results of Experiments 1 and 2 are remarkably similar, meaning that these findings generalize to all listeners encultured in Western music. Nevertheless, the pattern of findings is somewhat puzzling: Failures of selective attention suggest interaction between pitch and time, but their contributions to ratings were additive in these conditions, while interactive when rating both dimensions simultaneously.

One explanation to this pattern of findings has to do with task design, specifically the nature of the dependent measure. Participants rated the same abstract property of "goodness" regardless of which dimension they were attending. Accordingly, the reliable influence of pitch and time in both selective attention conditions may not indicate how participants processed pitch and time as separate dimensions, but instead be a result of evaluating both dimensions along a scale of the same basic psychological quality. A task in which information along one dimension is purely irrelevant to the other would provide a context that promotes the ability to process pitch and time more independently and therefore an opportunity to test how involuntary this interference is. To investigate this possibility, and to explore further the integration of pitch and time, Experiment 3 complemented the task of goodness ratings with an explicit classification—judging conformity to tonal and/or metric structure.

EXPERIMENT 3

The ratings of melodic goodness in Experiments 1 and 2 provide a number of interesting findings regarding how pitch and time combine in a musical context. The goal of Experiment 3 was to clarify these findings in a different task context and to provide convergent evidence for the pattern of pitch−time integration in the perception of melodies. By using a task that required explicit classifications—dichotomous judgements of conformity—of pitch or temporal structure in a melody, the stimulus manipulations should, theoretically, make a dimension truly irrelevant, rather than requiring active cognitive effort to suppress its contribution to the participant's response. Classification paradigms have a long history in research on dimensional integration (Garner, 1974; Pomerantz, 1983) and are still relevant today (for recent studies in audition see Dyson & Quinlan, 2010; Silbert, Townsend, & Lentz, 2009). Therefore Experiment 3 had the same melodies and the same instructions, but changed the task. Specifically, for the "classify pitch" condition, participants indicated whether the melody was tonal or atonal (did it establish a key area?). The "classify time" condition consisted of classifying whether the melody was metric or random (did it establish a regular beat?). The "classify both" condition required participants to evaluate whether the melody was both tonal and metric, or if either dimension was "off" (i.e., atonal, random, or both). Measuring the extent to which the irrelevant dimension contributed to participants' judgements in the selective attention conditions comprised the measure of interference between pitch and time. Given that these instructions require explicit knowledge of musical structure, all participants had formal musical training. Fortunately, Experiments 1 and 2 had established the similarity of pitch−time integration across expertise levels, attenuating concerns of generalizing findings to other populations.

Method

Participants

The participants in Experiment 3 were 24 musically trained performers; all of them had at least 8 years of musical training ($M = 9.7$, $SD = 2.4$), and their average age was 20.4 years ($SD = 2.9$). Participants were recruited using a flyer posted on campus and an online experiment database for an introductory psychology course. Recruitment occurred only at the University at Buffalo. Compensation was either course credit or $10. None of the participants from Experiment 1 participated in Experiment 3.

Stimuli

All stimuli were the same as those used in previous experiments.

Apparatus

The apparatus were the same as that used at the University at Buffalo in Experiment 1.

Procedure

The procedure was the same as that in the previous experiments except for the instructions provided for participants. Prior to beginning the practice trials, participants received thorough instructions. For the "classify pitch" condition, the on-screen instructions read: "Please classify if this melody is tonal (press T) or atonal (press A)." The "classify time" instructions were: "Please classify if this melody is metric (press M) or random (press R)." The term "random" was used instead of "ametric" such that the response key (R) would not be the same as the response key for "atonal" (A), used in the "classify pitch" block. Nevertheless, the participants received clear instructions to indicate whether the melody was metric or not (as in previous experiments). The "classify both" condition instructed participants: "Please classify if this melody is BOTH tonal AND metric (press B), or if it is EITHER atonal OR nonmetric (press E)." Particular care was given to explaining the "classify both" condition, as it had the greatest potential for confusion. The same message as that in Experiment 1 appeared when the instructions changed ("Now the experiment changes! You will . . .").

Results

Responses of "tonal", "metric", and "both tonal and metric" were coded as "yes" responses, whereas "atonal", "random", or "either atonal or random" were coded as "no" responses. Data were then analysed as the proportion of "yes" responses, as a function of conformity to pitch and temporal structure, and were averaged across melody. Intersubject correlations were again high (mean $r = .67$, $SD = .07$). Bonferroni-corrected pairwise comparisons of pitch variant levels showed that there was no significant difference between the second and third variant levels, for "classify pitch", $M = .04$, 95% CI [−.04, .11], $p = 1$, "classify time", $M = -.02$, 95% CI [−.09, .05], $p = 1$, and "classify both", $M = .04$, 95% CI [−.07, .16], $p = 1$. Time Variant Levels 2 and 3 were different for the "classify both" condition, $M = .11$, 95% CI [.04, .18], $p = .001$; however, they were not for "classify pitch", $M = -.00$, 95% CI [−.06, .06], $p = 1$, nor "classify time", $M = .01$, 95% CI [−.09, .12], $p = 1$. Accordingly, and as in the previous experiments, these two levels were averaged together (for pitch and time separately) in subsequent analyses, giving three levels of both pitch and time dimensions.

Further analyses of Experiment 3 data used the same three mixed ANOVAs with a 3 (pitch: Level 1, Level 2–3, Level 4) × 3 (time: Level 1, Level 2–3, Level 4) × 3 (block order) design to test the effects of pitch, time, and block order on classifications, for each instruction separately.

Classify pitch instruction

The "classify pitch" instruction ANOVA replicated the findings of the previous experiments (see Figure 4a)—there was a main effect of pitch, $F(2, 42) = 150.64$, $MSE = .06$, $p < .001$, $\eta^2 = .706$, and a main effect of time, $F(2, 42) = 21.15$, $MSE = .03$, $p < .001$, $\eta^2 = .044$, but no main effect of block order (the block in which participants classified pitch), $F(2, 21) = 1.37$, $MSE = 2.21$, $p = .28$, $\eta^2 = .005$. None of the interactions was significant: pitch and block order, $F(4, 42) < 1$, $MSE = .06$, $p = .45$, $\eta^2 = .009$; time and block order, $F(4, 42) = 1.17$, $MSE = .03$, $p = .34$, $\eta^2 = .005$; pitch and time, $F(4, 84) = 1.97$, $MSE = .03$, $p = .11$, $\eta^2 = .008$; pitch, time, and block order, $F(8, 84) < 1$, $MSE = .03$, $p = .67$, $\eta^2 = .006$.

Classify time instruction

For the "classify time" instruction, the ANOVA revealed an effect of pitch, $F(2, 42) = 41.97$, $MSE = .08$, $p < .001$, $\eta^2 = .240$, and time, $F(2, 42) = 96.83$, $MSE = .06$, $p < .001$, $\eta^2 = .397$, but not block order, $F(2, 21) = 1.93$, $MSE = 2.28$, $p = .17$, $\eta^2 = .008$. In contrast to

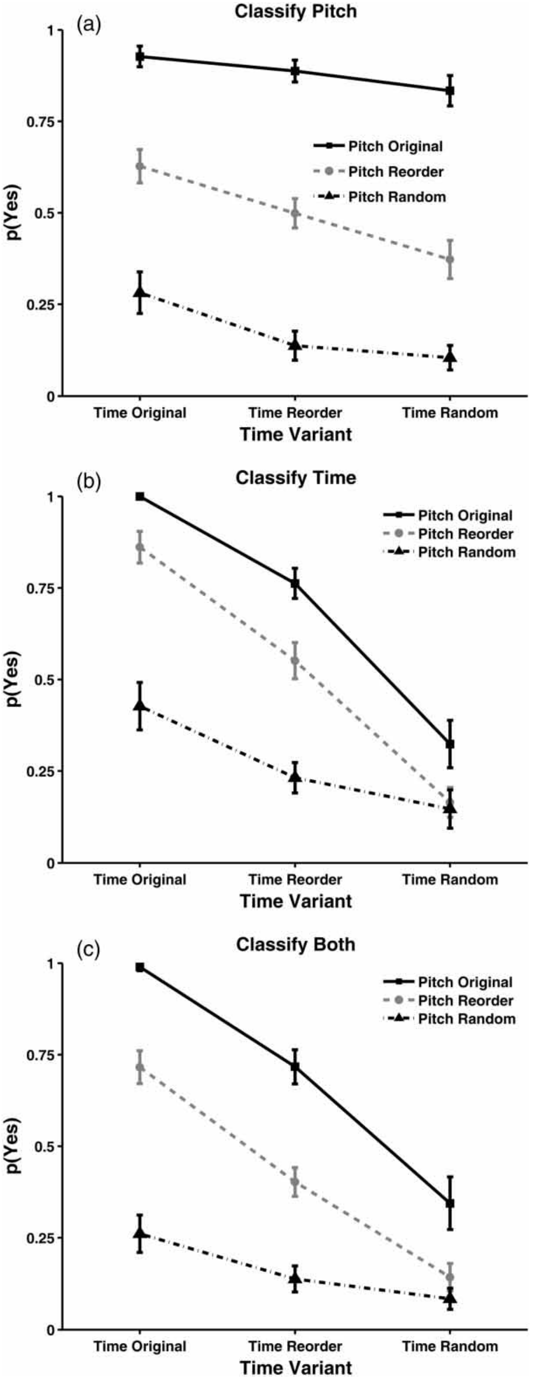

Figure 4. *Probability of classifications for melodies in each instructional condition of Experiment 3: (a) "classify pitch" condition; (b) "classify time" condition; (c) "classify both" condition. Error bars represent standard error of the mean.*

Experiments 1 and 2, the interaction between pitch and time was significant, $F(4, 84) = 20.93$, $MSE = .02$, $p < .001$, $\eta^2 = .055$ (shown in Figure 4b). Block order did not interact with pitch, $F(4, 42) = 1.97$, $MSE = .08$, $p = .12$, $\eta^2 = .022$, with time, $F(4, 42) = 2.41$, $MSE = .06$, $p = .06$, $\eta^2 = .020$, or with the pitch−time interaction, $F(4, 42) = 1.11$, $MSE = .02$, $p = .37$, $\eta^2 = .006$.

Like Experiments 1 and 2, and despite the task design and selective attention instructions, both pitch and time always affected classifications. Also, participants were able to emphasize either dimension in accordance with the instructions, as evidenced by the change in effect sizes of pitch and time across instruction. Unlike Experiments 1 and 2, where the effects of pitch and time combined linearly in the "rate time" instruction, there was an interaction between pitch and time for the "classify time" instruction (compare Figures 2b and 3b to Figure 4b).

Classify both instruction

The results of the "classify both" instruction revealed a main effect of pitch, $F(2, 42) = 122.19$, $MSE = .04$, $p < .001$, $\eta^2 = .374$, and time, $F(2, 42) = 58.51$, $MSE = .07$, $p < .001$, $\eta^2 = .297$, but not block order, $F(2, 21) < 1$, $MSE = 2.24$, $p = .58$, $\eta^2 = .003$. Figure 4c displays these data. There was an interaction between pitch and time, $F(4, 84) = 15.87$, $MSE = .03$, $p < .001$, $\eta^2 = .061$. The pitch−time interaction in the "classify time" and "classify both" instructions is similar to that observed in the "rate both" instruction of Experiments 1 and 2. That is, the effect of variation in one dimension was stronger as the other dimension adhered more closely to its original structure. Block order did not interact with pitch, $F(4, 42) < 1$, $MSE = .04$, $p = .57$, $\eta^2 = .004$, with time, $F(4, 42) < 1$, $MSE = .07$, $p = .10$, $\eta^2 = .004$, or with the pitch−time interaction, $F(8, 84) = 1.16$, $MSE = .03$, $p = .33$, $\eta^2 = .009$.

Discussion

As in Experiments 1 and 2, participants' responses followed the manipulations in pitch and temporal structural conformity within the stimuli. In fact, the overall pattern of data in Experiment 3 closely resembles that of the previous experiments. The data from all experiments showed essentially no difference between the second and third variant levels for pitch and time, with one exception for each (out of a possible 18). All experiments showed failures of selective attention when participants' instructions were to ignore either pitch or time. This result is particularly notable in Experiment 3 because the classification task design made the unattended dimension truly irrelevant in the selective attention conditions, as opposed to effortfully ignored (Experiments 1 and 2). Therefore, it appears that the contributions of pitch and time are involuntary in these two tasks. Another overlap in findings across experiment is that in the selective attention instructional conditions, the effect size of the dimensions varied according to the instructions. Additionally, for "rate both" and "classify both" instructions, there was always an interaction between pitch and time, and pitch was overall clearly the more influential dimension in all experiments.

Nevertheless, the differences between the experiments warrant further inspection. In the "rate pitch" instruction of Experiment 1, there was an interaction between pitch and block order, demonstrating that variations in tonal conformity influenced ratings more for those participants who rated pitch in their second block. In Experiment 2, block order interacted instead with time, again for the "rate pitch" condition. Neither of these interactions was significant in Experiment 3 (in fact, no interactions with block order emerged in Experiment 3), and neither lend themselves well to any particular theoretical interpretation.

The more interesting divergence in Experiment 3 is the presence of an interaction between pitch and time in classifications of metric conformity, where there was no such interaction in the temporal goodness ratings of the previous experiments. This interaction replicated the shape of the other pitch–time interactions (for "rate both" and "classify both" instructions), showing that as either dimension resembled more its original structural conformity, the effect of the other dimension increased. Accordingly, in these interactive cases, greater structural conformity in one dimension enhanced the ability to detect structural conformity in the other dimension, even when it was irrelevant to the task.

Overall, the variations in the pattern of findings across instruction and experiment suggest that differences in salience affected how the dimensions combine. Specifically, a more even matching of salience across dimensions fostered interactions between them; conversely larger mismatches in dimensional salience prevented such interactions. Table 1 illustrates this principle by listing the absolute difference score in effect size (i.e., $|\text{pitch } \eta^2 - \text{time } \eta^2|$) along with the strength of the interaction between the dimensions (pitch $\times$ time η^2) for all instructional conditions. For musically trained performers, when classifying metric conformity, the advantage in effect size for time, $|\text{pitch } \eta^2 - \text{time } \eta^2| = .157$ was smaller than the corresponding advantage when rating temporal goodness, $|\text{pitch } \eta^2 - \text{time } \eta^2| = .445$. Proportionately, the pitch–time interaction was more powerful when classifying time than when rating temporal goodness (interaction

Table 1. Differences in main effect sizes for pitch and time for each instruction and the corresponding effect size of the pitch–time interaction

Instruction	*Main effect size difference score* $|\text{pitch } \eta^2 - \text{time } \eta^2|$	*Interaction effect size* Pitch $\times$ Time η^2
Musically trained performers (Exp. 1)		
Rate pitch	.654	.004
Rate time	.445	.003
Rate both	.345	.014**
Musically trained performers (Exp. 3)		
Classify pitch	.663	.008
Classify time	.157	.055**
Classify both	.078	.061**
Musically untrained listeners (Exp. 2)		
Rate pitch	.410	.002
Rate time	.256	.004
Rate both	.084	.011*

$^*p < .01$. $^{**}p < .001$.

$\eta^2 = .055$ vs. $\eta^2 = .003$, respectively). Additionally, this observation may account for the increased effect size of the pitch–time interaction for the "classify both" instruction of Experiment 3 ($\eta^2 = .061$) versus the "rate both" instruction of Experiment 1 ($\eta^2 = .014$). The main effect sizes of pitch and time were more evenly matched in the "classify both" condition of Experiment 3 (difference score = .078) than in the "rate both" condition of Experiment 1 (difference score = .345), and there was a stronger interaction in Experiment 3. For musically trained performers, this inverse relation between the effect size difference score and pitch-time interaction effect size as shown in Table 1 is strong, $r(4) = -.90$, $p = .01$, even for such a small sample. Because of the replication of quantitative differences between musically trained and untrained listeners for explicitly musical tasks, the untrained listeners were examined separately (Experiment 2). In this population the same pattern emerged—as the main effect sizes of pitch and time became more evenly matched, interaction effect sizes grew (see Table 1). Overall, therefore, the principle of dimensional salience, explored initially in the context of single notes following a musical context (Prince, Schmuckler, et al., 2009; Prince, Thompson, et al., 2009) also applies to longer sequences such as entire melodies.

GENERAL DISCUSSION

Three experiments tested how the dimensions of pitch and time combine in the perception of complex melodies that varied in their degree of conformity to pitch and temporal structure. In the first experiment, musically trained performers rated the goodness of melodies based only on the sequence of pitches, durations, or both; in the second experiment, musically untrained listeners performed the same tasks. For Experiment 3, musically trained performers classified whether the melody adhered to typical pitch and/or temporal structure. Despite these task and participant differences, the results from all experiments were quite similar. There were main effects of both dimensions in all conditions, even for selective attention instructions; furthermore, these effects were consistent with the structural manipulations in the stimuli. For conditions in which the salience was more evenly matched between dimensions (as measured by main effect size), the two dimensions also showed a statistical interaction such that increasing structural conformity in either dimension exaggerated the effect of the other dimension. These interactions appeared in all experiments when participants responded on the basis of both dimensions, and also in the "classify time" instruction of Experiment 3.

There were joint contributions of both pitch and time in all conditions of these experiments. Moreover, the forms of these contributions were consistent with the degree of tonal and metric conformity—as the melody adhered more to its original pitch and/or temporal characteristics, goodness ratings and classifications of structural conformity increased. Neither instruction, nor task, nor whether there was a statistical interaction between the pitch and time variant levels affected this pattern. This finding is especially intriguing in the selective attention instructions, where structural conformity in the ignored/irrelevant dimension nevertheless influenced responses. Similar results have been reported elsewhere with respect to both dimensions. For instance, the tonal stability of individual events following a musical context can influence temporal change detection (Lebrun-Guillaud & Tillmann, 2007), phoneme monitoring (Bigand, Tillmann, Poulin, D'Adamo, & Madurell, 2001), and temporal classification (Prince, Thompson, et al., 2009). Similarly for time, the metric stability of an event can affect detection accuracy of a change in a pitch pattern (Jones, Johnston, & Puente, 2006), and pitch height comparison (Jones et al., 2006; Jones et al., 2002)—but only in a context lacking conformity to typical pitch structure (Prince, Schmuckler, et al., 2009). Therefore, it appears that irrelevant stimulus structure is automatically processed and can nonetheless affect perception, even when this operation confers no task benefit. Indeed, work in visual perception proposes exactly this concept—a coherent object can automatically attract attention when identifying the colour of a target object in a visual array (Kimchi, Yeshurun, & Cohen-Savransky, 2007).

An interesting extension of this notion arises from the ineffective manipulation of tonal and metric conformity between the second and third variant levels of pitch and time. That is, tonality and metre are automatically extracted from (and subsequently affect the processing of) stimuli that only weakly conform to such structures. The third variant levels were intended to decrease the degree of tonality and/or metre within the sequence while still retaining the original elements, and yet participants rated them as good as (and classified them as equally tonal or metric as) the second variant levels that preserved this structural conformity. The remaining vestiges of tonality and metre in the third variant level were sufficient to establish fully fledged corresponding percepts. Indeed, several reports show that to invoke these complex structures, only rudimentary levels of tonality (Cuddy & Badertscher, 1987; Krumhansl, 1990; Oram & Cuddy, 1995; Smith & Schmuckler, 2004) and metre (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Desain & Honing, 2003; Palmer & Krumhansl, 1990; Povel & Okkerman, 1981) are necessary. Given that impoverished levels of structural conformity can nonetheless invoke these organizational frameworks, listeners have most likely internalized the tonal and metric hierarchies over years of passive exposure and/or automatically use what little information there is in the stimulus to bootstrap the full structure. The difference between the first and second pitch variant levels (both of which exemplified tonality) is likely due to the disturbance of contour as a result of reordering, particularly the lack of stepwise motion in the second (and third) variant level as a result of order randomization. It is somewhat more challenging to account for the corresponding difference in the first and second time variant levels. One possible explanation is that the technique of correlating note onsets with the metric hierarchy did not capture entirely the contributing factors to temporal goodness/conformity. Indeed, the perception and reproduction of temporal sequences can be affected by other factors such as serial patterns (Jones, 1976), rhythmic consonance (Monahan & Carterette, 1987), and hierarchical simplicity (Povel & Essens, 1985). The reordering of elements in the second and third levels would destroy such information that presumably was present in the first variant level.

Even though listeners did not demonstrate sensitivity to the difference between the second and third variant levels, pitch and time always affected responses in every condition of all experiments. Interestingly, there was variation in how these dimensions combined—sometimes showing independent and additive properties, and other times demonstrating more complex interactive relations. As mentioned earlier, a good explanation for these patterns emerges from examining dimensional salience, indexed by the relative effect sizes of pitch and time as main effects. Interactions were more likely as the salience of pitch and time approached parity. Yet pitch remained the more salient dimension except when listeners attempted to consciously ignore it, and even in these conditions the advantage of time over pitch never approached the corresponding advantage of pitch over time when rating or classifying pitch. Previous work (Prince, Thompson, et al., 2009) found that when classifying the temporal characteristics of a note following a typical musical context, the tonal stability of the note influenced responses. However, variations in temporal location did not influence pitch classifications, even though both tasks were equally difficult. Thus, Prince, Thompson, et al. concluded that in such musical contexts, the dimension of pitch was more salient than that of time. In fact, follow-up work showed that only upon removing tonal pitch structure did time became sufficiently salient as to influence the processing of pitch (Prince, Schmuckler, et al., 2009).

However, many of the stimuli in the present experiments did not represent typical Western music in that they did not demonstrate tonality and metre. Accordingly, why was pitch consistently more salient than time? One possibility is that the nature of the tasks increased pitch salience. In previous work on single-note ratings of probe events, Prince, Thompson, et al. (2009) pointed out that responding to a pitched event inevitably makes the task inherently pitch based. In the context of the present experiments, listeners did not necessarily have to attend to the temporal properties of the melodies when attending selectively to its pitch characteristics (although they apparently did to some degree). However, when evaluating the temporal aspects of the stimuli, listeners had no choice but to listen to the pitches themselves in order to extract and process the temporal properties of the melody. Therefore, the tasks may have offered an intrinsic advantage to the dimension of pitch, because even the temporal instructions required processing of pitch on some level. Studies that use tasks that reverse this pattern of dependence (e.g., tapping to the beat of a melody) often show reverse patterns of dimensional salience, such that manipulations based on time predominate over pitch (London, Himberg, & Cross, 2009; Pfordresher, 2003; Snyder & Krumhansl, 2001). In these cases, the temporal position is of primary importance, and no pitch processing is necessary.

An issue that music cognition research faces regularly is the concern that the use of musically trained participants restricts the generality of the conclusions to that specific population. However, careful examination of the existing research shows that all encultured listeners perceive music remarkably similarly (Bigand & Poulin-Charronnat, 2006). With a specific view towards abstract organizational structures, both musically trained and untrained listeners show the same patterns of perception of tonality (Hébert, Peretz, & Gagnon, 1995; Koelsch, Gunter, Friederici, & Schröger, 2000; Loui, Wessel, & Kam, 2010; Tillmann & Poulin-Charronnat, 2010; Trainor & Trehub, 1994; van Egmond & Boswijk, 2007) and metre (Geiser, Ziegler, Jancke, & Meyer, 2009; Hannon et al., 2004). The present studies extend these general findings by revealing that musical training is not associated with any qualitative differences in pitch–time integration. Instead, it is more likely that the passive exposure to the statistical properties of music guides the development of music perception, regardless of formal musical training. Indeed, infant listeners (who therefore have not had the benefit of years of exposure to a musical culture) demonstrate more interactive patterns of pitch–time integration (Trehub & Hannon, 2009).

In fact, listeners encultured in Western music may have internalized a listening strategy that emphasizes pitch at the expense of time, thus creating a mental schema of pitch salience. Recent work on pitch–time integration suggests that cultural factors in Western music create a listening strategy that enhances the salience of pitch at the expense of time (Prince, Schmuckler, et al., 2009; Prince, Thompson, et al., 2009). This inference depends on two assertions: first, typical Western music emphasizes pitch over time; second, listeners passively learn these statistical properties through long-term exposure and internalize listening strategies that emphasize pitch.

The first of these assertions is uncontroversial, as discussed in the introduction. The second assertion requires some explanation. In any domain, perceivers learn to attend to the informative aspects of a stimulus at the expense of other aspects through experience provided by passive exposure (Bhatt & Quinn, 2011). Perhaps the most striking examples of this principle can be found in developmental studies of metre perception (Hannon & Trehub, 2005a, 2005b) and speech perception (Nazzi, Jusczyk, & Johnson, 2000; Werker & Tees, 1984). Essentially, although infants initially are able to learn a wide range of musical metres and are sensitive to all phonemic contrasts, variation that is not musically or linguistically meaningful in their culture is not informative. Over time, these listeners acquire strategies that maximize sensitivity to the meaningful information (separately for music and language), at the expense of sensitivity to other information. Accordingly, lifelong passive exposure to Western music, which features pitch above other musical dimensions, may cause listeners to internalize a strategy (i.e., mental schema) that reflects this imbalance, thus creating a salience of pitch. Importantly, this principle of pitch salience via passive exposure applies to all listeners, and not only musicians. Indeed, untrained listeners are nonetheless strikingly sensitive to abstract organizational principles of music (Bigand & Poulin-Charronnat, 2006). Also importantly, this listening strategy might only be invoked in the context of a stimulus that resembles Western music (Prince, Schmuckler, et al., 2009). However, this theory requires cross-cultural work comparing Western listeners to others encultured in musical systems that do not favour pitch, such as African, Eastern European, or Australian Aboriginal music.

What contributions do these findings offer? The ultimate aim of this research is to provide a theoretical framework on dimensional integration that can explain how stimulus dimensions combine to form a coherent percept and also to unify the complex pattern of findings specifically in the pitch–time integration literature. The current findings are relevant to existing theories of pitch–time integration. Regarding the local/global proposal (Bigand et al., 1999; Jones & Boltz, 1989; Tillmann & Lebrun-Guillaud, 2006), it is important to note that the present experiments always required attention to the entire melody (i.e., global tasks), yet the pattern of pitch–time integration varied. Instead, manipulating instructions affected whether interactions occurred, rather than showing overall evidence of interactions with these global tasks. Similarly for the early versus late processing stage theory of pitch–time interactions (Peretz & Kolinsky, 1993; Pitt & Monahan, 1987; Thompson et al., 2001), these experiments required cognitive processes that occur at a relatively late stage. However, again the pattern of pitch–time integration in the present data varied along the continuum of independent and interactive processing, not a homogenous pattern of interactive relations. Accordingly, explanations of pitch–time integration based on strong formulations of the local/global or stages of processing theories are unwarranted in the face of the present study. However, without further direct tests of these theories (i.e., finding independent and interactive contributions in local tasks and at early processing stages), the present data alone cannot discount these theories entirely.

These findings do support another possible unifying theory of dimensional integration, investigated here in the context of pitch–time integration. Assessing dimensional relations by their salience as well as relative perceptual difficulty (i.e., discriminability) shows promise for explaining many of the contradictory reports, both in the context of single notes and in longer sequences. Essentially, as salience and discriminability become more evenly matched, global (not asymmetric) interactions between these dimensions become more likely. It is important to note that salience can be dissociable from discriminability—Prince, Thompson, et al. (2009) demonstrated that in typical Western music, there was an imbalance in dimensional salience that caused asymmetric interference of pitch on time, even with equalized discriminability. Additionally, Garner's classic research on dimensional interactions showed that demonstrably independent dimensions can appear to interact when discriminability is unbalanced (Garner & Felfoldy, 1970). There are procedures for measuring discriminability, such as those provided by signal detection theory (Macmillan & Creelman, 1991), Stroop interference (Melara & Algom, 2003), and Garner interference (Garner & Felfoldy, 1970) on reaction times, but there are currently no methods to measure dimensional salience. However, the findings reported here suggest a possible objective measure of dimensional salience, based on comparing main effect sizes in selective attention conditions. In this case, relative main effect sizes ($|\text{pitch } \eta^2 - \text{time } \eta^2|$) between stimulus dimensions in selective attention conditions predicted the strength of an interaction ($\text{pitch} \times \text{time } \eta^2$) in conditions of attending both dimensions jointly. Thus, when equalized in terms of discriminability, the difference in main effect size between stimulus dimensions could provide an indicator of relative dimensional salience. Hopefully, these findings can guide future work on dimensional interactions and increase our understanding of the cognitive mechanisms of perception.

Original manuscript received 10 May 2010
Accepted revision received 28 February 2011
First published online day month year

REFERENCES

Aldwell, E., & Schacter, C. (2002). *Harmony and voice leading* (3rd ed.). Belmont, CA: Thomson/Schirmer.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, *93,* 154–179.

Atkinson, A. P., Tipples, J., Burt, D. M., & Young, A. W. (2005). Asymmetric interference between sex and emotion in face perception. *Perception and Psychophysics*, *67,* 1199–1213.

Berkowitz, S., Fontrier, G., & Kraft, L. (1997). *A new approach to sightsinging* (4th ed.). New York, NY: W. W. Norton & Company.

Bhatt, R. S., & Quinn, P. C. (2011). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, *16,* 2–38.

Bigand, E. (1997). Perceiving musical stability: The effect of tonal structure, rhythm, and musical expertise. *Journal of Experimental Psychology: Human Perception and Performance*, *23,* 808–822.

Bigand, E. (2003). More about the musical expertise of musically untrained listeners. In G. Avanzini, C. Faienza, D. Minciacchi, L. Lopez, & M. Majno (Eds.), *Neurosciences and music* (Vol. 999, pp. 304–312). New York, NY: New York Academy of Sciences.

Bigand, E., Madurell, F., Tillmann, B., & Pineau, M. (1999). Effect of global structure and temporal organization on chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, *25,* 184–197.

Bigand, E., & Poulin-Charronnat, B. (2006). Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition*, *100,* 100–130.

Bigand, E., Tillmann, B., Poulin, B., D'Adamo, D. A., & Madurell, F. (2001). The effect of harmonic context on phoneme monitoring in vocal music. *Cognition*, *81,* B11–B20.

Boltz, M. G., Marshburn, E., Jones, M. R., & Johnson, W. W. (1985). Serial-pattern structure and temporal-order recognition. *Perception and Psychophysics*, *37,* 209–217.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10,* 433–436.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science*, *14,* 362–366.

Castellano, M. A., Krumhansl, C. L., & Bharucha, J. J. (1984). Tonal hierarchies in the music of North India. *Journal of Experimental Psychology: General*, *113,* 394–412.

Cousineau, M., Demany, L., & Pressnitzer, D. (2009). What makes a melody: The perceptual singularity of pitch sequences. *The Journal of the Acoustical Society of America*, *126,* 3179–3187.

Cuddy, L. L., & Badertscher, B. (1987). Recovery of the tonal hierarchy—some comparisons across age and levels of musical experience. *Perception and Psychophysics*, *41,* 609–620.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1993). Harmonic accents in inference of metrical structure and perception of rhythm patterns. *Perception and Psychophysics*, *54,* 794–807.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1994). Inference of metrical structure from perception of iterative pulses within time spans defined by chord changes. *Music Perception*, *12,* 57–76.

Dawe, L. A., Platt, J. R., & Racine, R. J. (1995). Rhythm perception and differences in accent weights for musicians and nonmusicians. *Perception and Psychophysics*, *57,* 905–914.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception*, *32,* 341–365.

Dibben, N. (1994). The cognitive reality of hierarchical structure in tonal and atonal music. *Music Perception*, *12,* 1–25.

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

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

Dyson, B. J., & Quinlan, P. T. (2010). Decomposing the Garner interference paradigm: Evidence for dissociations between macrolevel and microlevel performance. *Attention, Perception & Psychophysics*, *72,* 1676–1691.

Eerola, T., Järvinen, T., Louhivuori, J., & Toiviainen, P. (2001). Statistical features and perceived similarity of folk melodies. *Music Perception*, *18,* 275–296.

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance*, *35,* 264–280.

Everett, W. (2000). Pitch down the middle. In W. Everett (Ed.), *Expression in pop-rock music: A collection of critical and analytical essays* (pp. 111–174). New York, NY: Garland.

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (1st ed., pp. 149–180). New York, NY: Academic Press.

Gagnon, L., & Peretz, I. (2000). Laterality effects in processing tonal and atonal melodies with affective and nonaffective task instructions. *Brain and Cognition*, *43,* 206–210.

Garner, W. R. (1974). *The processing of information and structure* (Vol. 203). Oxford, UK: Lawrence Erlbaum Associates.

Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, *8,* 98–123.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, *1,* 225–241.

Geiser, E., Ziegler, E., Jancke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex*, *45,* 93–102.

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *30,* 956–974.

Hannon, E. E., & Trehub, S. E. (2005a). Metrical categories in infancy and adulthood. *Psychological Science*, *16,* 48–55.

Hannon, E. E., & Trehub, S. E. (2005b). Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, *102,* 12639–12643.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4,* 223–233.

Hébert, S., & Peretz, I. (1997). Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? *Memory and Cognition*, *25,* 518–533.

Hébert, S., Peretz, I., & Gagnon, L. (1995). Perceiving the tonal ending of tune excerpts—the roles of preexisting representation and musical expertise. *Canadian Journal of Experimental Psychology–Revue Canadienne De Psychologie Experimentale*, *49,* 193–210.

Jones, M. R. (1976). Levels of structure in reconstruction of temporal and spatial serial patterns. *Journal of Experimental Psychology: Human Learning and Memory*, *2,* 475–488.

Jones, M. R. (1987). Dynamic pattern structure in music—recent theory and research. *Perception and Psychophysics*, *41,* 621–634.

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

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

Jones, M. R., Johnston, H. M., & Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cognitive Psychology*, *53,* 59–96.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13,* 313–319.

Kessler, E. J., Hansen, C., & Shepard, R. N. (1984). Tonal schemata in the perception of music in Bali and in the west. *Music Perception*, *2,* 131–165.

Kimchi, R., Yeshurun, Y., & Cohen-Savransky, A. (2007). Automatic, stimulus-driven attentional capture by objecthood. *Psychonomic Bulletin & Review*, *14,* 166–172.

Koelsch, S., Gunter, T. C., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12,* 520–541.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York, NY: Oxford University Press.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126,* 159–179.

Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89,* 334–368.

Krumhansl, C. L., & Schmuckler, M. A. (1986, July). *Key-finding in music: An algorithm based on pattern matching to tonal hierarchies*. Paper presented at the 19th Annual Meeting of the Society of Mathematical Psychology, Cambridge, MA.

Lebrun-Guillaud, G., & Tillmann, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception and Psychophysics*, *69,* 1450–1459.

Lerdahl, F. (2001). *Tonal pitch space*. New York, NY: Oxford University Press.

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

London, J., Himberg, T., & Cross, I. (2009). The effect of structural and performance factors in the perception of anacruses. *Music Perception*, *27,* 103–120.

Loui, P., Wessel, D. L., & Kam, C. L. H. (2010). Humans rapidly learn grammatical structure in a new musical scale. *Music Perception*, *27,* 377–388.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York, NY: Cambridge University Press.

Mathworks. (2004). Matlab [Computer software]. Natick, MA: Author

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110,* 422–471.

Monahan, C. B., & Carterette, E. C. (1987). The influence of temporal contour and interval information on the discrimination of rhythmic patterns. *Journal of the Acoustical Society of America*, *81,* S91–S92.

Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech-perception. *Perception and Psychophysics*, *47,* 379–390.

Nazzi, T., Jusczyk, P. W., & Johnson, E. K. (2000). Language discrimination by English-learning 5-month-olds: Effects of rhythm and familiarity. *Journal of Memory and Language*, *43,* 1–19.

Oram, N., & Cuddy, L. L. (1995). Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research*, *57,* 103–118.

Ottman, R. W. (1986). *Music for sight-singing* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

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

Palmer, C., & Krumhansl, C. L. (1987b). Pitch and temporal contributions to musical phrase perception—effects of harmony, performance timing, and familiarity. *Perception and Psychophysics*, *41,* 505–518.

Palmer, C., & Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance*, *16,* 728–741.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6,* 688–691.

Peretz, I., & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination—a neuropsychological perspective. *Quarterly Journal of Experimental Psychology*, *46A,* 301–325.

Pfordresher, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception*, *20,* 431–464.

Pitt, M. A., & Monahan, C. B. (1987). The perceived similarity of auditory polyrhythms. *Perception and Psychophysics*, *41,* 534–546.

Pomerantz, J. R. (1983). Global and local precedence—selective attention in form and motion perception. *Journal of Experimental Psychology: General*, *112,* 516–540.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception*, *2,* 411–440.

Povel, D. J., & Okkerman, H. (1981). Accents in equitone sequences. *Perception and Psychophysics*, *30,* 565–572.

Prince, J. B., Schmuckler, M. A., & Thompson, W. F. (2009). The effect of task and pitch structure on pitch–time interactions in music. *Memory and Cognition*, *37,* 368–381.

Prince, J. B., Thompson, W. F., & Schmuckler, M. A. (2009). Pitch and time, tonality and meter: How do musical dimensions combine? *Journal of Experimental Psychology: Human Perception and Performance*, *35,* 1598–1617.

Schenker, H. (1979). *Free composition (Der freie satz)*, E. Ostler, Trans. New York, NY: Longman, (Original work published 1935).

Schmuckler, M. A. (2004). Pitch and pitch structures. In J. Neuhoff (Ed.), *Ecological psychoacoustics* (pp. 271–315). San Diego, CA: Elsevier Science.

Schweinberger, S. R., Burton, A. M., & Kelly, S. W. (1999). Asymmetric dependencies in perceiving identity and emotion: Experiments with morphed faces. *Perception and Psychophysics*, *61,* 1102–1115.

Silbert, N. H., Townsend, J. T., & Lentz, J. J. (2009). Independence and separability in the perception of complex nonspeech sounds. *Attention, Perception & Psychophysics*, *71,* 1900–1915.

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of music*. Oxford, UK: Clarendon Press.

Smith, N. A., & Schmuckler, M. A. (2004). The perception of tonal structure through the differentiation and organization of pitches. *Journal of Experimental*

Psychology: Human Perception and Performance, *30,* 268–286.

Snyder, J. S., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, *18,* 455–489.

Temperley, D. (2001). *The cognition of basic musical structures*. Cambridge, MA: MIT Press.

Thompson, W. F., Hall, M. D., & Pressing, J. (2001). Illusory conjunctions of pitch and duration in unfamiliar tone sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *27,* 128–140.

Tillmann, B., & Lebrun-Guillaud, G. (2006). Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. *Psychological Research*, *70,* 345–358.

Tillmann, B., & Poulin-Charronnat, B. (2010). Auditory expectations for newly acquired structures. *Quarterly Journal of Experimental Psychology*, *63,* 1646–1664.

Tong, Y. X., Francis, A. L., & Gandour, J. T. (2008). Processing dependencies between segmental and suprasegmental features in mandarin Chinese. *Language and Cognitive Processes*, *23,* 689–708.

Trainor, L. J., & Trehub, S. E. (1994). Key membership and implied harmony in Western tonal music: Developmental perspectives. *Perception and Psychophysics*, *56,* 125–132.

Trehub, S. E., & Hannon, E. E. (2009). Conventional rhythms enhance infants' and adults' perception of musical patterns. *Cortex*, *45,* 110–118.

Trehub, S. E., Schellenberg, E. G., & Kamenetsky, S. B. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception and Performance*, *25,* 965–975.

Treisman, A. M., & Gelade, G. (1980). Feature-integration theory of attention. *Cognitive Psychology*, *12,* 97–136.

van Egmond, R., & Boswijk, M. (2007). The perception of key: The role of music training. *Music Perception*, *25,* 31–41.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior & Development*, *7,* 49–63.

White, B. W. (1960). Recognition of distorted melodies. *American Journal of Psychology*, *73,* 100–107.