Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC11). Seattle, Washington, USA.
S.M. Demorest, S.J. Morrison, P.S. Campbell (Eds)

PITCH/TIME DIVERSITY EFFECTS ON PERCEPTION AND PRODUCTION

Jon B. Prince and Peter Q. Pfordresher

University at Buffalo, State University of New York

ABSTRACT

The perception and production of complex musical sequences was tested, while varying either the number of major diatonic pitches or inter-onset intervals (IOIs), but not both. One group of participants rated the complexity of each sequence (perception). A second group reproduced each sequence on a keyboard and rated the difficulty of production (perception and production). For both dimensions, increasing the number of unique elements in the sequences led to greater perceived complexity, rated difficulty, repetitions, and lower production accuracy. Paradoxically, increases in the number of unique IOIs had a smaller influence on perceived complexity than increases in unique pitches, whereas the opposite effect was found in production. Potential explanations include a perception-action mismatch, and an inferior ability to differentiate between IOIs with increasing temporal complexity.

1. INTRODUCTION

Integration of pitch and time is necessary to perceive or perform music, but how this combination occurs remains unclear. In contrast to the swath of research in pitch-time integration in perception, there is scant work on this issue in production, let alone comparisons of the two. This research begins a systematic investigation of whether pitch and time combine similarly in perception and production. This approach is based on the idea that listeners form a mental representation of a musical sequence that necessarily includes both pitch and timing information. However, stimulus factors determine the relative *salience* of the dimensions such that measured behavior suggests independence, interaction, or asymmetric interaction. Dimensional salience refers to the degree to which a given dimension exerts dominance over another, in the absence of differences in discriminability (Prince, Thompson, & Schmuckler, 2009). Therefore, when dimensions are equally salient, they are more likely to obtain interactive effects.

There are several stimulus variables that may affect how pitch and time contribute to music perception and production tasks. However, the experiment reported here tests only the effect of dimensional diversity on dimensional salience. Dimensional diversity refers simply to the number of different elements within a dimension that are present. For instance, a melody including only three pitches (e.g., C – E – F) is less diverse than a melody including all 12 pitch classes. Increased diversity may correspond to improved dimensional salience.

There are also task variables that can influence how pitch and time combine in music. This experiment focuses on the nature of the output response. In the existing literature, tasks labeled as "perceptual" typically employ an output response that is decisional and discrete. By contrast, in "production tasks" the participant is required to reproduce the entire sequence in the correct order, with the correct element identities (pitches), and with response timing that matches the original stimulus. Furthermore, pitch-based tasks such as responding to a note or melody may differ from time-focused tasks such as tapping (Pfordresher, 2003). There are several ways in which such tasks differ; at issue here is whether the output response of these tasks influences how listeners form a mental representation of the musical sequence, with respect to the integration of pitch and time.

In order to assess the role of dimensional diversity on dimensional salience, each dimension must be varied systematically while holding the other dimension constant. This procedure reveals the relative ease of task completion in each dimension separately. Additionally, this variation in diversity may or may not have similar effects on perception and production, therefore both tasks are tested individually.

2. METHOD

2.1 Participants

Two groups of participants were recruited for this experiment, one for the perceptual task and the other for the production task. The production group consisted of 18 musicians (*M* age = 20.8, *SD* = 5.2; *M* years of training = 11.3, *SD* = 3.9) whose primary instrument was piano; there was no restriction on primary instrument for the 18 participants (*M* age = 19.2, *SD* = 1.5; *M* years of training = 10.2, *SD* = 2.1) in the perception group.

2.2 Stimuli

Stimuli were created by composing variations on 9 "seed" melodies that were selected from a set of sightsinging melodies (Ottman, 1986). These seed melodies used all 7 diatonic pitches and remained within a single octave. All of the melodies were in duple meter. Of the twelve 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 (quarter notes). 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.

2.3 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 code, and FTAP (Finney, 2001) was used to record the participants' performances.

Variant	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	𝅘𝅥𝅯 ♪ ♪. ♩ ♩. 𝅗𝅥

Table 1: Pitch classes and IOIs used in melody variants.

2.4 Procedure

Participants in the perception task heard a melody and then rated how "complex" (also described as "complicated, difficult") the melody was, on a scale of 1 to 7. Participants completed 4 practice trials prior to rating the entire set, and the procedure took about 40 minutes.

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). During the listening phase, participants were not allowed to vocalize, tap, finger, or in any way practice performing the melody. Once a participant indicated readiness to perform, they attempted to reproduce the melody 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. The experiment then progressed to the following melody. Participants completed 2 practice trials prior to the full set, and the experiment lasted about 1 hour. Participants reproduced as many of the 108 melodies as they could within the one-hour limit of experiment duration; this number ranged from 12 to 60 based on the participant.

3. RESULTS

For each participant in the perception group, complexity ratings were averaged across melody, resulting in 12 unique ratings corresponding to the 12 variant conditions. A 2 X 6 repeated measures ANOVA using Dimension (pitch, time) and Variant (1-6) as within-subjects variables revealed a main effect of Dimension, $F(1,17) = 16.58$, $p < .001$, $\eta^2 = .05$ and Variant, $F(5,85) = 1856.89$, $p < .001$, $\eta^2 = .86$. There was also an interaction, $F(5,85) = 16.80$, $p < .001$, $\eta^2 = .02$, indicating that the perceived complexity of the pitch and time variants was equal at low variant numbers, and that pitch variants were rated as more difficult that time variants at higher variant numbers. This interaction is shown in Figure 1.

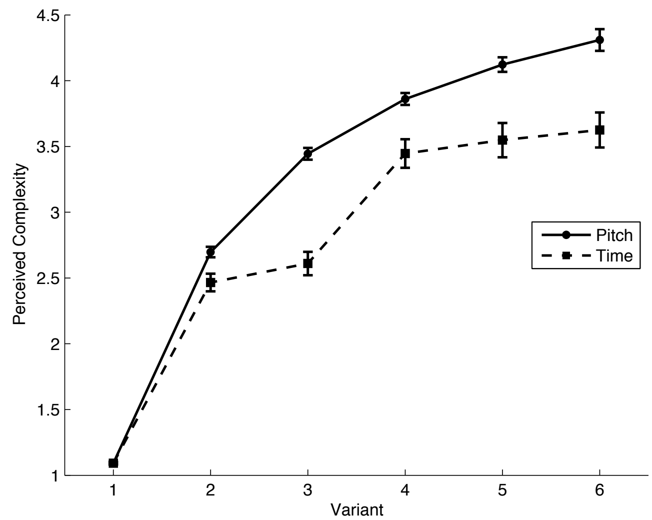

Figure 1: Perceived complexity for pitch and temporal variants as a function of variant (number of unique pitches or IOIs).

In the production task, error rates were calculated using a dynamic matching algorithm (Large & Rankin, 2008) implemented in MATLAB. Each performance was matched to the MIDI notation of the corresponding melody variant, and further programming was used to calculate the error rate of the pitch and temporal variants separately. For the isochronous variants, standardized 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, temporal error rate was calculated by quantizing the matched performances as multiples of 16th notes and comparing them to the similarly quantized notation. The number of incorrect quantized durations in the performance divided by the number of notes constituted the standardized error rate. Error rate and repetitions did not correlate ($r = .03$), but error rate and difficulty did ($r = -.32$, $p < .001$), as did difficulty and repetitions ($r = .47$, $p < .001$).

Error rate, difficulty and repetitions were analyzed in separate 2 X 6 repeated measures ANOVAs. As before, Dimension and Variant were the within-subjects factors for all analyses. Results for Error rates are shown in Figure 2. The ANOVA yielded significant main effect of Dimension and Variant, $F(1,17) = 8.56$, $p < .01$, $\eta^2 = .06$, $F(5,85) = 34.39$, $p < .001$, $\eta^2 = .47$, and a significant Dimension x Variant interaction, $F(5,85) = 2.56$, $p < .05$, $\eta^2 = .02$. The interaction shows that the difference between pitch and temporal error rates increased with dimensional diversity (variant).

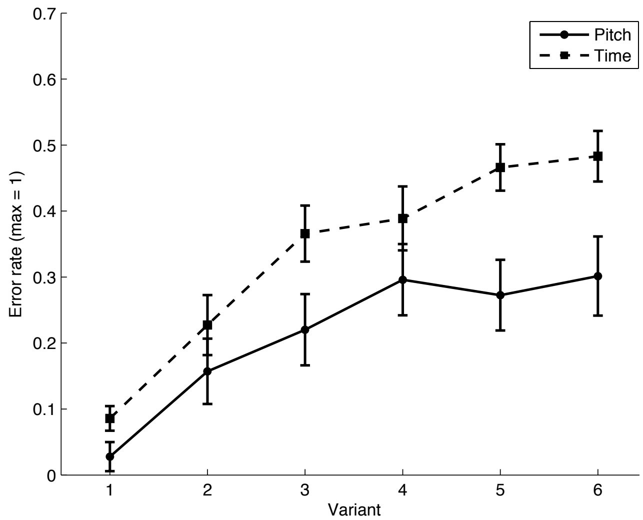

Figure 2: Error rate for pitch and temporal variants as a function of the number of unique elements (pitches or IOIs).

Figures 3 and 4 show results from difficulty ratings and number of repetitions. In each case, the ANOVA only yielded a significant main effect of variant, $F(5,80) = 40.14$, $p < .001$, $\eta^2 = .46$ for difficulty ratings, and $F(5,85) = 21.98$, $p < .001$, $\eta^2 = .36$ for repetitions.

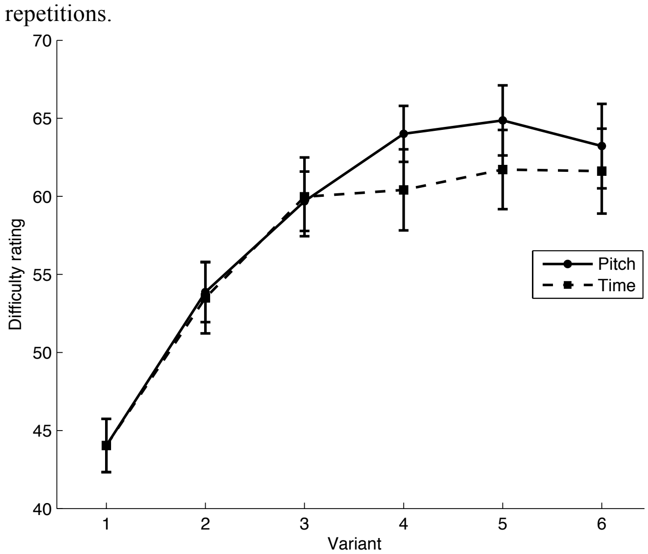

Figure 3: Rated difficulty to perform the pitch and temporal sequences for each variant level (unique pitches or IOIs).

To explore the role of individual differences, difference scores for error rate and difficulty rating were calculated (pitch minus time) for each participant, averaged across variant. There was significant agreement between these difference scores ($r = -.65$, $p <.001$), indicating that the participants who had more errors in time than pitch did not show as much the tendency to rate the pitch variants as more difficult. Nevertheless, pitch and time variants were rated overall as equally difficult, even though they made more errors in time. Figure 5 depicts these data.

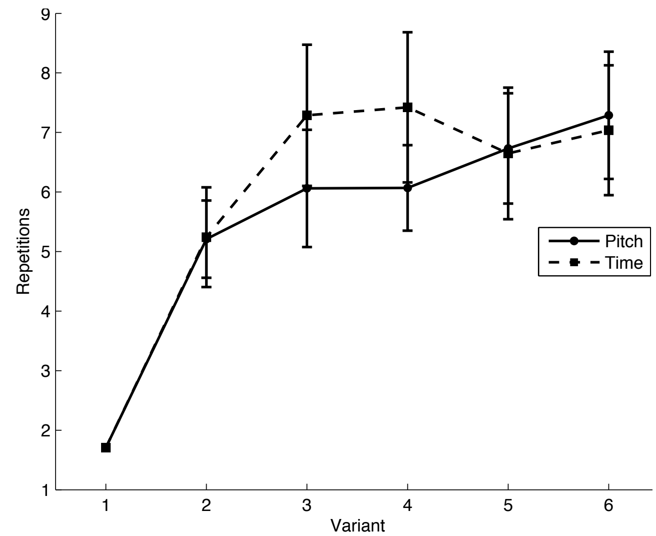

Figure 4: Repetitions heard before performance based on the pitch or temporal variant (number of unique pitches or IOIs).

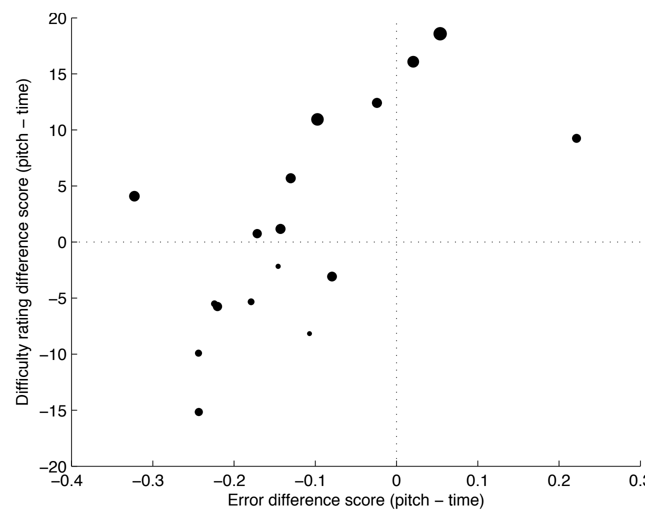

Figure 5: Relation between error difference score (rightwards indicates more pitch errors) and difficulty rating difference score (upwards indicates higher rated difficulty for pitch). Each dot represents a participant. Dot size indicates overall error rate.

4. DISCUSSION

Equalizing the number of unique pitches and IOIs in complex sequences yielded no difference across dimension in the number of repetitions heard prior to performance or ratings of production difficulty. However, pitch variants were perceived as more complex than the temporal variants, despite the fact that errors were lower for pitch variants. The main implication of these data is that equalizing the number of unique pitch and temporal elements in a sequence does not ensure equal discriminability or salience. A second implication is that there appears to be a mismatch between perceptual and performance tasks in the evaluation of pitch-time sequences, given that the error rates conflict with both difficulty and complexity ratings, as well as number of repetitions.

Ensuring that stimulus dimensions are equally discriminable is a critical first step for investigations of dimensional interactions. In the absence of equalized discriminability, dimensional interactions (especially asymmetric interactions) can occur with demonstrably independent dimensions (Garner & Felfoldy, 1970). Nevertheless, dimensional interactions can still occur even with equal discriminability. Usually such interactions are global, that is, both dimensions mutually interfere (Melara & Algom, 2003). However, asymmetric interactions (one dimension interferes with the other but not vice versa), typically thought to indicate unequal discriminability, may instead be due to underlying differences in dimensional salience. That is, equal performance does not ensure equal salience. Therefore, ensuring equal performance in baseline conditions allows stronger inferences about dimensional salience when the two dimensions are recombined.

Therefore, these data fill an important role in establishing levels of equal discriminability for experiments on dimensional integration in performance of sequences that vary simultaneously in pitch and time. In order to ensure equal accuracy in the dimensions of pitch and time, perhaps each level of time should be paired with the next higher level of pitch (e.g., sequences with 2 IOIs and 3 pitches). However, this equalization may not be the same for perception contexts. Indeed, the divergence between the error rate and ratings (as well as repetitions) is intriguing. Why would these measures conflict? A perception-production mismatch is one possible explanation for these data. Even though participants' accuracy difference score covaried with their difficulty ratings (see Figure 5), they still rated pitch as harder while making more temporal errors – there were no cases of the converse (i.e., rating time as harder and making more pitch errors). Despite the growing popularity of a common perception-action mechanism, there is also accumulating evidence of a mismatch between the two (Loui, Guenther, Mathys, & Schlaug, 2008; Repp, 2009). These data contribute to this literature.

Another possible explanation posits that relatively lengthy and complex sequences such as these, the ability to consciously differentiate between IOIs declines as their number increases. Given that typical Western music tends to employ only 2-3 quantized IOIs, listeners may not have developed the cognitive strategies necessary to process optimally larger numbers of unique IOIs in a sequence. In Western music, pitch in tends to be considerably more complex, using a variety of structural features and, most relevant for the present study, about 7 unique pitch classes. After years of exposure to music with such statistical probabilities, more attentional resources may be devoted to processing pitch than time, allowing the differentiation of more unique levels of pitch than IOI. As a result, both performers and listeners would be less able to notice the difference between temporal variants as the number of unique IOIs increased (resulting in more errors), whereas they would have less difficulty noticing the change in numbers of pitch classes used in the pitch variants. Consequently, errors continued to increase with the number of IOIs, while these increases yielded no change in perceived complexity.

These data first step towards a more comprehensive investigation of pitch-time integration in perception and production, by determining how dimensional diversity affects dimensional salience, and what levels of diversity in pitch and time correspond to equal performance. Subsequent experiments can use selective attention instructions (in which both dimensions vary concurrently) to test if pitch and time function independently or interactively in these tasks. Accordingly, this experiment sets the foundation for these selective attention tasks by providing baseline measures of perception and production of musical sequences that vary either in pitch or time. Further research will address some of the questions raised by this initial experiment, and contribute to a fuller understanding of dimensional interactions in complex sequences such as music.

5. REFERENCES

Finney, S. A. (2001). Ftap: A linux-based program for tapping and music experiments. *Behavior Research Methods Instruments & Computers, 33*(1), 65-72.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*(3), 225-241.

Large, E. W., & Rankin, S. K. (2008). Matlab performance matcher.

Loui, P., Guenther, F. H., Mathys, C., & Schlaug, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology, 18*(8), R331-R332.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110*(3), 422-471.

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

Pfordresher, P. Q. (2003). The role of melodic and rhythmic accents in musical structure. *Music Perception, 20*(4), 431-464.

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*(5), 1598-1617.

Repp, B. H. (2009). Segregated in perception, integrated for action: Immunity of rhythmic sensorimotor coordination to auditory stream segregation. *Quarterly Journal of Experimental Psychology, 62*(3), 426-434.