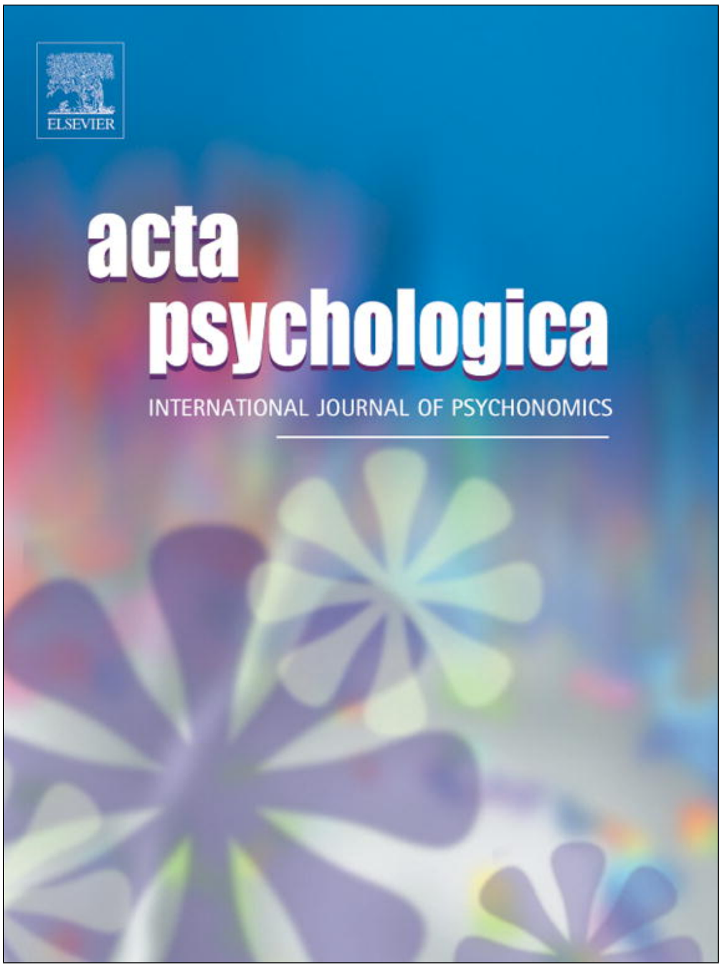

(This is a sample cover image for this issue. The actual cover is not yet available at this time.)

Acta Psychologica 141 (2012) 184–198

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/ locate/actpsy

The role of pitch and temporal diversity in the perception and production of musical sequences

Jon B. Prince [a,*], Peter Q. Pfordresher [b]

[a] Murdoch University, Australia
[b] University at Buffalo, State University of New York, United States

ARTICLE INFO

Article history:
Received 10 January 2012
Received in revised form 24 July 2012
Accepted 26 July 2012
Available online xxxx

PsycINFO codes:
2326
2330
2340

Keywords:
Music cognition
Pitch
Time
Music performance

ABSTRACT

In two experiments we explored how the dimensions of pitch and time contribute to the perception and production of musical sequences. We tested how dimensional diversity (the number of unique categories in each dimension) affects how pitch and time combine. In Experiment 1, 18 musically trained participants rated the complexity of sequences varying only in their diversity in pitch or time; a separate group of 18 pianists reproduced these sequences after listening to them without practice. Overall, sequences with more diversity were perceived as more complex, but pitch diversity influenced ratings more strongly than temporal diversity. Further, although participants perceived sequences with high levels of pitch diversity as more complex, errors were more common in the sequences with higher diversity in time. Sequences in Experiment 2 exhibited diversity in both pitch and time; diversity levels were a subset of those tested in Experiment 1. Again diversity affected complexity ratings and errors, but there were no statistical interactions between dimensions. Nonetheless, pitch diversity was the primary factor in determining perceived complexity, and again temporal errors occurred more often than pitch errors. Additionally, diversity in one dimension influenced error rates in the other dimension in that both error types were more frequent relative to Experiment 1. These results suggest that although pitch and time do not interact directly, they are nevertheless not processed in an informationally encapsulated manner. The findings also align with a dimensional salience hypothesis, in which pitch is prioritised in the processing of typical Western musical sequences.

1. Introduction

In auditory cognition, the dimensions of pitch and time are critical to defining an object. Kubovy (1981; Kubovy & Van Valkenburg, 2001), posits that pitch (as the psychological correlate of frequency) and time (patterns of event duration and onset) are in fact indispensable attributes—that differentiation along these dimensions is necessary to distinguish between two objects (i.e., achieve perceptual numerosity). Indeed, the ability to perceive, recognise, and perform a musical sequence depends on the preservation of distinct patterns of change in pitch (sometimes referred to as *melody*), and temporal patterns formed by ratio relationships among inter-onset intervals (sometimes referred to as *rhythm*). Rescaling these patterns in absolute terms (within musically reasonable bounds) does not interfere with recognition; rescaling pitch amounts to a change in key, and rescaling rate amounts to a change in tempo. Likewise, changes in timbre or absolute loudness levels do not affect melody recognition. However, when patterns of relative pitch or relative timing are distorted, recognition becomes difficult or even impossible (Hébert & Peretz, 1997; Jones & Ralston, 1991; White, 1960). Preserving sequential patterns along one dimension is not always sufficient for recognition. Consider, for instance, the first five notes of "The first Noel", versus "Mary had a little lamb"—these notes are identical with respect to their pitch pattern and are distinct only by virtue of rhythm (Palmer, personal communication). Other examples involve comparing the Dragnet theme with the main motif of Schubert's 8th symphony, and failures in recognition of the Bernstein tune "America" with changes to its rhythmic framework (Monahan, 1993).

Given the essential role of pitch and time in the mental representation of music, the way in which these dimensions combine is a critical question in music cognition. Yet despite much research there is no clear answer in the literature (for reviews, see Ellis & Jones, 2009; Krumhansl, 2000; Prince, Thompson, & Schmuckler, 2009). Many results suggest that pitch patterns contribute to perception independently of time patterns, predicting additive contributions of pitch and time. Initial demonstrations of independence of pitch and time come from ratings of melodic completion judgments (Palmer & Krumhansl, 1987a, 1987b). In these experiments, ratings from isochronous (neutralise temporal information, preserve melody) and monotonic (neutralise melodic information, preserve rhythm) versions of melodies made additive contributions in predicting ratings from an intact (both pitch and time), as well as phase-shifted, versions of the same melodies.

* Corresponding author at: School of Psychology, Murdoch University, 90 South Street, Murdoch, WA 6150, Australia. Tel.: +61 8 9360 6670; fax: +61 8 9360 6492.
E-mail address: j.prince@murdoch.edu.au (J.B. Prince).

0001-6918/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.actpsy.2012.07.013

Some authors support a modular approach to pitch–time combination—that the dimensions are processed in cognitively and neurally separate modules that only integrate information at later stages in perceptual processing (Peretz & Coltheart, 2003; Thompson, Hall, & Pressing, 2001). However, other findings suggest interaction, with properties of pitch patterns influencing one's ability to perceive patterns of time and vice-versa. For instance, listeners are better able to detect deviations in pitch when the rhythmic pattern highlighted the temporal position of the changed note (Jones, Boltz, & Kidd, 1982), and the relation between rhythm and melody patterns can influence the perception of melodic completion and overall duration (Boltz, 1989; Jones & Boltz, 1989).

1.1. Dimensional salience

Due to the lack of consensus on the exact nature of how pitch and time combine in perception, several authors have explored the idea that the relation between these dimensions is not fixed but instead is flexible and can change on the basis of stimulus and task factors (e.g., Prince, 2011; Tillmann & Lebrun-Guillaud, 2006). Prince, Thompson, et al. (2009) proposed *dimensional salience* as a new framework for understanding how pitch and time combine. Dimensional salience refers to the prioritisation of one stimulus dimension in perceptual processing, which leads to more effective encoding of all information defined along that dimension. A dimension with higher salience contributes more strongly to the mental representation of the stimulus, providing a structure (i.e., schema) on which to encode information from additional dimensions. Salience therefore enhances sensitivity to a dimension, such that listeners are better able to recognise and retrieve information defined along it. Essentially, more salient dimensions are processed better. Importantly, the salience of a dimension is independent of its perceptual difficulty relative to other dimensions. That is, even after equating dimensions in terms of discriminability (or other equivalent perceptual measure), a more salient dimension will still be disproportionately emphasised in perceptual processing. Accordingly, one result of this prioritisation is that a more salient dimension is likely to interfere with (and less likely to experience interference from) a less salient dimension.

There are a number of possible sources of dimensional salience, including inherent differences across dimensions, characteristics of the stimulus, task and pre-learned schemas. For example, as the mental representation of a stimulus must, by definition, include indispensable attributes, such dimensions will necessarily be more salient than other dimensions of less central importance. Also, stimulus dimensions that feature greater informative value are likely to exhibit preferential status in perceptual processing, both in the domain of audition (Prince, Thompson, et al., 2009; Warrier & Zatorre, 2002) and vision (Ellison & Massaro, 1997; Melara & Algom, 2003). Further, task design may influence the salience of a dimension; for example, inherently temporal tasks such as tapping to a beat may highlight time over pitch (Pfordresher, 2003; Snyder & Krumhansl, 2001), whereas pitch-based tasks such as judging the goodness of a note or melody may favour pitch (Prince, 2011; Prince, Thompson, et al., 2009). Lastly, a pre-learned schema such as a culture-specific hierarchical pitch structure may influence dimensional salience in a musical context, such that when activated, pitch becomes more salient and reduces the influences of temporal manipulations on pitch judgements (Prince, Schmuckler, & Thompson, 2009).

1.2. Pitch salience in music

Some evidence from perceptual studies suggests that in musical sequences, pitch patterns may be more salient to listeners than temporal patterns (e.g., Bigand, 1997; Cousineau, Demany, & Pressnitzer, 2009; Dawe, Platt, & Racine, 1994; Eerola, Jarvinen, Louhivuori, & Toiviainen, 2001; Prince, 2011). When temporal variability (rhythm) is neutralised, melody recognition does not deteriorate as much as when pitch variability (melody) is neutralised (Hébert & Peretz, 1997). Furthermore, when listeners classified the timing of probe tones relative to a preceding musical context, the pitch class of probe tones was found to affect classification of durations, but timing did not influence pitch classifications, despite equal discriminability of exemplars within the dimensions of pitch and time (Prince, Thompson, et al., 2009). Metrical grouping of sequences varying in pitch and time showed much larger effects of pitch than time, after having been equated for strength of grouping induction in a baseline experiment (Ellis & Jones, 2009). Finally, goodness ratings of melodies varying in their degree of conformity to typical pitch and temporal structure show stronger effect sizes of pitch than of time (Prince, 2011). As such, in many cases there may be an inherently stronger role for pitch than for timing in Western musical contexts, and differences across studies may simply reflect the salience of pitch patterns relative to temporal patterns.

If pitch is more salient in music perception, what may be its basis? One factor has to do with the degree of informative value typically associated with these dimensions. At the simplest level, the amount of informative value could be the number of different categories used in each dimension in a given exemplar t. Whereas tonal melodies routinely present all categories of the diatonic scale and even non-diatonic tones (Järvinen, 1995; Knopoff & Hutchinson, 1983; Krumhansl, 1990), frequency counts of duration are much less diverse, typically within the range of 2–3 durations (Fraisse, 1982). As such, salience of pitch may derive from a higher relative usefulness than time in information-theoretic terms. In support of this explanation, effects of time on pitch (suggesting decreased pitch salience) are more likely when pitch patterns do not conform to standard tonal conventions of Western music (Prince, Schmuckler, et al., 2009).

It is possible, of course, that within a musical context, a schematic salience of pitch may result from the auditory system gradually adjusting the weighting of these dimensions based on the presence of greater pitch diversity than temporal diversity in musical sequences. As a rule, perceptual systems learn through experience to prioritise sources with more informative value (Goldstone, 1998), in vision (Bhatt & Quinn, 2011), speech (Werker & Tees, 2005), and music (Hannon, Soley, & Ullal, 2012). Accordingly, dimensional salience of pitch may reflect the priority of processing given to this more diverse dimension of musical sequences. In experimental terms, the imbalance between pitch and time typically found in musical patterns is problematic because of the possibility that pitch salience may be an artefact of diversity, a matter we turn to next.

1.3. Dimensional diversity

In the present research, diversity refers to the range of categorical values represented along each dimension within a stimulus. As mentioned in the previous section, the diversity on the dimension of time is typically lower in musical sequences (often just 2 IOI categories) than for pitch (often 7 categories).[1] Such observations are problematic for a dimensional salience hypothesis because previous reports of the relative importance of pitch may be due to the diversity of pitch classes in the immediate sequence, rather than an inherent, schematic salience of pitch. Research in both auditory and visual perception has demonstrated that the diversity of a stimulus dimension (how many unique values are presented along that dimension) affects how it combines with other dimensions (Melara & Mounts, 1994; Pansky & Algom, 1999; Sabri, Melara, & Algom, 2001). Specifically, dimensions with high diversity may dominate in perceptual processing and can thus create asymmetric interactions with dimensions of

[1] Disregarding slight fluctuations in pitch and time that reflect nuances like expressive time or noise in the system that can affect intonation (pitch) or the regularity of motor movements.

lower diversity. Unequal diversity across pitch and temporal dimensions may therefore change how they combine in auditory patterns. However this factor has not been controlled in much of the perceptual or production research on how pitch and time combine, although some have attempted to control the frequency and organisation of accents that arise from changes to pitch versus time (e.g., Ellis & Jones, 2009; Jones & Pfordresher, 1997; Pfordresher, 2003). One exception is Prince, Thompson, et al. (2009), who found dominance of pitch in a perceptual task in contexts of both equal and unequal dimensional diversity. However, that study was not designed to assess the effects of varying diversity across the dimensions of pitch and time.

In the experiments reported here, we systematically varied dimensional diversity in musical sequences by manipulating the number of different pitch classes (from the diatonic major scale) as well as the number of IOI categories (based on levels of the metrical hierarchy) that were present in a given sequence. This manipulation served two interlinked purposes. First, it allowed us to test for pitch salience in a way that is not confounded by diversity. Moreover, we can assess whether there are certain levels of diversity in pitch and time that lead to equivalent performance (cf. Ellis & Jones, 2009). Second, the effect of varying diversity across many levels provides a measure of salience. The dimensional salience hypothesis predicts that in the context of typical Western music, the magnitude of the effect associated with diversity will vary across the dimensions of pitch and time. Specifically, we predict that participants will be better able to process different levels of pitch diversity than temporal diversity. As a result, pitch diversity will have a stronger effect than temporal diversity, because the salience of pitch will enhance participants' awareness of differences in diversity along this dimension.

1.4. Perception versus production

Thus far, the evidence for pitch salience in musical sequences comes exclusively from perceptual tasks. Applied to the domain of music production, the dimensional salience hypothesis predicts better retrieval of pitch information than temporal information during performance of typical Western music, particularly when diversity is controlled. If pitch has prioritised status in a performance context, performers' mental representation of a musical sequence should rely primarily on pitch information. When accessing this largely pitch-based representation during performance, the temporal features will not enjoy the same fidelity. Accordingly, performers would be more likely to recall and produce accurately the sequence of pitches than durations, especially if the sequence is complex and difficult. Indeed, in a challenging performance situation performers would likely sacrifice temporal accuracy to maximise accuracy of the prioritised information (pitch). Furthermore, the fact that melodies typically exhibit greater pitch than temporal diversity may cause melodies that have (unusually) high levels of temporal diversity to seem unusual and thus difficult to encode or reproduce. Comparatively, high levels of pitch diversity would be expected, and much easier to produce accurately. Overall, therefore, dimensional salience predicts that temporal errors are more likely than pitch errors.

In contrast to the dimensional salience prediction and the perceptual work described earlier, the few existing findings on pitch–time combination in performance suggest a dominance of time over pitch. Using piano performance of notated melodies, Drake and Palmer (1993) explored the way in which melodic and temporal accents influence the timing of keyboard production when these accents aligned or conflicted in their temporal position. This research suggested that temporal accents dominated melodic accents in measures of intensity, timing, and articulation of piano performance. Later work explored the occurrence of pitch errors and temporal errors (separately) across successive piano performances (Drake & Palmer, 2000). At early stages, temporal errors were more frequent, but with practise there were proportionately more pitch errors than temporal errors, suggesting greater dominance of temporal information during production for trained performers. Similar findings suggesting temporal dominance have emerged from research in which participants synchronise with the rhythm of a sequence, or beat (Snyder & Krumhansl, 2001), suggesting that temporal information may be prioritised for tasks that require repetitive periodic movements, which occurs in both music performance and sensorimotor synchronisation. However, more recent research from Ellis and Jones (2009) suggests that such temporal dominance effects may be based on an imbalance of accent strength across dimensions, factors not controlled in the studies mentioned here.

Beyond the issue of which dimension is dominant in performance, there is also the question of whether pitch and time are independent or interactive in this context. Given that the goal of the production system is to integrate information into a single action command (e.g., a key press on a keyboard), one might expect more interactive relations between pitch and time in performance. Some evidence suggesting interaction in performance comes from the aforementioned error analyses reported by Drake and Palmer (2000). They found that the probability of joint pitch/duration errors was greater than that predicted by a multiplicative combination of separately occurring pitch and temporal errors, as would be predicted by a model based on independence. Drake, Dowling, and Palmer (1991) tested child singing/tapping and adult pianist performance in reproducing melodies while manipulating melodic accents (contour changes), rhythmic grouping (duration change on final note of group), and intensity (increasing loudness of a note). Both melodic and rhythmic changes affected pitch accuracy (equivalently), but not temporal accuracy. Consistent with the assumption of integration, (pitch) performance accuracy was best when accent types were concordant, and worst when they conflicted.

Although there is some (albeit sparse) research addressing the contributions of pitch and time to the reproduction of melodies in performance, to our knowledge there is no work comparing pitch–time combination systematically across perception and production of music. Comparisons between perception and production are nevertheless both relevant and theoretically important. Based on the assumption that perception must precede performance (with the possible exception of improvisation), perceptual relations between pitch and time likely carry over from perception to production. This notion resonates with recent theories proposing a common shared representation for perception and action (Hommel, Musseler, Aschersleben, & Prinz, 2001). There is also interest in perception–production relations specifically within the musical domain (Repp, 1998, 2005; Repp & Knoblich, 2007; Repp, London, & Keller, 2011). Some of this research suggests dissociations between the systems that support music perception and performance (Loui, Guenther, Mathys, & Schlaug, 2008; Pfordresher & Brown, 2007; Repp, et al., 2011; Zatorre, Chen, & Penhune, 2007). Thus addressing pitch–time combination in the context of perception–production relations offers an important contribution to several areas of enquiry.

1.5. Current experiments

The aforementioned issues motivate systematic investigation of pitch–time combination in music perception and production. Due to a lack of shared methodology, a deeper understanding of the relation between these dimensions requires direct and systematic comparisons between perception and production tasks, using a common set of stimuli and similar designs. In particular, it is important to assess the role of dimensional salience in both the perception and production of musical sequences. Thus the main goal of this research was to explore how dimensional diversity influences the contribution of pitch and time in the perception and production of musical sequences.

To accomplish this goal, pianists and other musicians heard melodies that varied in their degree of pitch and temporal diversity, achieved by manipulating the number of unique pitch classes and inter-onset intervals (IOIs) in each sequence. Pianists reproduced the melodies from memory after hearing the melody as many times as they wished, and then rated how difficult the melody was to perform. The other musicians rated the complexity of the melodies but did not perform them. In Experiment 1, melodies varied only in pitch (isochronous) or time (monotonic). In Experiment 2, various levels of dimensional diversity were recombined, and a new set of participants performed the same tasks. For both experiments, the dependent measures were pitch error rate, temporal error rate, number of repetitions heard (prior to performance), difficulty rating, and complexity. The pianists completed all but the complexity rating task, whereas the other sample of musicians completed only the complexity rating task. The main experimental question was how the changes in dimensional diversity would affect the dependent measures, and if perception (complexity rating, difficulty rating), and production measures (pitch and temporal error rate) would yield similar patterns.

2. Experiment 1

In Experiment 1 we tested the effects of pitch and temporal diversity separately, by presenting melodies that varied only in one dimension in each trial. Two groups of musically trained participants were assigned to either perceptual or production tasks. The perception group rated the complexity of heard melodies. The production group first listened to and then reproduced melodies without viewing notation (they were allowed to listen as many times as they wished), and afterwards rated the difficulty of the melody. This design maximises comparability between perception and production tasks by only providing an auditory input to complete the task, in contrast to providing visual information (notation), or learned motor plans created by allowing a practise period. Diversity of pitch patterns was manipulated by varying the number of pitch categories present from 1 to 7, all within the major diatonic scale system. Diversity of temporal patterns was manipulated by varying the number of durational categories within a binary metrical framework. Analyses focused on the effects of diversity across dimensions of pitch and time, separately for perception and production.

Experiment 1 was designed to address two critical questions. First, would both pitch and temporal dimensional diversity have similar effects? Previous results suggesting pitch salience lead to the prediction that the effects of varying diversity along each dimension would differ. Second, would the effects of dimensional diversity be similar for perception and production? As discussed earlier (Section 1.4), results of previous studies differ with respect to support for “shared representations” for perception and action (hinting at similar effects) whereas other data suggest dissociations across the perception and action systems (implying distinct effects).

2.1. Method

2.1.1. Participants

There were two groups of participants in this experiment: a production group (18 musicians whose primary instrument was the piano) and a perception group (18 musicians with no restrictions on primary instrument). All participants were students from the University at Buffalo. The average age of participants in the production group was 20.8 ($SD = 5.2$), and they had an average of 11.3 years of formal training in piano performance ($SD = 3.9$). In the perception group, the average age was 19.2 ($SD = 1.5$), and they had an average of 10.2 years of training on their instrument ($SD = 2.1$). The two groups did not differ in years of training, $t(17) = 1.1$, $p = .3$.

2.1.2. Stimuli

Stimuli were created by composing 12 variations on each of 9 “seed” melodies that were selected from a set of sightsinging melodies (Ottman, 1986). These seed melodies stayed within a major tonality (using all 7 diatonic pitches), remained within a single octave, and were in duple meter. Variants were designed to mimic the practice of deriving reductions of typical musical structure (cf. Lerdahl & Jackendoff, 1983; Schenker, 1935/1979), in order to create patterns with systematically varying diversity of pitch and time that nonetheless all shared a common deep syntactic structure. Of the 12 variants, 6 varied in the number of unique pitch classes used (2 to 7) while being isochronous, and the other 5 varied in the number of unique inter-onset intervals (IOIs) while being monotonic. The final variant was both isochronous and monotonic. Each variant was preceded by a cadence in a major key using four chords, and lasted for 16 beats (crotchets). Table 1 shows the pitch and IOI settings for each variant. Pitch values are shown as scale degree, IOI values are shown as denominations within a duple metrical framework. Original seed melodies were not used as stimuli.

The choice of pitch classes and IOIs across diversity levels was designed to maximise the tonal and metrical stability within each level of diversity, and to eliminate confounds between diversity and pitch height, or diversity and IOI range. Thus the first pitch variant started with the 1st scale degree (tonic), the second variant added the 5th scale degree (dominant), the third added the 3rd scale degree (mediant), and so on. Similarly, the temporal variants began with the tactus (crotchet; also the duration of each chord in the cadence preceding the sequence), and added shorter and longer IOIs at increasing levels of diversity.[2]

Pitch and temporal variants were constructed so as to preserve correlations with the tonal and metric hierarchies (Krumhansl & Kessler, 1982; Palmer & Krumhansl, 1990). For pitch variants 2 to 7, the duration profile (summing all occurrences across the entire variant) of each of the 12 pitch classes was correlated with the major

Table 1
Pitch classes and IOIs used in melody variants.

Diversity	Pitch classes (scale degree)	IOIs (duration denomination)
1 (pitch/time)	1	♩
2 (pitch)	1, 5	♩
3 (pitch)	1, 3, 5	♩
4 (pitch)	1, 2, 3, 5	♩
5 (pitch)	1, 2, 3, 5, 6	♩
6 (pitch)	1, 2, 3, 5, 6, 7	♩
7 (pitch)	1, 2, 3, 4, 5, 6, 7	♩
2 (time)	1	♪ ♩
3 (time)	1	♪ ♩ ♩.
4 (time)	1	𝅘𝅥𝅯 ♪ ♩ ♩.
5 (time)	1	𝅘𝅥𝅯 ♪ ♩ ♩. 𝅗𝅥
6 (time)	1	𝅘𝅥𝅯 ♪ ♪. ♩ ♩. 𝅗𝅥

[2] It is possible that the diversity manipulation is not equally musically meaningful across dimensions (e.g., does the difference between scale degrees 1 and 5 equal that between a crotchet and quaver?), but focusing on the variable of diversity necessitates selecting some value. To minimise musically important differences across dimensions we aimed to increase the diversity of pitches/durations in a manner that progressed from higher levels of the tonal/metric hierarchies (see text) to lower levels, while still avoiding confounding of pitch height with pitch class. Thus the second level of diversity picked the most stable members of the tonal hierarchy and the corresponding durations from the metric hierarchy. Regardless, the issue of whether the tonic is equivalent to the tactus remains an interesting question for further research.

tonal hierarchy (Krumhansl & Kessler, 1982; Krumhansl & Schmuckler, 1986), yielding average values of .81, .86, .85, .87, .87, and .94, respectively. Similarly, for the temporal variants 2 to 6, the frequency of occurrence of each metric position within a 4/4 measure was correlated with the metric hierarchy (Palmer & Krumhansl, 1990), giving values of .86, .84, .82, .80, and .80, respectively. There were no reliable pairwise differences among these correlations, despite the nominal pattern. Accordingly, cues for tonality and meter were equally distributed across variants and did not vary with diversity.

The musical key (as established by the chord cadence) was not the same across all melodies, however it remained consistent within a melody. That is, all variants from a single seed melody were in the same key, but it could either be C, F, or G major. Additionally, the crotchet IOI was varied across melody, resulting in three different tempi—86 bpm (698 ms IOI), 92 bpm (652 ms IOI), and 100 bpm (600 ms IOI). As with the key manipulation, all variants from a single seed melody retained the same tempo. Fig. 1 depicts an example seed melody and the 12 variants derived from it.

2.1.3. Apparatus

Melody variants were constructed as MIDI files using Finale Songwriter 2010, and converted to .wav format using a piano soundfont in MIDI Converter Studio 6.1. For the perception task, MATLAB was used to program the experimental interface, which was presented on a Macintosh G5. Participants wore Sennheiser HD280pro headphones to listen to each melody. For the production task, participants listened to the melodies using Sony MDR-7500 professional headphones and performed the melodies using a FATAR CMK 49 unweighted keyboard. The experimental interface was programmed in C-shell scripts, and FTAP (Finney, 2001) was used to record the participants' performances.

2.1.4. Procedure

Participants in the perception task heard a melody and then rated how "complex" (also described as "complicated, difficult") it was, on a scale of 1 to 7. Participants completed 4 practise trials prior to rating the entire set (all 12 variants from all 9 seed melodies = 108 trials) in a randomised order; the procedure took about 40 min.

For the production task, participants heard the melody and could ask to hear it again as many times as they would like (the total number of repetitions were recorded). Participants were informed of the key of the melody, such that they would know the starting note of each sequence without possessing absolute pitch. During the listening phase, participants were not allowed to vocalise, tap, finger, or in any way practise performing the melody. Once a participant indicated readiness to perform, they attempted to reproduce the melody on the piano keyboard without stopping to correct mistakes. Upon completion of the performance, participants entered a rating indicating the perceived difficulty of reproducing the melody, using the piano keyboard, where the leftmost key indicated the easiest possible production, and the rightmost indicated the hardest production (MIDI note range 36 to 84). Participants did not hear a pitch in response to their selection. The experiment then progressed to the following melody; melodies were blocked such that participants performed all 12 variants of a single melody (in a randomised order) before progressing to the 12 variants of the next melody. This arrangement means that participants' recall may have benefited from similarity across variants within a particular melody, potentially interfering with performance measures across variant type. However, variants were in a different random order for each melody (and for each participant). Moreover, the possibility that any such benefits might lead to a ceiling effect was largely offset given the high level of difficulty of the task. Participants completed 2 practise trials prior to the full set, and the experiment lasted about 1 h. Participants reproduced as many of the 108 melodies as they could (at least one of each of 12 variants) within the one-hour limit of experiment duration; this number ranged from 12 to 60 based on the participant.

2.2. Results

For each participant in the perception group, complexity ratings were averaged across melody, resulting in 12 unique ratings per participant corresponding to the 12 melody variant conditions.[3] A 2 (dimension = pitch, time) × 6 (Diversity level = 1–6) repeated measure ANOVA revealed significant main effects of Dimension, $F(1,17)=8.33$, $p=.01$, $\eta^2=.07$[4], Diversity, $F(5,85)=115.10$, $p<.001$, $\eta^2=.60$, and a Dimension × Diversity interaction, $F(5,85)=5.72$, $p<.001$, $\eta^2=.02$. The interaction indicated that perceived complexity of the pitch and time variants was equal at low diversity variants, and that pitch variants were rated as more complex than time variants at higher levels of diversity. Fig. 2 shows this interaction.

In the production task, error rates were calculated using a dynamic matching algorithm (Large & Rankin, 2007) implemented in MATLAB. For the isochronous variants, this algorithm matched performances to MIDI notation of the heard melody; standardised pitch error rate was calculated by counting the number of incorrect notes divided by the number of notes in the notation. For the monotonic variants, IOIs were quantised to multiples of semiquaver durations, and then coded as pitch values (e.g., semiquaver = 1, quaver = 2, etc.). This quantisation and coding procedure was also done for the MIDI notation. These "pitch" values (quantised IOIs coded as pitches) were matched to the MIDI notation using the same matching algorithm, and then errors were calculated using the same procedure as before.

Each dependent measure in the production task was averaged across melody, resulting in 12 unique data points per participant (as in the perception data). Due to a computer error, perceived difficulty data were lost for one participant. Averaged across participants, error rates were correlated with repetitions, ($r(11)=.83$, $p=.002$), and with difficulty ratings, ($r(11)=.74$, $p=.009$). Furthermore, difficulty ratings were correlated with repetitions ($r=.90$, $p<.001$). These correlations indicate that the effects of diversity were similar across measures in the production group.

As in the complexity ratings, pitch variants with 6 and 7 pitch classes did not differ in error rate, $t(17)=.45$, $p=.66$, perceived difficulty, $t(16)=1.0$, $p=.33$, or repetitions, $t(17)=.01$, $p=1.0$. Therefore subsequent analyses omitted pitch variants with 7 unique pitch classes, allowing 2 × 6 repeated measures ANOVAs on each dependent measure. As before, Dimension and Diversity were the within-subjects factors for all analyses.

Results for Error rates are shown in Fig. 3. The ANOVA yielded significant main effects of Dimension, $F(1,17)=16.55$, $p=.001$, $\eta^2=.11$, Diversity, $F(5,85)=29.17$, $p<.001$, $\eta^2=.41$, and a significant Dimension × Diversity interaction, $F(5,85)=3.43$, $p<.01$, $\eta^2=.02$. Participants made more temporal errors than pitch errors overall, and errors were increasingly frequent in more diverse sequences. The interaction shows that the difference between pitch and temporal error rate increased with dimensional diversity. To test for a speed–accuracy tradeoff, one-way ANOVAs analysed the effect of stimulus tempo (averaged across instance within a participant) on

[3] As expected, variants with more diversity received higher ratings of complexity (see Fig. 2). However, there was no significant difference between the variants with pitch diversity of 6 and 7 pitch classes, $t(17)=.28$, $p=.78$, the only non-significant difference in all comparisons of pitch diversity levels. Thus including level 7 in the statistical analysis provides no additional information to the experimental question at hand. Accordingly, we omitted level 7, leaving 6 levels of diversity (unique pitches/IOIs) in both dimensions. This adjustment has the convenient feature of allowing an evenly matched 2 × 6 ANOVA design.

[4] Please note that the effect sizes reported throughout this paper are eta-squared values, not partial eta-squared (Cohen, 1973).

Fig. 1. Example seed melody and derived pitch and temporal variants for Experiment 1.

pitch error rates and temporal error rates, with no significant results, $F(2,25) < 1$ in both cases. Thus error rates were not linked to tempo, ruling out a speed–accuracy tradeoff.

Figs. 4 and 5 show results from difficulty ratings (scaled between 0 and 100) and number of repetitions. In each case, the ANOVA only yielded a significant main effect of Diversity, $F(5,80) = 35.65$,

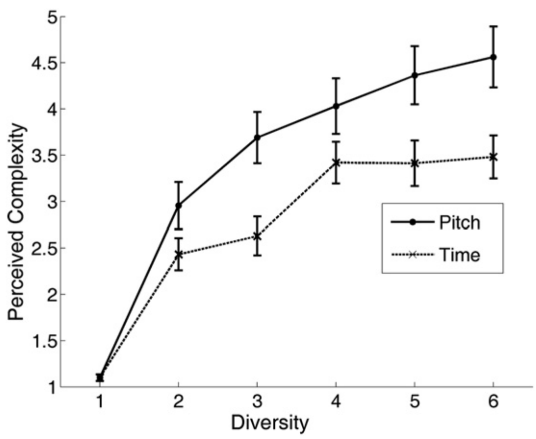

Fig. 2. Complexity ratings for pitch and temporal melody variants in Experiment 1. Error bars represent standard error of the mean.

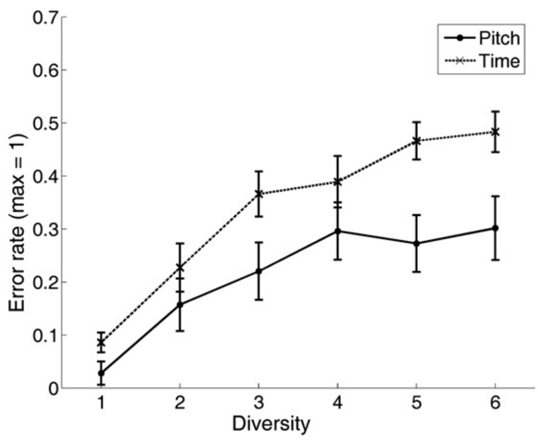

Fig. 3. Pitch and temporal error rates across diversity levels in Experiment 1. Error bars represent standard error of the mean.

$p<.001$, $\eta^2=.44$ for difficulty ratings, and $F(5,85)=21.70$, $p<.001$, $\eta^2=.36$ for repetitions. Both measures were higher for more diverse sequences. Interestingly, neither difficulty ratings nor repetitions differed across Dimension, $F(1,16)<1$, $p=.56$, $F(1,17)<1$, $p=.55$, respectively, and neither exhibited an interaction between Dimension and Diversity, $F(5,80)<1$, $p=.74$, $F(5,85)=1.04$, $p=.4$, respectively.

2.3. Discussion

In Experiment 1, a sample of pianists heard melodies varying either in pitch or temporal diversity, performed them, and rated their difficulty; a sample of musicians with no restriction on instrument rated the melodies' complexity. Results demonstrated that listeners are sensitive to variations in dimensional diversity when making perceptual ratings, and when reproducing melodies from memory. Importantly, primary measures of perception (complexity ratings) and production (errors) suggested differing effects of diversity within the dimensions of pitch and time. Whereas diversity always increased errors as well as perceived complexity, the magnitude of the effect differed across dimensions. Complexity ratings increased with increasing diversity at a faster rate for the dimension of pitch than for time. As a result, pitch variants were perceived as significantly more complex than temporal variants at high, but not low, levels of diversity. Within the domain of production, the effects of diversity differed, with diversity increasing errors at a faster rate for time than for pitch. Consequently, temporal error rates moved further above pitch error rates with increasing diversity.

The fact that the interaction between dimension and diversity went in opposite directions for perception and production measures reflects the predictions of the dimensional salience hypothesis as described in the Introduction. Specifically, both effects could be accounted for by the notion that, in these sequences largely typical of Western music, participants prioritise pitch information at the expense of time. This prioritisation leads to less effective encoding of temporal categories relative to pitch. This fundamental difference may not be apparent for low levels of diversity, where the number of categories to be processed is unchallenging (e.g., 1–2 categories) but appears for higher (more difficult) levels of diversity. As a result, highly diverse temporal patterns are hard to recall and yet do not receive appropriately high levels of complexity ratings due to an insufficient encoding of fine timing information in the mental representation of the sequence. The difficulty of temporal categories may result from relative unfamiliarity with melodies that include high levels of temporal diversity (Fraisse, 1982). In this context, the tendency for functions relating diversity to complexity to level off is informative (see Fig. 2). If participants are less familiar with high levels of diversity in time than pitch, then they may have more difficulty processing temporal diversity. In turn, they would be less likely to notice the difference between temporal diversity levels beyond 2–3 IOIs.

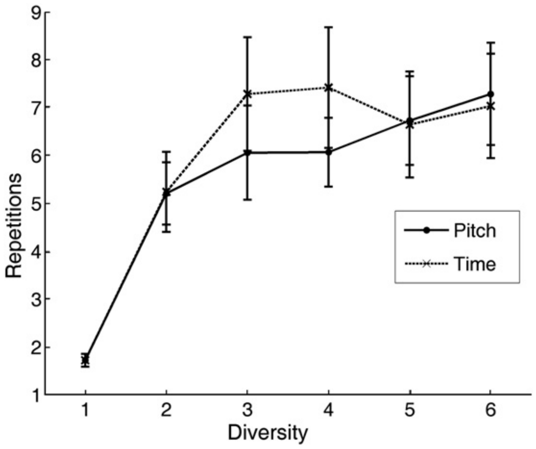

Fig. 4. Repetitions requested prior to performance for pitch and temporal melody variants in Experiment 1. Error bars represent standard error of the mean.

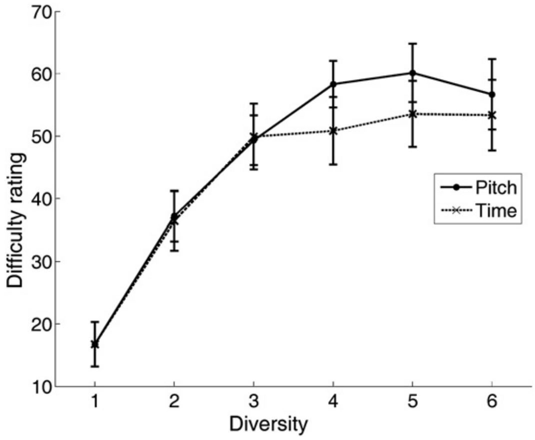

Fig. 5. Performance difficulty ratings for pitch and temporal melody variants in Experiment 1. Error bars represent standard error of the mean.

An important implication of these findings has to do with sources of melodic complexity. Dimensional diversity, as manipulated in this experiment, can be considered as a stimulus driven, or “bottom-up” source of complexity. The present data show, not surprisingly (though unique to this study), that this simple objective measure does in fact influence perceived complexity, error rates, perceived difficulty, and repetitions requested prior to performance. At the same time, the present data clearly suggest that an entirely stimulus-driven approach is incomplete. Although in the present experiment (and in contrast to many previous studies) diversity was controlled across the dimensions of pitch and time, equal diversity levels did not yield equivalent effects. This is true both for the effects of individual levels of diversity and (more importantly) for the effect of varying diversity within each dimension. Of course, one could argue that diversity levels are not directly comparable across dimensions (e.g., pitches scale in cents, which may not be comparable to the scaling of IOIs in milliseconds). However, that is essentially our point. In other words, we have shown here that categories of pitch, within the system of Western tonal music, do not scale equivalently with categories of time.

For both pitch and time, performance error rates in the current study were high in comparison with the existing literature in music performance. For instance, performance of technically difficult finger exercises found error rates between 10 and 15% for a comparable range of tempi and participant expertise (Pfordresher, Palmer, & Jungers, 2007). The higher error rates found here (M pitch = 21%; M time = 34%) likely reflect the fact that the sequences were long, intricate, and were performed from memory by ear (i.e., no notation was provided) without the benefit of practise.

These findings also raise additional questions. First, how will combining pitch and temporal diversity in sequences affect their perception and production? The effects of diversity may be additive, suggesting independent contributions of pitch and time; conversely there may be a more complex pattern in the data when pitch diversity and temporal diversity covary, suggesting interactive relations. Second, will dimensional salience effects in Experiment 2 mirror those from Experiment 1? Experiment 2 was designed to address these issues.

Making inferences about dimensional salience requires that any differences between dimensions are not due to inherent mismatches in their level of difficulty. Work on dimensional integration shows that the relative difficulty (e.g., discriminability) of two dimensions can influence their relations, such that a more discriminable dimension will show asymmetric influence on a dimension with lesser discriminability (Garner, 1974; Melara & Mounts, 1993). Fortunately, this experiment provides measures of relative difficulty between dimensions (similar to a baseline measure of discriminability) that can be used to select levels of pitch and temporal diversity that yield comparable levels of difficulty, based on error rate. Thus the goal of Experiment 2 was to investigate the effects of diversity on how pitch and time combine in the perception and performance of musical sequences, using diversity levels that demonstrated the smallest differences in error rate across the dimensions of pitch and time.

3. Experiment 2

Experiment 2 was designed to test how varying dimensional diversity in both pitch and time affects the perception and production of musical sequences. Thus this experiment had different stimuli than Experiment 1, but the same tasks. Based on the error rates of Experiment 1, a subset of the levels of pitch and temporal diversity were selected, that is, those that exhibited the smallest possible difference across dimension. When recombined, these levels constituted new variants that varied both in pitch and time. By minimising differences in difficulty (as indexed by error rates), residual interference between dimensions could provide evidence of other factors influencing how the dimensions combined, such as dimensional salience.

3.1. Method

3.1.1. Participants

As in Experiment 1, two groups of participants took part in this experiment, namely a production group of 16 pianists, and a perception group of 16 musicians with no primary instrument restriction. The pianists were on average 19.9 years old ($SD = 2.1$) and had on average 10.4 years of training ($SD = 2.1$). The perception group had 9.9 years of training ($SD = 1.9$) and was 18.8 years old ($SD = 1.8$). The two groups did not differ in years of training, $t(15) = .71$, $p = .48$.

3.1.2. Stimuli

Based on Experiment 1 data, four levels of diversity in both pitch and time that were most closely matched in error rate were selected for use in Experiment 2. Specifically, pitch diversity levels 3–6 and temporal diversity levels 2–5 gave near-equivalent ranges of error rates (see Fig. 3), so these levels were used to form the stimuli in this Experiment. Each level of pitch diversity was combined with each level of temporal diversity, yielding 16 variations for each of the 9 seed melodies (144 unique stimuli in total). Fig. 6 depicts the 16 variants of an example melody (derived from the same seed melody as the examples in Fig. 1).

3.1.3. Apparatus

All aspects of the experimental apparatus were the same as in Experiment 1, for both the perception and production groups.

3.1.4. Procedure

All aspects of the procedure other than the number of trials were the same as in Experiment 1, for both participant groups. The number of trials was equal to the number of stimuli (144) for the perception group, and the production group completed as many of the 144 stimuli as possible within the 1-hour session, ranging from 16 to 32.

3.2. Results

As in Experiment 1, all complexity ratings for the perception group were averaged across melody, yielding 16 data points per participant (one for each of the 16 variant conditions). The effects of dimension and diversity on complexity ratings were tested with a 4 (pitch diversity: 3 to 6)×4 (temporal diversity: 2 to 5) repeated measures ANOVA. There was a main effect of pitch diversity, $F(3,45) = 23.48$, $p < .001$, $\eta^2 = .17$; all levels of pitch diversity except 5 and 6 were significantly different. There was also a main effect of temporal diversity, $F(3,45) = 27.48$, $p < .001$, $\eta^2 = .26$, again showing an increase in perceived complexity with diversity, however diversity levels 2 and 3 did not differ, nor did diversity levels 4 and 5. There was no interaction between pitch and temporal diversity, $F(9,135) = 1.66$, $p = .11$, $\eta^2 = .03$. Fig. 7 depicts these data.

The dynamic matching algorithm used in Experiment 1 to evaluate pitch and temporal errors was also used in Experiment 2. Overall, pitch error rate was lower than temporal error rate, as indicated by a one-way repeated measures ANOVA with dimension as a factor and error rate as the dependent variable, $F(1,15) = 5.65$, $p = .03$, $\eta^2 = .06$. For pitch errors, the 4×4 ANOVA (as used in the complexity ratings above) yielded no main effects of pitch diversity, $F(3,45) < 1$, ns. However, temporal diversity affected pitch error rate, $F(3,45) = 4.14$, $p = .01$, $\eta^2 = .04$, because sequences with 3 IOIs had fewer pitch errors than sequences with 5 IOIs, 95% CI [.02, .10], $p = .007$. No other pairwise comparisons were significantly different (e.g., pitch errors for sequences with 2 IOIs did not differ from those with 5 IOIs). There was no interaction between pitch and temporal diversity.

The effects of pitch and temporal diversity were similar for temporal errors. Specifically, pitch diversity did not affect temporal errors, $F(3,45) < 1$, ns, yet temporal diversity did, $F(3,45) = 4.20$, $p = .011$, $\eta^2 = .09$; pairwise comparisons revealed that sequences with 4 IOIs had more errors than those with 3 IOIs, 95% CI [.02, .15], $p = .005$. Overall, errors were most common in 4 IOI sequences, but were significantly different only from 3 IOI sequences. No other pairwise comparisons showed significant differences; there was no interaction between pitch and temporal diversity. Table 2 shows the pitch and temporal error rate data.

When applied to difficulty ratings, the 4×4 ANOVA revealed no significant main effects of pitch or temporal diversity, $F(3,45) = 1.37$, $p = .26$, $\eta^2 = .01$ and $F(3,45) = 1.09$, $p = .36$, $\eta^2 = .01$, respectively; there was no pitch–time interaction, $F(9,135) < 1$, ns. The same pattern emerged in repetitions, $F(3,45) < 1$, ns for both main effects, and no interaction, $F(9,135) = 1.01$, $p = .43$, $\eta^2 = .04$.

3.3. Comparisons across Experiments 1 and 2

Further analyses compared the results of Experiments 1 and 2, to see how introducing variability along one dimension modulated the effect of diversity of a second dimension. This comparison required some re-processing of the data from Experiment 2, specifically matching each level of diversity in one dimension (e.g., pitch) across experiment, while averaging across diversity in the other dimension (e.g., time). For example, to see how temporal diversity affected perceived complexity of pitch diversity, the Experiment 2 data for each level of pitch diversity (e.g., 3 unique pitches) consisted of an average across all levels of temporal diversity (e.g., [pitch 3, time 2], [pitch 3, time 3], [pitch 3, time 4], and [pitch 3, time 5]). These values were then compared with the Experiment 1 perceived complexity data from the condition whose variants had 3 unique pitches and no temporal diversity (isochronous). The opposite averaging procedure (i.e., averaging across levels of pitch diversity for each level of temporal diversity) occurred for measuring how pitch diversity influenced the effect of temporal diversity on perceived complexity.

A 2 (dimension exhibiting diversity)×4 (diversity level)×2 (experiment) mixed ANOVA analysing complexity ratings revealed a main effect of the dimension exhibiting diversity, $F(1,32) = 13.28$, $p = .001$, $\eta^2 = .15$, because sequences with pitch diversity (averaged across levels of temporal diversity) received higher ratings of perceived complexity than sequences with temporal diversity (averaged across levels of pitch diversity). There was also a main effect of the level of diversity, $F(1,32) = 98.72$, $p < .001$, $\eta^2 = .19$, as greater diversity yielded higher ratings. Surprisingly, there was no overall difference between experiment in perceived complexity, $F(1,32) = 1.14$, $p = .29$, $\eta^2 = .00$. These main effect results were qualified by two-way interactions. Dimension and diversity interacted, $F(3, 96) = 4.61$, $p = .005$,

Fig. 6. All 16 variants resulting from one seed melody (not shown) used in Experiment 2.

$\eta^2 = .01$, because ratings increased in a continuous fashion across pitch diversity, but not for time (described above). Dimension also interacted with experiment, $F(1,32) = 13.28$, $p = .001$, $\eta^2 = .15$, because introducing pitch diversity to temporal variants greatly increased perceived complexity, but introducing temporal diversity to pitch variants had no effect, $F(1,32) < 1$, ns. Lastly, there was an interaction between diversity and experiment, $F(3, 96) = 2.91$, $p = .04$, $\eta^2 = .01$, reflecting the fact that the range of complexity ratings (averaged across dimension) was more compressed in Experiment 2 (min: 3.51; max: 4.2) than the same diversity range of Experiment 1 (min: 3.06; max: 3.99). This interaction may be due to the fact that all stimuli in Experiment 2 exhibited pitch diversity (which in general leads to higher complexity ratings), whereas Experiment 1 stimuli included monotonic variants that (without pitch diversity) received lower ratings than isochronous stimuli. Fig. 8 depicts these data.

A stepwise regression analysis tested if the pitch and time complexity ratings from Experiment 1 predicted the complexity ratings of Experiment 2, and also if the pitch and time ratings had independent or interactive contributions. There were thus three independent variables, corresponding to the Experiment 1 complexity ratings of the pitch variants and the time variants (step 1), and a multiplicative interaction term (Experiment 1 pitch complexity ratings ∗ time complexity ratings; step 2). The complexity ratings of the 16 variant conditions of Experiment 2 (averaged across participant) served as the dependent variable. The equation predicted the Experiment 2 complexity ratings, $F(3, 12) = 84.29$, $p < .001$, $R^2 = .95$. All three predictors contributed to the

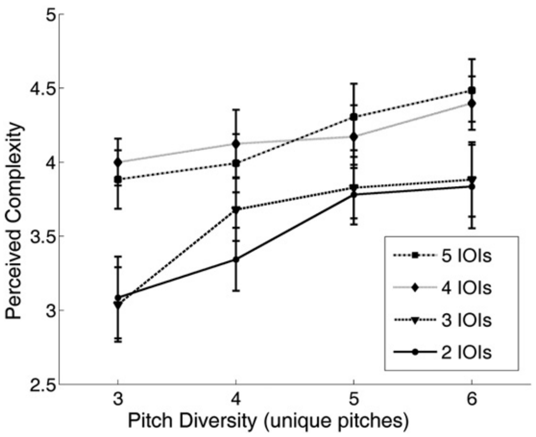

Fig. 7. Perceived complexity as a function of pitch and temporal diversity in Experiment 2. Error bars represent standard error of the mean.

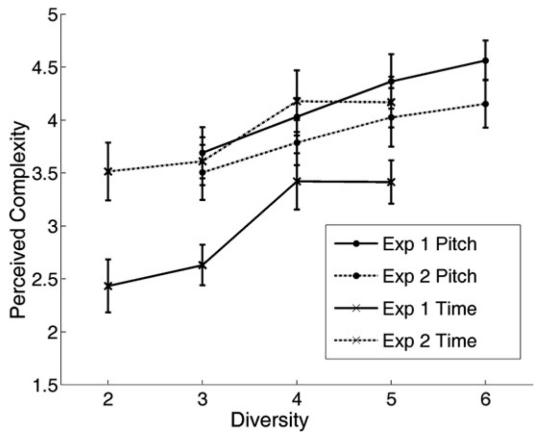

Fig. 8. Comparison across experiment of the effects of diversity on complexity ratings. Experiment 2 pitch ratings are averaged across levels of temporal diversity; Experiment 2 time ratings are averaged across pitch diversity. Error bars represent standard error of the mean.

equation: Experiment 1 pitch complexity, $sr^2 = .06$, $p = .002$, Experiment 1 time complexity, $sr^2 = .04$, $p = .005$, and the multiplicative interaction factor (Experiment 1 pitch complexity∗time complexity), $sr^2 = .02$, $p = .002$. As the sum of these squared semipartial correlations (sr^2, indicating unique variance accounted for) amounts only to 12% of the total 95% explained, there was much overlapping variance (77%) accounted for by these three (intercorrelated) predictors.

The same averaging technique and 2×4×2 ANOVA design described above was used to compare across experiment the dependent measures of the production group. For error rates, the dimension factor in this ANOVA refers to pitch errors or temporal errors. This analysis thus tested the effect of introducing temporal diversity on pitch error rates, and pitch diversity on temporal error rates. A main effect of Experiment occurred because error rates were higher for Experiment 2, $F(1,32) = 10.58$, $p = .003$, $\eta^2 = .02$. There was a significant effect of dimension on error rate, $F(1,32) = 10.62$, $p = .003$, $\eta^2 = .11$, indicating that temporal errors were more common than pitch errors. Dimension did not interact with experiment, $F(1,32) < 1$, ns, however there was an interaction between diversity and experiment, because the effect of diversity was larger in Experiment 1 than in Experiment 2, $F(3, 96) = 6.57$, $p < .001$, $\eta^2 = .04$. In other words, the influence of diversity in either dimension on error rates was less when performers encountered diversity along both dimensions (Experiment 2) compared to along a single dimension (Experiment 1). Fig. 9 compares the error rates across experiment as a function of dimension and diversity.

We also compared effects of diversity on difficulty ratings and number of listening repetitions for participants who performed the melodies. With respect to difficulty ratings, participants rated Experiment 2 sequences as more difficult than Experiment 1, $F(1,31) = 8.08$, $p = .008$, $\eta^2 = .21$. Experiment interacted with diversity, $F(3, 93) = 8.82$, $p < .001$, $\eta^2 = .02$, but not with dimension, $F(1,31) = 1.78$, $p = .19$, $\eta^2 = .01$. As with error rates, the effect of diversity on difficulty ratings was larger in Experiment 1 than Experiment 2. Interestingly, the number of learning repetitions used for practice did not vary across experiments, $F(1,32) < 1$. Experiment again interacted with diversity for listening repetitions, in the same manner as difficulty ratings and error rates, $F(3, 96) = 3.41$, $p = .02$, $\eta^2 = .02$. There was no interaction between Experiment and dimension, $F(1,32) < 1$, ns.

The same stepwise regression analysis used on the complexity ratings was also applied to the dependent variables from the production group (pitch error rate, time error rate, difficulty rating, requested repetitions). However, none of the stepwise regression equations were significant.

3.4. Discussion

Experiment 2 expanded on Experiment 1 by testing the perception and performance of sequences that varied both in pitch and temporal diversity. By picking a subset of the values of pitch and temporal diversity that demonstrated the closest possible match in Experiment 1, Experiment 2 presented the opportunity to test how these dimensions

Table 2
Pitch and temporal error rates across levels of pitch and temporal diversity in Experiment 2.

Temporal diversity	Pitch diversity				
	3 pitches	4 pitches	5 pitches	6 pitches	Mean
Pitch error rate					
2 IOIs	41%	46%	45%	43%	44%
3 IOIs	40%	38%	39%	44%	40%
4 IOIs	45%	47%	46%	49%	47%
5 IOIs	50%	39%	48%	48%	46%
MEAN	44%	43%	45%	46%	
Temporal error rate					
2 IOIs	46%	42%	50%	54%	48%
3 IOIs	45%	47%	50%	53%	49%
4 IOIs	60%	54%	54%	62%	58%
5 IOIs	52%	52%	57%	48%	52%
Mean	51%	49%	53%	54%	

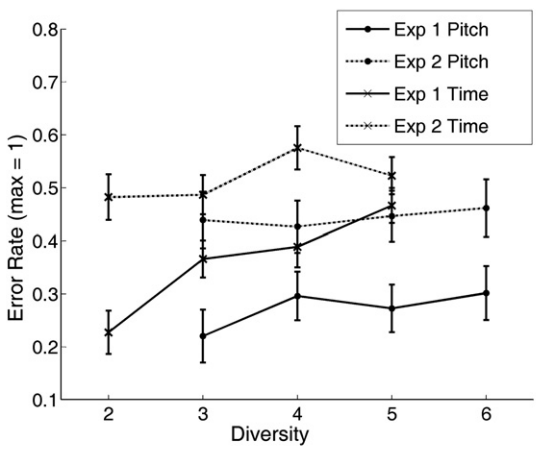

Fig. 9. Comparison across experiment of the effects of diversity on error rates. Experiment 2 pitch errors are averaged across levels of temporal diversity; Experiment 2 time errors are averaged across pitch diversity. Error bars represent standard error of the mean.

combined in perception and performance while minimising differences in difficulty across dimensions. Diversity associated with pitch did not interact with time for the sequences of Experiment 2. Rather, both dimensions had independent effects on complexity ratings, and only temporal diversity affected error rates.

Introducing diversity on two dimensions affected both perception and performance measures (i.e., Experiment 2 relative to Experiment 1). Adding pitch variability to temporal patterns increased their perceived complexity relative to the level of the isochronous pitch variants from Experiment 1. However, the addition of temporal diversity to pitch sequences in Experiment 2 did not similarly affect the level of complexity associated with different levels of pitch diversity. That is, perceived complexity of these sequences appears linked primarily to the presence of pitch diversity and minimally affected by temporal diversity, in keeping with the dimensional salience hypothesis described in the Introduction.

In production measures, introducing diversity on either dimension increased overall error rates in Experiment 2, and to roughly equivalent extents for pitch errors and for timing errors. As in Experiment 1, pitch sequences were performed more accurately than the temporal patterns. However, the data from Experiment 2 did not suggest a greater influence of diversity on timing errors than on pitch errors. Within each error type (pitch or time) there was no interaction across the two dimensions defining diversity, and no interaction was apparent in pooled analyses across Experiments 1 and 2 that compared errors of different types. The disappearance of the interaction between the two error types found in Experiment 1 may reflect the fact that in Experiment 2, every sequence featured diversity on both dimensions. This fact may also explain why introducing diversity on a second dimension decreased the overall influence of diversity (relative to Experiment 1, in which only one dimension exhibited diversity). Specifically, with more total diversity (by virtue of both dimensions varying), the difference between high and low diversity was not as large, yielding an accordingly smaller effect of diversity. The data in Table 2 support this intuition, as the combined (added) diversity across the dimensions of both pitch and time correlated positively with the pitch error rate ($r = .47$) and temporal error rate ($r = .55$). We return to the implications of this result in the General discussion.

4. General discussion

We report the results of two experiments that tested how the diversity of categorical pitch and temporal elements present in an auditory sequence influences the perception and production of that sequence. For Experiment 1 the sequences exhibited diversity in only one dimension, and revealed that adding diversity in either dimension generally increased perceived complexity as well as production errors. In Experiment 2, sequences included diversity across both dimensions. Temporal diversity had larger effects on production errors than diversity of pitch in Experiment 1. This pitch/time difference in the effect of diversity on errors was exaggerated in Experiment 2, where diversity levels (a subset of those from Experiment 1) had no effect at all on errors. Furthermore, an apparent perception/action dissociation was found. The aforementioned effect of diversity on errors reversed for measures of complexity ratings, with pitch diversity having a larger effect than temporal diversity. Finally, combining diversity in both dimensions increased error rates and diminished the influence of diversity overall (relative to Experiment 1), but there were no statistical interactions between pitch and time.

4.1. Pitch–time combination

These experiments were set up to address the way in which pitch and time combine in the production and perception of musical sequences. As mentioned in Section 1.1, a great deal of debate in the music cognition community has concerned whether pitch and time combine independently or interactively. The data from Experiment 2 support the notion that framing the debate in this way may constitute an oversimplification (cf. Ellis & Jones, 2009; Prince, 2011; Prince, Thompson, et al., 2009). No interactions emerged between pitch and temporal diversity for any dependent measure, in apparent support of independence. At the same time, Experiment 2 demonstrated that production errors in one dimension (pitch or time), are influenced by the auditory dimension that is technically not relevant for accuracy. Thus, the accuracy of timing in production was influenced by the presence of variability in the pitch dimension, though not sensitive to the degree of variability (i.e., there was no main effect of pitch diversity on temporal errors).

Overall, these data suggest that the processing of pitch and time does not occur in a separate and informationally encapsulated fashion (cf. Peretz & Coltheart, 2003). Admittedly, when diversity manipulations in Experiment 2 were significant, the dimensions had independent effects. Yet overall diversity (summing pitch and temporal diversity) correlated with both pitch and time error rates, despite the negligible effects of diversity in each dimension separately. In a similar way, although complexity ratings yielded no pitch–time interactions in Experiment 2, including diversity along the dimensions of both pitch and time elevated complexity ratings compared to patterns with only temporal diversity from Experiment 1. Additionally, an interaction term (multiplying Experiment 1 pitch complexity ratings by Experiment 1 time complexity ratings) predicted Experiment 2 complexity ratings beyond the contributions of these variables separately. Lastly, diversity in both dimensions did not similarly elevate complexity ratings compared with patterns including only pitch diversity from Experiment 1. These results again suggest that pitch and temporal processing are not entirely separate, and is also supportive of another key hypothesis of the current research: pitch salience.

4.2. Dimensional diversity and pitch salience

As discussed in the Introduction, several studies suggest that pitch dominates time in many perceptual tasks, but it has not been clear whether this apparent difference in salience is a by-product of diversity, which is greater for pitch than for time in typical Western music (Fraisse, 1982; Järvinen, 1995; Krumhansl, 1990). Data from Experiment 1 clearly show that pitch salience is not simply a by-product of diversity in the immediate sequence. Participants were more sensitive to pitch diversity than temporal diversity when rating perceived complexity, but were at the same time better able to reproduce pitch patterns as diversity increased than they were able to reproduce temporal patterns. Experiment 2 revealed that pitch primarily drove complexity ratings, as introducing temporal diversity to pitch variants did not increase the ratings. This finding suggests that pitch was the more salient dimension in perceived complexity.

Why would more diversity make pitch more salient? Certainly having a greater diversity of elements increases the processing load of any dimension, but greater diversity also leads to greater informativeness. Repeated exposure to stimuli with greater diversity along one dimension may improve the processing ability for that dimension. In turn, although more diversity is always more difficult, the effect will be less noticeable for habitually diverse dimensions because listeners have developed the necessary skills to handle it successfully. Recall that in typical Western music, pitch is more variable and more elaborately structured than time—using (on average) 7 unique pitch classes but 2 unique IOIs. Internalising these statistical properties would then lead to a tendency to prioritise pitch in the context of such music. Relative to these norms, in Experiment 2 the number of pitch classes was lower (3–6), whereas the number of unique IOIs was higher (2–5). Attenuating differences between pitch and time involved changing the relative variability, or informative value, of the dimensions in the immediate sequence, from the typical norms of

Western music. Yet these changes had no noticeable effect on how pitch and time combined. Accordingly, although changing dimensional diversity in the immediate sequence may not affect dimensional salience, it is possible that because these sequences at least loosely resembled Western music, they invoked a learned, or schematic, prioritisation of pitch.

The lack of statistical interactions between pitch and temporal diversity reveals that the particular level of diversity in a dimension did not influence the effect of diversity in another dimension. Additionally, effects of pitch and temporal diversity in Experiment 2 yielded much smaller effects on errors when both forms of diversity were present in sequences, in contrast to Experiment 1. It is interesting that pitch diversity in Experiment 2 yielded no significant effect on either pitch errors or on time errors, whereas temporal diversity did. Yet the effect of temporal diversity was smaller in Experiment 2 than Experiment 1, and more importantly did not uphold the same relationship between diversity and error rates as was found in Experiment 1. Why did the effect of diversity change in Experiment 2, in contrast with previous research on (albeit perceptual) auditory dimensional integration (Melara & Mounts, 1994)? One possibility is that the Experiment 2 error data may represent a limit in how much total diversity (summed across pitch and time) participants could accurately produce. With finite perceptual processing resources (and memory capacity), participants may have engaged in a tradeoff between dimensions, such that accuracy would decrease in one dimension in order to achieve some accuracy in the other dimension. Similar results have occurred in judgements of melodic similarity (Monahan & Carterette, 1985); participants tended to base similarity ratings on one dimension (pitch or time) at the expense of the other. Some authors have suggested that this ability to emphasise selectively one dimension over another is suggestive of independent processing (Palmer & Krumhansl, 1987a, 1987b; Prince, 2011). In any case these data suggest that pitch has a special role in the perception and production of auditory sequences, relative to time.

Of course, it is not certain that pitch salience would be found in all situations. Indeed, research summarised earlier suggests a dominance of temporal over pitch factors in synchronisation tasks (e.g., Snyder & Krumhansl, 2001). A possible critical factor in the current study is that participants were required to make perceptual judgements that were highly abstract, or to engage in production tasks that involved retrieval of sequential information. Perhaps temporal complexity is more in higher-order factors that involve the extraction of global time structure, critical for synchronisation tasks. Additionally, other domains of auditory cognition may reveal different patterns of salience. As we have pointed out, musical sequences often possess greater pitch complexity than temporal complexity. However, in other domains this may not be the case, leading to greater balance of pitch and time, or even temporal dominance.

4.3. Perception and production

A final issue that was critical to the experiments reported here was to compare the effects of diversity across perception and production tasks. We found several differences across perception and production here. First, in Experiment 1, we found reversed effects of dimension for perception and production tasks, as mentioned earlier (Section 2.3). Second, despite the fact that there was no difference between dimensions in Experiment 2 complexity ratings (Fig. 8: Experiment 2 Pitch and Experiment 2 Time), there were still fewer pitch errors than temporal errors. Participants may have prioritised the dimension of pitch as it was more salient, despite the time dimension being equally as complex. Lastly, the level of diversity affected complexity ratings in Experiment 2, but not production measures.

Overall, therefore, Experiment 2 reinforces the disparity between perceptual and performance measures found in Experiment 1. This mismatch adds to the growing literature on perception/action mismatches in music cognition (Loui et al., 2008; Pfordresher & Brown, 2007; Repp et al., 2011; Zatorre et al., 2007). It is possible that this disparity results from a between-subjects comparison (perception group vs. production group), but two factors argue against this interpretation. First, this pattern emerged in both experiments—two independent tests of unique participants whose data nonetheless suggest a perception/production mismatch. Second, comparing error rates to difficulty ratings and repetitions provides a within-subjects perception/production comparison. Although the difficulty ratings and repetitions do not differ across dimension, they nonetheless diverge from the error rates that show more errors in time than in pitch.

4.4. Limitations

A limitation of these results stems from the assignment of pitch classes and durations to diversity levels. Specifically, this assignment was fixed, that is, not randomised within participant nor counterbalanced across participants. For example, the pitch classes used in variants with two unique categories (i.e., diversity level of 2) were always the tonic and dominant scale degrees; the duration denominations for the homologous temporal variants were always quavers and crotchets. In a more complete design, the assignment of pitch and temporal categories to diversity levels could vary, however there would be drawbacks. In particular, such a design would need many more trials, and would likely have an unclear tonality and/or metric framework. Accordingly, the present design is a compromise between the interests of ecological validity in preserving the musical nature of the sequences, and controlled scientific conditions. This is a common issue in the area of music cognition, and deserves careful treatment.

A potential concern with our error measures is that equating the number of pitch errors to temporal errors may not be a valid assumption. Accordingly, the observed differences between the rate of pitch and temporal errors may be an artefact of our particular error measures. For instance, in Experiment 1 the 1-IOI and 1-pitch (no temporal nor pitch diversity) is essentially an isochronous tapping task, at which the trained participants should perform without trouble. Instead, the average error rate was 7%. Consider, however, that varying the duration of a semiquaver (scaled to each performance) for a single mistake in the entire 16-note sequence would result in an error rate of 6.25% (1 note out of 16). Further inspection of the data revealed that all except one participant produced this error rate in these isochronous sequences. These errors therefore represent more an issue of timing precision than categorical production errors. This potential artefact is one reason why we assessed performance using effect sizes associated with manipulations of diversity.

Another possible limitation of the current study had to do with the fact that performers were made to learn sequences “by ear”, without being exposed to (visual) music notation, which was unquestionably challenging and incurred high demands on memory. This experimental design was intended to remove the contribution of stored motor programs (i.e., practise) to keyboard performance, enabling a purer investigation of the contributions of pitch and time to performance. In exchange, this approach has the limitation of taxing participants' memory resources heavily. Future research in this area is necessary to disentangle comprehensively the contributions of memory and motor practise from performance.

It is worth noting that our measures of production focused on the accuracy with which participants sequence pitches and durations during recall, whereas other research on the role of pitch and time has focused on expressive timing (Drake & Palmer, 1993) and sensorimotor synchronisation (e.g., Jones & Pfordresher, 1997; Pfordresher, 2003; Snyder & Krumhansl, 2001). Our choice of measures was designed to provide a strong test of the dimensional salience hypothesis in production, and are similar to other measures used in studies of recall (e.g., Drake & Palmer, 2000). Nevertheless, it is possible that the independent effects of pitch and time found in our data would differ in a task that focused on temporal nuances of production such as synchronisation tapping. For instance, other research suggesting interactive effects of pitch and time has focused on the way in which accents created by serial changes along these dimensions contribute to the formation of higher-order temporal structure (Ellis & Jones, 2009; Jones & Pfordresher, 1997; Pfordresher, 2003). Whereas the current data suggest independent contributions of pitch and time during recall, they do not speak to whether melodic and rhythmic accents in sequences create independent time structures, or whether accents interact as per the joint accent structure construct of Jones (1987).

4.5. Generalising dimensional salience

Developed within the context of pitch–time combination, research exploring the dimensional salience hypothesis remains limited to the domain of music cognition. Nevertheless, this concept may prove fruitful as a theoretical framework of perceptual processing with further investigation and expansion to additional domains. For instance, interactions across the dimensions of pitch and time in auditory patterns influence auditory organisation (Bigand, Madurell, Tillmann, & Pineau, 1999; Brochard, Drake, Botte, & McAdams, 1999; Griffiths & Warren, 2004; e.g., Mondor & Terrio, 1998; van Noorden, 1975). However, to date important questions remain regarding exactly how these dimensions combine (cf. Justus & List, 2005; Silbert, Townsend, & Lentz, 2009; Winkler, Denham, Mill, Bohm, & Bendixen, 2012). Thus examining the role of dimensional salience in pitch–time combination for auditory contexts beyond music is a promising area of future research.

Dimensional salience may also apply to cognition more generally, as it complements existing work on object perception. Indeed, the dimensional salience hypothesis has strong roots in an information processing approach to perception. Garner's seminal work (1974) proposed that the physical dimensions of a stimulus may be processed as separable (independent) or integral (interactive). Separable dimensions (e.g., shape and colour) can be processed independently and experience no mutual interference, whereas integral dimensions are by nature processed as an integrated whole (e.g., saturation and brightness). He also showed that separable dimensions can falsely appear to be integral if one is more discriminable (easier to process) than the other. A more discriminable dimension will interfere with the other dimension, and be immune to the less discriminable dimension. Yet there are multiple domains in which asymmetric interference occurs despite equal discriminability, such as the perception of faces (Atkinson, Tipples, Burt, & Young, 2005), speech (Tong, Francis, & Gandour, 2008), and music (Prince, Thompson, et al., 2009). Proposed explanations of such phenomena based on physical primacy (Wood, 1974), or invariant versus changeable attributes (Haxby, Hoffman, & Gobbini, 2000) may be subsumed within a dimensional salience hypothesis.

Another influential body of work that relates to dimensional salience is Massaro's Fuzzy Logic Model of Perception, or FLMP (Massaro & Friedman, 1990; Oden & Massaro, 1978). The FLMP has multiple serial processing stages (evaluation, integration, assessment, response selection) to accomplish the task of perceiving multidimensional stimuli. In other words, the perceiver must form a mental representation of the stimulus, in part by evaluating the relative importance of multiple dimensions and weighting them accordingly. This arrangement presumes independence of dimensions at the stage of feature evaluation, similar to work in music cognition that proposes a stage model of pitch–time combination (Peretz & Coltheart, 2003; Thompson, et al., 2001). FLMP includes matching the stimulus information to a stored prototype, also similar to research in music perception in which more culturally prototypical stimuli confer a processing advantage (Lebrun-Guillaud & Tillmann, 2007; Tillmann & Bharucha, 2002).

Perhaps the most immediately relevant attribute of the FLMP to dimensional salience emerges from the evaluation stage, when each information source (i.e., dimension) is weighted according to its degree of ambiguity (similar to the degree of informative value of dimensional salience). Less ambiguous sources receive greater weight (a higher “fuzzy truth value”), derived from the extent to which the exemplar matches a stored prototype of the stimulus in question. However a notable difference between these models is in the definition of this ambiguity. Schwarzer and Massaro (2001) varied ambiguity by adjusting the relative distinctiveness between eye and mouth features in face identification. This adjustment would unquestionably influence the psychophysical discriminability of a dimension (Garner, 1974). In contrast, dimensional salience is independent of discriminability (Prince, Thompson, et al., 2009), such that the salience of a dimension could be influenced by manipulations that have no effect on the perceptual difficulty (or ambiguity, in FLMP terminology).

A final relation between dimensional salience and the FLMP comes from the creation of a non-categorical, continuous representation of an object through integration of independent information sources (Massaro & Cohen, 1990). Dimensional salience is not a categorical all-or-none dominance of one dimension over another, but a prioritising in accordance with its informative value. Thus the relative dimensional salience influences observed relations between dimensions (Prince, 2011). Although the FLMP was developed in the context of speech perception, it has successful applications within numerous perceptual domains and across modalities (cf. Massaro, 1987; Massaro, 1998). More generally, dimensional salience may similarly contribute to the understanding of the binding problem in object perception (Treisman, 1996), applied primarily to visual perception but generalised to multiple modalities and the domains of both perception and production (Hommel, 2004; Zmigrod, Spape, & Hommel, 2009).

In conclusion, the current results support the view that pitch is more salient than time in the context of perception and performance of typical musical sequences. Increases in the diversity of pitch categories had larger effects on complexity ratings, yet smaller effects on performance errors, than did increases in the diversity of temporal event categories. Moreover, the current experiments showed that pitch salience is not simply an artefact of dimensional diversity, which in most studies is confounded with dimension. When sequence events vary in both pitch and time, independent effects emerged; however, the influence of diversity in each dimension was reduced (compressed) for sequences that included variability in both pitch and time. Taken together, these data suggest that pitch and time are in fact not “equal partners”, as proposed before (Hébert & Peretz, 1997). Further, pitch and time may contribute additively rather than interactively (cf. Palmer & Krumhansl, 1987a), but given the lack of information encapsulation, these dimensions do not seem to function entirely separately (cf. Peretz & Coltheart, 2003).

Acknowledgements

This research was sponsored in part by NSF grant BCS-0642474. We are grateful to Anastasiya Kobrina for help with stimulus preparation and data collection.

References

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.

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., 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.

Boltz, M. G. (1989). Rhythm and good endings — Effects of temporal structure on tonality judgments. *Perception & Psychophysics*, *46*, 9–17.

Brochard, R., Drake, C., Botte, M. C., & McAdams, S. (1999). Perceptual organization of complex auditory sequences: Effect of number of simultaneous subsequences and frequency separation. *Journal of Experimental Psychology. Human Perception and Performance*, *25*, 1742–1759.
Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, *33*, 107–112.
Cousineau, M., Demany, L., & Pressnitzer, D. (2009). What makes a melody: The perceptual singularity of pitch sequences. *Journal of the Acoustical Society of America*, *126*, 3179–3187.
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.
Drake, C., Dowling, W. J., & Palmer, C. (1991). Accent structures in the reproduction of simple tunes by children and adult pianists. *Music Perception*, *8*, 315–334.
Drake, C., & Palmer, C. (1993). Accent structures in music performance. *Music Perception*, *10*, 343–378.
Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, *74*, 1–32.
Eerola, T., Jarvinen, 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.
Ellison, J. W., & Massaro, D. W. (1997). Featural evaluation, integration, and judgment of facial affect. *Journal of Experimental Psychology. Human Perception and Performance*, *23*, 213–226.
Finney, S. A. (2001). Ftap: A linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, *33*, 65–72.
Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). (1st ed.). New York: Academic Press.
Garner, W. R. (1974). *The processing of information and structure*, *Vol. 203*, Oxford, England: Lawrence Erlbaum.
Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, *49*, 585–612.
Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, *5*, 887–892.
Hannon, E. E., Soley, G., & Ullal, S. (2012). Familiarity overrides complexity in rhythm perception: A cross-cultural comparison of American and Turkish listeners. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 543–548.
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.
Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences*, *8*, 494–500.
Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *The Behavioral and Brain Sciences*, *24*, 849–937.
Järvinen, T. (1995). Tonal hierarchies in jazz improvisation. *Music Perception*, *12*, 415–437.
Jones, M. R. (1987). Dynamic pattern structure in music — Recent theory and research. *Perception & 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 & Psychophysics*, *32*, 211–218.
Jones, M. R., & Pfordresher, P. Q. (1997). Tracking musical patterns using joint accent structure. *Canadian Journal of Experimental Psychology*, *51*, 271–291.
Jones, M. R., & Ralston, J. T. (1991). Some influences of accent structure on melody recognition. *Memory and Cognition*, *19*, 8–20.
Justus, T., & List, A. (2005). Auditory attention to frequency and time: An analogy to visual local–global stimuli. *Cognition*, *98*, 31–51.
Knopoff, L., & Hutchinson, W. (1983). Entropy as a measure of style: The influence of sample length. *Journal of Music Theory*, *27*, 75–97.
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). Key-finding in music: An algorithm based on pattern matching to tonal hierarchies. *19th Annual Meeting of the Society of Mathematical Psychology. Cambridge, MA.*
Kubovy, M. (1981). Concurrent pitch-segregation and the theory of indispensable attributes. In M. Kubovy, & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 55–98). Hillsdale, NJ: Erlbaum.
Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*, *80*, 97–126.
Large, E. W., & Rankin, S. K. (2007). Matching performance to notation. In T. Eerola, & P. Toivanen (Eds.), *Midi toolbox: Matlab tools for music research*. Jyväskylä, Finland: University of Jyväskylä Available at , http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox/
Lebrun-Guillaud, G., & Tillmann, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception & Psychophysics*, *69*, 1450–1459.
Lerdahl, F., & Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, Massachusetts: MIT Press.
Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action–perception mismatch in tone-deafness. *Current Biology*, *18*, R331–R332.
Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry.* : Erlbaum Associates.
Massaro, D. W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle.* : MIT Press.
Massaro, D. W., & Cohen, M. M. (1990). Perception of synthesized audible and visible speech. *Psychological Science*, *1*, 55–63.
Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review*, *97*, 225–252.
Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*, 422–471.
Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimensions — Effects of base-line discriminability, response-mode, and practice. *Memory and Cognition*, *21*, 627–645.
Melara, R. D., & Mounts, J. R. W. (1994). Contextual influences on interactive processing: Effects of discriminability, quantity, and uncertainty. *Perception & Psychophysics*, *56*, 73–90.
Monahan, C. B. (1993). Parallels between pitch and time and how they go together. In T. J. Tighe, & W. J. Dowling (Eds.), *Psychology and music: The understanding of melody and rhythm* (pp. 121–154). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc..
Monahan, C. B., & Carterette, E. C. (1985). Pitch and duration as determinants of musical space. *Music Perception*, *3*, 1–32.
Mondor, T. A., & Terrio, N. A. (1998). Mechanisms of perceptual organization and auditory selective attention: The role of pattern structure. *Journal of Experimental Psychology. Human Perception and Performance*, *24*, 1628–1641.
Oden, G. C., & Massaro, D. W. (1978). Integration of featural information in speech perception. *Psychological Review*, *85*, 172–191.
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 & 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.
Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology. Human Perception and Performance*, *25*, 39–58.
Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.
Pfordresher, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception*, *20*, 431–464.
Pfordresher, P. Q., & Brown, S. (2007). Poor-pitch singing in the absence of "Tone deafness". *Music Perception*, *25*, 95–115.
Pfordresher, P. Q., Palmer, C., & Jungers, M. K. (2007). Speed, accuracy, and serial order in sequence production. *Cognitive Science: A Multidisciplinary Journal*, *31*, 1–36.
Prince, J. B. (2011). The integration of stimulus dimensions in the perception of music. *Quarterly Journal of Experimental Psychology*, *64*, 2125–2152.
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.
Repp, B. H. (1998). Variations on a theme by Chopin: Relations between perception and production of timing in music. *Journal of Experimental Psychology. Human Perception and Performance*, *24*, 791–811.
Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.
Repp, B. H., & Knoblich, G. (2007). Action can affect auditory perception. *Psychological Science*, *18*, 6–7.
Repp, B. H., London, J., & Keller, P. E. (2011). Perception–production relationships and phase correction in synchronization with two-interval rhythms. *Psychological Research*, *75*, 227–242.
Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology. Human Perception and Performance*, *27*, 515–537.
Schenker, H. (1935/1979). *Free composition (Der freie satz).* New York: Longman.
Schwarzer, G., & Massaro, D. W. (2001). Modeling face identification processing in children and adults. *Journal of Experimental Child Psychology*, *79*, 139–161.
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.
Snyder, J. S., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, *18*, 455–489.
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., & Bharucha, J. J. (2002). Effect of harmonic relatedness on the detection of temporal asynchronies. *Perception & Psychophysics*, *64*, 640–649.
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.

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.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171–178.

van Noorden, L. (1975). Temporal coherence in the perception of tone sequences. Unpublished doctoral dissertation, Technical University Eindhoven, Eindhoven, The Netherlands.

Warrier, C., & Zatorre, R. (2002). Influence of tonal context and timbral variation on perception of pitch. *Attention, Perception, & Psychophysics*, *64*, 198–207.

Werker, J. F., & Tees, R. C. (2005). Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology*, *46*, 233–251.

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

Winkler, I., Denham, S., Mill, R., Bohm, T. M., & Bendixen, A. (2012). Multistability in auditory stream segregation: A predictive coding view. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *367*, 1001–1012.

Wood, C. C. (1974). Parallel processing of auditory and phonetic information in speech discrimination. *Perception and Psychophysics*, *15*, 501–508.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*, 547–558.

Zmigrod, S., Spape, M., & Hommel, B. (2009). Intermodal event files: Integrating features across vision, audition, taction, and action. *Psychological Research-Psychologische Forschung*, *73*, 674–684.