



OPEN Pitch biases sensorimotor synchronization to auditory rhythms

Jesse K. Pazdera¹ & Laurel J. Trainor^{1,2,3}

Current models of rhythm perception propose that humans track musical beats using the phase, period, and amplitude of sound patterns. However, a growing body of evidence suggests that pitch can also influence the perceived timing of auditory signals. In the present study, we conducted two experiments to investigate whether pitch affects the phase and period of sensorimotor synchronization. To do so, we asked participants to synchronize with a repeating tone, whose pitch on each trial was drawn from one of six different octaves (110–3520 Hz). In Experiment 1, we observed U-shaped patterns in both mean asynchrony and continuation tapping rates, with participants tapping latest and slowest when synchronizing to low and extremely high (above 2000 Hz) pitches, and tapping earliest and fastest to moderately high pitches. In Experiment 2, we found that extremely high pitches still produced slower timing than moderately high pitches when participants were exposed to an exclusively high-pitched context. Based on our results, we advocate for the incorporation of pitch into models of rhythm perception and discuss possible origins of these effects.

Keywords Auditory, Continuation tapping, Perceptual bias, Pitch, Sensorimotor synchronization, Tempo

In the field of music perception, several predominant models exist to explain how humans synchronize their movements with rhythmic sounds^{1,2}. *Rhythm* refers to a series of events that are predictably organized across time, and humans—as well as some nonhuman animals^{3–5}—are able to extract a sense of underlying *pulse* or *beat* from these predictable event structures. Some models envision this process of perceiving and synchronizing with the beat as an emergent property of neural dynamics, in which the auditory and motor systems resonate when driven by rhythmic stimulation^{6–9}. Others conceptualize the brain as a Bayesian prediction engine that works to infer the structure underlying predictable patterns in time, in accordance with predictive coding theory^{10–13}. Still others propose biophysical pacemaker models that learn stimulus timing through error correction^{14,15}.

Regardless of category, these established models track rhythms based on the phase, period, and amplitude of sound patterns, but not their pitch. However, there is a wealth of evidence to suggest that pitch affects the perceived tempo of a rhythm^{16–18}, as do the size^{19–22} and direction of pitch changes^{23–26}. In the present study, we sought to identify whether previously identified effects of pitch height on subjective tempo ratings¹⁸ also influence the timing of sensorimotor synchrony. If they do, it would support the incorporation of pitch as a factor in models of rhythm perception, either as an influence on neural dynamics or as an informative cue in a Bayesian prediction process.

The idea that pitch and time perception are integrated dates back decades, originating as an extension of research on the *kappa* effect, in which observing spatial movement distorts time perception^{27,28}. This concept of integrality was formalized by Jones²⁹, who proposed that the brain integrates pitch, time, and loudness into a shared representational space. She hypothesized that music is represented as a trajectory of movement through this multidimensional space, and lawful relations between movements along different dimensions guide musical expectancy. To the extent that pitch and pitch changes predict the tempo or changes in timing of real-world auditory signals, the brain will learn these statistical relations and use pitch as a tempo cue^{16,29}. A similar idea underlies the auditory pitch-motion hypothesis³⁰, in which the movements of an auditory signal in frequency space guide the predicted timing of that signal. Boltz³¹ has further suggested that pitch and time share an integrated representation in memory, and our own lab has recently presented evidence that pitch and timing bidirectionally influence one another in auditory perception³².

In general, higher pitch has been associated with faster and earlier perceived timing^{16,17,24–26,31,32}. For example, Collier and Hubbard¹⁷ observed that participants rated higher-pitched repeating tones and musical scales as

¹Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada. ²McMaster Institute for Music and the Mind, McMaster University, Hamilton, ON, Canada. ³Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada. ✉email: pazderaj@mcmaster.ca

faster and as speeding up more than those played in a lower octave. Similarly, Boltz¹⁶ found that participants rated short melodies as faster when played in a higher octave compared to a lower one, and rated ascending melodies as faster than descending ones. Neural correlates of pitch-induced illusory tempo have since been discovered in both magnetoencephalography (MEG)²⁵ and functional magnetic resonance imaging²⁶. Most relevant to models of rhythm perception is the discovery that ascending and descending pitch patterns shift the phase of entrained neural activity in opposite directions relative to a rhythmic auditory stimulus, while also improving neural tracking of tempo changes in the congruent direction²⁵. These findings offer evidence that pitch biases rhythm perception at the level of neural tracking, supporting the incorporation of pitch into entrainment-based rhythm perception models such as Neural Resonance Theory^{6,8}.

Consistent with Jones's original conceptualization, there is a general consensus that the brain learns this higher-faster association from real-world correlations between pitch and timing²⁹. The most common suggestion is that pitch and timing co-vary with arousal in human vocal production^{16,17,33,34}, although there may also be physical constraints on how quickly low-pitched instruments can be played, due to their larger size and slower attack transients³⁵. However, little work has attempted to demonstrate these proposed pitch-timing correlations^{35,36}, so empirical evidence for their existence remains limited.

Complicating matters, a recent study from our lab found that pitch may influence perceived tempo nonmonotonically, rather than through a simple higher-faster bias. Pazdera and Trainor¹⁸ conducted five experiments examining the generalizability of previous research on subjective tempo^{16,17} to a frequency range wider than previously tested. Whereas Collier and Hubbard¹⁷ compared the perceived timing of the pitches C4 (261.6 Hz), C5 (523.2 Hz), and C6 (1046.4 Hz) and Boltz¹⁶ tested melodies starting on C3 (130.8 Hz) or C5 (523.2 Hz), our previous study tested tempo perception for tones as low as A2 (110 Hz; close to the average male speaking voice) to as high as A7 (3520 Hz; the fourth-highest note on a piano). In this prior study, we asked participants to rate the tempo of repeating tones relative to a metronome. The metronome had the same tempo on every trial, acting as a standard reference, whereas the repeating tone varied in pitch and tempo across trials. In order to disambiguate effects of absolute and relative pitch, we tested some participants' tempo perception across the full 110–3520 Hz range (Pazdera & Trainor, Experiments 1–3¹⁸), while exposing other participants to stimuli from only the lower or upper half of this range (Pazdera & Trainor, Experiment 5¹⁸).

Contrary to the generally accepted view that higher pitches are perceived as faster, we observed an inverted U-shaped relation between pitch and perceived tempo among participants exposed to the full five-octave range of stimuli (Fig. 1A). Between 110 Hz and 440 Hz, higher pitches were consistently perceived as faster than lower pitches; however, perceived tempo peaked somewhere between 440 Hz and 1760 Hz (varying across experiments), with 3520 Hz reliably perceived as slower than 440 Hz. In contrast, participants exposed to only one half of the pitch range showed a more traditional response, perceiving higher pitches within their assigned register as faster, regardless of the absolute pitch height of that register (Fig. 1B). Together, we interpreted these data as suggesting that perceived tempo depends on relative pitch within a context, and may reverse direction as pitch increases beyond two to three octaves above baseline—perhaps tying the effect to implicit knowledge about the span of human vocal ranges³⁷.

One limitation of a subjective rating paradigm like those used by our recent study and others^{16–18} is that they cannot distinguish whether biases in people's ratings arise at the level of perception or decision making. That is, we cannot tell in Fig. 1 whether participants rated certain tones as faster because they actually experienced

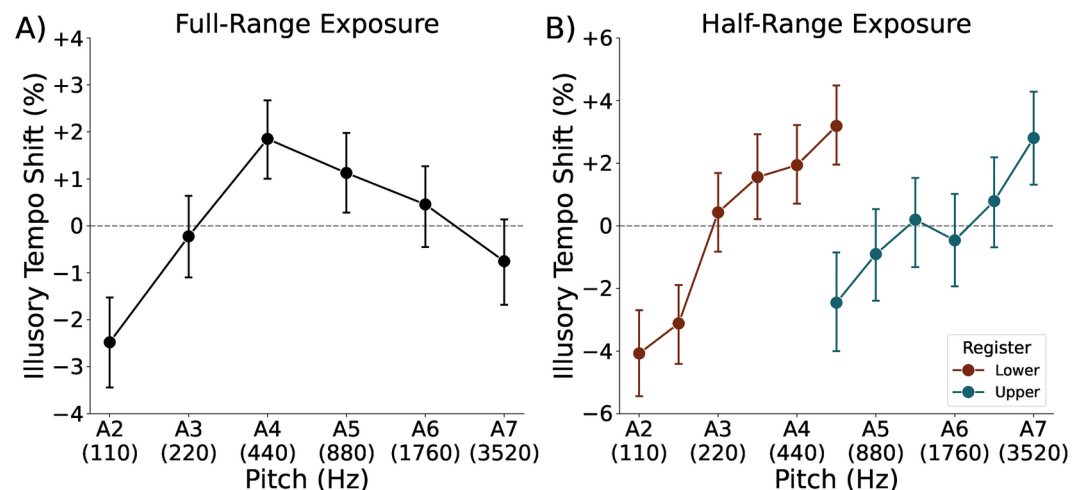


Figure 1. Effect of pitch on subjective tempo (data replotted from Pazdera and Trainor¹⁸). Subjective tempo ratings for isochronous repeating tones as a function of pitch, as observed when participants (A) heard tones between 110 Hz and 3520 Hz (Pazdera and Trainor¹⁸ Experiments 1–3), or (B) heard tones from only the lower or upper half of that range (Pazdera and Trainor¹⁸ Experiment 5). Error bars indicate within-subject 95% confidence intervals³⁸. Exposure to a five-octave range of frequencies produced an inverted U-shaped relation between pitch and perceived tempo, whereas exposure to either half of that range produced a bias to perceive higher pitches as faster regardless of absolute pitch.

them as faster in real-time while listening to them, or because they implicitly (or explicitly) used pitch as post hoc evidence for how fast the tones were. To address this limitation, we designed the present study to test the perceived tempo of the same pitches as our previous experiments¹⁸ in the absence of an explicit decision-making task. If the previously observed effects of pitch on timing are preserved even in the absence of decision-making, then it would provide evidence for a perceptual origin of this illusory tempo.

Outside of subjective ratings, the most common measurement of human timing is through sensorimotor tasks, usually involving tapping^{39,40}. Both synchronization^{21,23} and continuation^{19,20,41} tapping tasks have been used previously to study the effects of pitch distance, pitch direction, and contour changes on perceived tempo. We therefore adopted a synchronization–continuation tapping task for the present study, to determine whether pitch height also influences sensorimotor timing. In a synchronization–continuation task, participants first synchronize their movements to a metronome or other pacing signal, and then try to continue moving at the same tempo even after the signal ends^{42–44} or switches to the participant's control, such that their taps now trigger the stimulus^{19,45}. If the pitch of the pacing signal biases the perception of its tempo, then participants should synchronize to different pitches as if they were played at different tempos. For example, participants might tap at an earlier phase when synchronizing to pitches they perceive as speeding up^{21,23}, or may tap with a shorter period when asked to continue moving at the tempo of the pacing signal^{19,20,41}.

In the present study, we tested whether the same effects of pitch on subjective tempo ratings would carry over to sensorimotor timing in a synchronization–continuation task. In Experiment 1, we attempted to replicate the U-shaped effect from the subjective rating task by Pazdera and Trainor, Experiments 1–3¹⁸ (Fig. 1A) by exposing participants to pitches spanning a five-octave range from A2 (110 Hz) to A7 (3520 Hz). In Experiment 2, we attempted to replicate the simple higher–faster bias from Pazdera and Trainor, Experiment 5¹⁸ (Fig. 1B) by exposing participants to tones from only the lower half or upper half of that range.

If pitch biases timing on-line during perception, we should observe a greater negative mean asynchrony (the extent to which tapping anticipates the stimulus) and shorter inter-tap intervals for the same pitches that participants rated as fastest in our recent subjective tempo experiments¹⁸. Specifically, when participants in Experiment 1 are exposed to stimuli spanning 110–3520 Hz, we should observe a U-shaped effect in which the most negative mean asynchrony and shortest inter-tap intervals fall between A4 (440 Hz) and A6 (1760 Hz). Participants in Experiment 2 should instead show greater negative mean asynchrony and shorter inter-tap intervals for relatively high pitches within their assigned register, regardless of absolute pitch. However, if pitch affects subjective tempo at a later stage of processing—such as during decision-making—we should not observe the same effects of pitch on sensorimotor timing that have been observed during subjective judgment tasks.

Methods

Given the similarity of our methods for Experiments 1 and 2, we present the methods for both experiments together, making note of any differences between the experiments where relevant. In both experiments, participants completed a synchronization–continuation tapping task in which they first synchronized finger taps to an isochronous, repeating complex tone, and then attempted to keep the tone playing at the same tempo by continuing to tap at the same rate after being given control of its timing. In Experiment 1, the repeating tone on any given trial was an A-note ranging from A2 (110 Hz) to A7 (3520 Hz). In Experiment 2, each participant heard tones only from the lower half (110–622.3 Hz) or upper half (622.3–3520 Hz) of the full range, and the repeating tone was instead any A or D♯ from that assigned register. By exposing some participants to the full five-octave range and other participants to half that range, we sought to disambiguate the effects of absolute pitch versus relative pitch within a context. We also tested participants in both experiments at two tempos: a 600 ms interonset interval (IOI) on some trials, and a faster 400 ms IOI on other trials. Our primary measures of interest were the asynchrony of participants' taps during synchronization tapping and the duration of their inter-tap intervals during continuation tapping.

Participants

Data for Experiment 1 were collected between April 2022 and January 2023, under special COVID-19 safety protocols, including mask requirements for all participants and experimenters. Thirty-nine undergraduate students (32 female, 6 male, 1 unreported) from McMaster University completed the experiment for course credit. Ages ranged from 17–30 ($M = 18.8$, $SD = 2.2$). An additional two participants completed the experiment, but were excluded from analysis based on performance criteria (see Data Analysis for details). One was excluded for having an outlier mean phase of synchronization that indicated either anti-phase synchronization or tapping in response to tones, preventing alignment of their taps with the synchronization tones. The other was excluded both for lack of synchronization and for tapping with an inter-onset interval faster than 450 ms in the 600 ms IOI condition, indicating rushing through the continuation tapping task.

Data for Experiment 2 were collected between May and October 2023, with the COVID-19 safety protocols from Experiment 1 remaining in place for all participants collected before August. Forty-eight undergraduate students (36 female, 12 male) from McMaster University completed Experiment 2 for course credit. Ages ranged from 17–24 ($M = 18.3$, $SD = 1.1$). Twenty-four participants (16 female) were randomly assigned to hear lower-register tones only, while the remaining 24 (20 female) were assigned to hear upper-register tones only. An additional four participants (two lower register, two upper register) completed the experiment, but were excluded. Two were excluded for tapping with an inter-onset interval faster than 450 ms in the 600 ms IOI condition, indicating rushing through the task. The remaining two were excluded both for lack of synchronization (see Data Analysis) and for tapping at approximately 600 ms inter-onset intervals regardless of the stimulus IOI, suggesting a misunderstanding of the instructions. Additionally, three of the non-excluded participants showed anomalous synchronization or continuation tapping behavior for the first one to two blocks of the experiment,

but performed the task correctly on all remaining blocks. We chose to include these participants by dropping only the affected blocks.

Both experiments were performed in line with the principles of the Declaration of Helsinki, and approval for the present research was granted by the McMaster Research Ethics Board (No. 2609). All participants provided informed consent prior to their participation in our study. For information on our participants' musical training and how it influenced performance on our task, see the Supplementary Information.

Materials

All text used in both experiments was displayed in white, 72-point Arial font on a black background. All auditory stimuli were 250 ms complex tones, which we generated in Python by summing sine waves for the fundamental frequency and the first two overtones. We summed the sine waves with random phase, and reduced the amplitude of the first and second overtones by 6 dB and 9.5 dB respectively, relative to the fundamental. We then applied a percussive amplitude envelope⁴⁶ by adding a 10 ms linear rise followed by a 240 ms exponential fade. We applied an additional linear fade to the final 10 ms of the tone so that the amplitude ended at zero. Finally, we normalized the output loudness of all tones through the headphones to 75 dBA using a Brüel & Kjær Type 2239 sound-level meter. This calibration was performed once for each experiment prior to participant recruitment, and experimenters were instructed not to adjust the Arduino's loudness setting throughout the study. For Arduino compatibility, tones were saved as WAV files with a sampling rate of 22 kHz instead of the standard 44.1 kHz (see Apparatus). To account for the reduced sampling rate, we ensured that none of the frequencies used in our tone design (the highest being 10560 Hz, the second overtone of A7) exceeded the Nyquist frequency of our audio (11025 Hz).

Apparatus

We conducted audiometry for both experiments using a Grason–Stadler GSI-61 Clinical Audiometer. During the main experimental task, an Arduino Uno running custom C++ code controlled auditory stimulus presentation and tap detection. The Arduino used two attachments: an Adafruit Wave Shield for audio presentation and an Ohmite FSR01CE force sensitive resistor (FSR) as a tapping pad. Stimuli were presented over Senheiser HD280 Pro headphones plugged into the Wave Shield, and participants tapped on the FSR. We selected the Arduino, Wave Shield, and FSR configuration for two reasons. First this hardware configuration has been demonstrated to provide < 3 ms of latency and sub-millisecond precision for presenting stimuli in response to tapping⁴⁷. Second, tapping on an FSR helped minimize the audibility of the participant's taps, reducing its potential use as an auditory feedback cue.

The Arduino communicated over USB with a Windows 7 computer running a Python (version 3.6) program that used the PsychoPy library⁴⁸ to control the trial order and to log all data collected. All visuals were displayed on a 19-inch Dell 1908FP monitor with a resolution of 1280 × 1024 and a frame rate of 60 Hz. The computer sent the Arduino instructions regarding which experimental condition to use on each trial, and the Arduino returned a timestamp for each stimulus onset, tap onset, and tap release. In Experiment 2, we also recorded the peak pressure value of each tap on the FSR, and the timestamp of that peak pressure. Although we do not report on the pressure data in the present manuscript, we have included these measures in our open-access dataset (see Data Availability).

Design

Experiment 1 followed a 6 (Pitch Height) × 2 (Interonset Interval) within-subjects design. The pitch of the stimulus varied randomly between trials, and was one of A2 (110 Hz), A3 (220 Hz), A4 (440 Hz), A5 (880 Hz), A6 (1760 Hz), or A7 (3520 Hz) on each trial. The IOI of the synchronization tones also varied randomly between trials, and was either 600 ms (100 BPM) or 400 ms (150 BPM).

Experiment 2 maintained the same IOI manipulation, but instead followed a 2 (Register) × 6 (Pitch Height) × 2 (Interonset Interval) mixed design. Participants were randomly assigned to hear six unique pitches from either the lower half (lower register) or upper half (upper register) of the pitch range used in Experiment 1. The pitch of the stimulus varied between trials, with participants assigned to the lower register hearing one of A2 (110 Hz), D \sharp 3 (155.6 Hz), A3 (220 Hz), D \sharp 4 (311.1 Hz), A4 (440 Hz), or D \sharp 5 (622.3 Hz) on each trial. Those assigned to the upper register instead heard one of D \sharp 5 (622.3 Hz), A5 (880 Hz), D \sharp 6 (1244.5 Hz), A6 (1760 Hz), D \sharp 7 (2489.0 Hz), or A7 (3520 Hz) on each trial.

Procedure

Prior to the main synchronization–continuation task, we collected an audiogram from each participant in a sound-attenuated booth adjacent to the main testing room. Participants were instructed to detect the presence of a tone that was presented binaurally over headphones. In Experiment 1, hearing thresholds were tested for pure tones at the frequencies 125 Hz, 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz. We first tested frequencies in ascending order from 1000 Hz to 8000 Hz, then tested in descending order from 1000 Hz to 125 Hz. At each frequency, we measured hearing threshold via a staircase procedure that began at +30 dB HL, reduced by 10 dB for every correct detection (to a minimum of -10 dB HL), and increased by 5 dB for every miss. We averaged the two measurements taken at 1000 Hz to obtain a single threshold value for this frequency. Experiment 2 followed the same audiometry procedure, with the exception that participants assigned to the lower register received only the descending half of the audiogram, while those assigned to the upper register received only the ascending half of the audiogram.

The participant then returned to the main testing room, and we measured their spontaneous motor tempo by asking them to tap on the FSR with the index finger of their dominant hand “at the rate that feels most natural

and comfortable” to them. A fixation cross appeared on the screen to indicate when they should begin tapping, and the measurement ended after 30 taps.

We then introduced the participant to the synchronization–continuation task. We informed them that they would hear several tones playing at a steady pace on each trial, and that after several beats they would gain control of the tones so that their tapping would cause the tones to play. We emphasized to the participant that their goals were to first synchronize their taps so that their finger “lands exactly when the next tone will begin,” and then to “keep the tones playing at the same steady pace throughout the trial by continuing to tap at the same rate” even after gaining control of the tones.

Each trial consisted of 24 repetitions of a tone, which varied in pitch across trial. A fixation cross preceded the first tone of the trial by a uniformly jittered 1000 to 1500 ms. The first eight tones of the trial (the “synchronization tones”) then played at an isochronous IOI of either 600 ms or 400 ms, with no auditory feedback provided in response to the participant’s taps. The remaining 16 tones (the “continuation tones”) were instead initiated by the participant’s tapping. Continuation tones played 20 ms after the participants’ taps, a delay which we introduced in order to help stabilize the transition from synchronization to continuation⁴⁵. If the participant tapped while the previous tone was still playing, no new tone would be generated until the participant lifted their finger off the FSR and tapped again. Immediately after the final continuation tone finished playing, the fixation cross disappeared and participants were shown the standard deviation (in milliseconds) of their continuation-phase inter-tap intervals alongside instructions to keep this score as low as possible by tapping steadily¹⁹. We selected the standard deviation as feedback, so as not to inform participants whether they had sped up or slowed down during continuation tapping. This score remained onscreen for 2 s, after which a blank screen was displayed for 2 s before the next trial began.

Participants completed six practice trials and 120 experimental trials of the synchronization–continuation task. All practice trials used an IOI of 500 ms, and each used a different one of the six pitch height conditions, randomly ordered. The experimental trials were divided into five blocks of 24. Each block contained two trials of each pitch height and IOI combination, randomly ordered. Self-paced breaks were intended to be administered between blocks, but due to a software error were instead administered after the 12th, 24th, 36th, and 48th trials in Experiment 1. This error was corrected for Experiment 2.

Data analysis

Tap debouncing

To pre-process our tapping data, we first removed any falsely-detected taps through a debouncing procedure. Early testing revealed that the force sensitive resistor sometimes erroneously detected multiple onsets from a single tap if the tap pressure fluctuated above and below the minimum detection threshold during initial contact. To avoid analyzing these false taps and releases, we programmed the Arduino to automatically ignore any below-threshold pressure reading within 20 ms of a tap onset, treating the participant’s finger as remaining on the FSR. Similar pressure fluctuations could also occur as the participant lifted their finger off of the resistor; therefore, the Arduino also ignored any above-threshold pressure reading within 80 ms of the previous tap’s release, as if the participant’s finger remained off the FSR. A review of our tap duration data suggested that some additional false releases occurred up to 30 ms after tap onset, and so were not automatically rejected by the Arduino. Any time the participant’s finger remained on the FSR for longer than the 80 ms debounce period, a false tap onset was detected exactly 80 ms after these false releases. We therefore marked any tap that began 80 ms after a tap whose recorded duration was 30 ms or less as a false tap, totaling 0.7% of all taps. We then merged each false tap with its preceding tap by replacing the preceding tap’s release time with the release time of the false tap.

After applying these debouncing methods, we separated taps into synchronization taps and continuation taps. We categorized any tap occurring between the third and ninth tone as a synchronization tap (including the tap that triggered the first continuation tone), and any tap occurring after the ninth tone as a continuation tap. We did not analyze taps before the third tone, as a review of the data suggested that synchronization tended to stabilize during the third inter-stimulus interval.

Synchronization tap processing

For each synchronization tap, we identified the pair of tones between which it occurred, and calculated the fraction of the stimulus IOI that had elapsed before the tap occurred. We then converted this fraction to a phase value relative to the preceding tone, from 0 to 2π . Taps in the interval between the final synchronization tone and the first continuation tone (i.e., eighth-interval taps) were a unique case, in that their phase values depended on the virtual time that the next synchronization tone would have fallen, had the pacing signal continued. We only analyzed eighth-interval taps that fell before this hypothetical ninth synchronization tone would have played (92.7% and 91.7% of eighth-interval taps in Experiment 1 and Experiment 2, respectively).

Because the Arduino can only perform one process at a time, it can only read from the FSR in between other operations. As such, longer operations can create “blind spots” in tap detection. In particular, we identified two operations lasting longer than 1 ms: loading an audio file from the Wave Shield’s memory card took approximately 6 ms and initializing audio playback took approximately 2 ms. In Experiment 1, file loading occurred immediately before playback; therefore, whenever the participant tapped within the 8 ms prior to a tone onset, the Arduino would instead detect the tap immediately after tone onset, producing an apparent relative phase of 0. In Experiment 2, we reduced the pre-stimulus blind spot to 2 ms by adjusting the Arduino routine to pre-load each tone immediately after the previous tone ended—shifting the majority of the blind spot close to the offbeat, when participants were unlikely to tap. Due to the blind spot, a majority of taps recorded as occurring at the same millisecond as a tone were actually taps initiated during the preceding blind spot (see

Supplementary Fig. S1 for an illustration). We therefore excluded from further analysis all synchronization taps with a relative phase of 0, totaling 2.5% of synchronization taps in Experiment 1 and 2.0% in Experiment 2.

We next converted each tap's phase within its inter-stimulus interval into an asynchrony relative to the tone at which it was most likely targeted. Figure 2 illustrates this conversion process in one example participant. For each participant, i , we first calculated the circular mean of their tapping phases, μ_i , as well as the anti-phase of their average tap, $\mu_i - \pi$ (the dashed and dotted lines in Fig. 2A, respectively). Our conversion method next relies on two assumptions: 1) taps at the mean phase were made in anticipation of the subsequent tone (as participants were instructed), and 2) each person's tap timing is symmetrically distributed around the mean. Under these two assumptions, any taps with a relative phase in the range $[\mu_i - \pi, 2\pi)$ (i.e., taps to the right of the dotted line in Fig. 2A) were most likely targeted at the subsequent tone, and any taps with a relative phase in the range $(0, \mu_i - \pi)$ (those to the left of the dotted line) were most likely targeted at the preceding tone. To ensure that the first assumption held, we excluded participants whose mean phase fell closer to the preceding tone than to the subsequent tone, as well as those with a mean resultant vector length less than 1/3 (indicating lack of synchronization; see Participants for exclusion counts). Having aligned taps with the tones they most likely targeted, we concluded by calculating the percent of an interonset interval that each tap occurred before or after its target. This procedure produced a percent asynchrony for each tap (Fig. 2B), where a negative asynchrony indicates that the tap preceded its target tone and a positive asynchrony indicates the tap occurred after its target.

Continuation tap processing

For each continuation tap, we calculated the preceding inter-tap interval as the number of milliseconds elapsed since the onset of the previous tap. We then divided the inter-tap interval by the IOI of the synchronization tones on that trial to obtain a score we will refer to as the *relative inter-tap interval*. A relative inter-tap interval of 1 would indicate perfectly accurate continuation tap timing, with the inter-tap interval being equal to the IOI. Scores greater than 1 indicate that the participant tapped slower than the synchronization tones, whereas scores less than 1 indicate faster tapping. To remove participants who attempted to rush through the experiment by tapping quickly regardless of the stimulus IOI, we excluded participants with a mean relative inter-tap interval of less than 3/4 in the 600 ms condition. Finally, to eliminate individual intervals where the participant paused or tapped excessively quickly, we excluded from further analysis any continuation tap with a relative inter-tap interval greater than 2 or less than 1/2, respectively (0.2% of continuation taps in Experiment 1 and 0.5% in Experiment 2).

Audiometry

As the frequencies tested during the audiogram did not match the fundamental frequencies of our stimuli, we used a cubic spline procedure to interpolate or extrapolate hearing thresholds for each stimulus fundamental frequency. To account for the log-linear perception of frequency, we fit the cubic spline to hearing thresholds as a function of the log of frequency, rather than frequency in Hz.

Statistical analysis

We analyzed the effect of pitch on sensorimotor timing in Experiment 1 using a pair of linear mixed models, one to predict percent asynchrony and one to predict relative inter-tap interval. Models were fit to data at the level of participant averages, with all taps from a participant averaged within each condition. Each model contained fixed slopes for interonset interval (coded as a categorical factor), the linear and quadratic effects of pitch height (coded as a second-order polynomial regression), as well as the interaction between IOI and pitch height. We fit

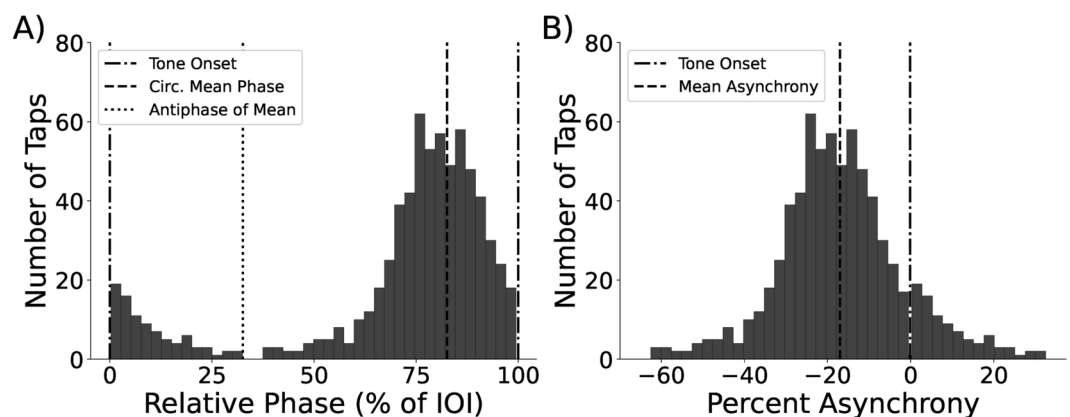


Figure 2. Example asynchrony calculation for one participant. (A) The distribution of synchronization tap phases, relative to the preceding tone onset, from one example participant. The dashed line denotes their circular mean tapping phase, while the dotted line indicates the antiphase of their tapping. We treated all tap times that were closer to the preceding stimulus than to the participant's mean tapping phase (i.e., all taps to the left of the antiphase line) as positive asynchronies targeted at the preceding tone. We treated all other tap times as negative asynchronies made in anticipation of the next tone. (B) The resulting asynchrony values, expressed as a percent of the interonset interval of the stimulus.

random participant-level intercepts and slopes for interonset interval, under the assumption that participants would differ in their ability to synchronize with and maintain different tempos.

Each model also included each participant's hearing threshold for the fundamental frequency of each pitch as a covariate, as determined from the participant's audiogram. We mean centered thresholds within-subject, such that the models would test whether participants tapped at an earlier phase or faster tempo when synchronizing with pitches to which they were relatively (in)sensitive. Although we calibrated all tones to be equally loud for participants with normal hearing, individual hearing loss may still cause some pitches to be perceived as quieter than others. As there is some evidence that loudness can influence perceived tempo, with increasing loudness associated with a faster tempo¹⁶, differences in perceived loudness across our stimuli introduce a potential confound. For example, the prevalence of high frequency hearing loss may cause extremely high pitches to be perceived as quieter, and therefore slower, than other stimuli. By accounting for any hearing loss in our participants, we can separate the effects of pitch and loudness on sensorimotor timing.

We repeated the same statistical analysis for Experiment 2, with the exception that we fit separate mixed models for each of the two register conditions, as participants in each register were exposed to largely non-overlapping pitch ranges.

As a final test of the overall impact of hearing thresholds in our study, we also separately tested their effect on mean asynchrony and relative inter-tap interval, pooling the data from both experiments. To do so, we used a pair of minimal mixed effects models with only a fixed slope for hearing threshold and random participant-level intercepts.

To evaluate the significance of all fixed effects in all models, we performed *F*-tests using the lmerTest package⁴⁹. In cases where we observed a significant effect of pitch height, we determined whether each of the linear and/or quadratic terms were significant by using lmerTest to perform post hoc *t*-tests with Holm–Bonferroni correction on the slope estimates.

Results

We begin by discussing the effect of pitch on mean asynchrony during synchronization tapping in each experiment, and then discuss its effect on the tempo of continuation tapping. Audiometry detected no hearing loss (threshold > 25 dB) in any participants for any of the fundamental frequencies used in the main experimental task. Pooling across both experiments, we identified a significant negative correlation between pitch height and hearing thresholds, $r(520) = -.34$, $p < .001$, such that participants tended to have more sensitive hearing (lower thresholds) for higher pitches. For information on spontaneous motor tempo and its relation to sensorimotor timing in our synchronization-continuation task, see the Supplementary Information.

Synchronization tapping

Figure 3 shows mean asynchronies for synchronization tapping to isochronous tones of each combination of pitch and interonset interval, expressed as a percent of that IOI. Figure 3A presents the data from Experiment 1, in which participants heard the full range of pitches spanning from A2 (110 Hz) to A7 (3520 Hz). Figure 3B presents the data from Experiment 2, in which each participant was only exposed to tones from the lower (110–622.3 Hz) or upper (622.3–3520 Hz) half of this range. Larger negative asynchronies indicate earlier tapping relative to the tone being anticipated. In Experiment 1, participants tapped earliest while synchronizing to A4

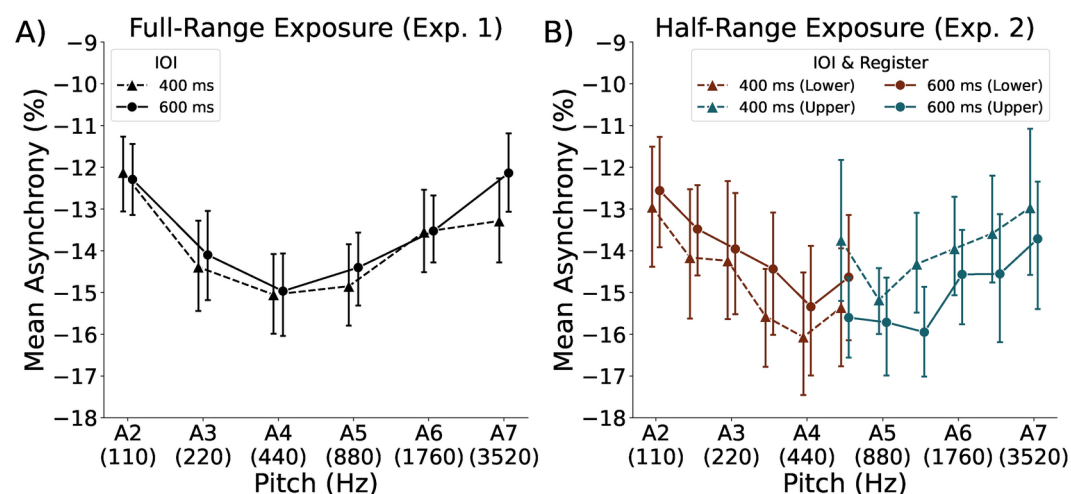


Figure 3. Effect of pitch on mean asynchrony. Markers indicate participants' average asynchrony when synchronizing to a pacing signal of a given pitch and interonset interval (IOI), expressed as a percentage of the IOI. Larger negative values indicate earlier tapping, relative to the beat. Pitch height exerted a U-shaped effect on asynchrony across the frequency range of 110–3520 Hz, regardless of whether participants heard (A) the full five-octave range in Experiment 1 or (B) tones from only one half of the full range in Experiment 2. Participants tapped earliest when synchronizing to pitches between A4 and A5, and latest when synchronizing to A2 and A7. Error bars indicate within-subject 95% confidence intervals³⁸.

(440 Hz) and latest while synchronizing to A2 (110 Hz) and A7 (3520 Hz). In Experiment 2, participants in the lower register generally tapped earlier to higher tones, whereas those in the upper register tapped later to higher tones.

Our linear mixed model analysis for Experiment 1 confirmed that pitch height significantly affected the percent asynchrony of synchronization taps, $F(2, 390) = 36.39, p < .001$. The effect of pitch was characterized by a significant positive quadratic slope, $t(390.0) = 8.47, p_{adj} < .001$, and a nonsignificant linear slope, $t(390.0) = 0.69, p_{adj} = .493$, consistent with the mostly symmetrical U-shape of the data in Fig. 3A. The main effect of IOI was nonsignificant, $F(1, 39) = 0.19, p = .665$, suggesting that participants tapped at a consistent percent asynchrony regardless of the IOI of the stimulus. The interaction between IOI and pitch was also nonsignificant, $F(2, 390) = 1.00, p = .368$, indicating that pitch did not differentially affect sensorimotor synchronization at different tempos. Higher (less sensitive) hearing thresholds predicted slightly later tapping on average ($\beta = 0.038, SE = 0.028$), but not significantly so, $F(1, 401.4) = 1.83, p = .177$.

Among participants assigned to the lower register in Experiment 2, pitch again significantly affected the percent asynchrony of synchronization tapping, $F(2, 240) = 19.12, p < .001$. The effect of pitch was characterized by both a significant negative linear slope, $t(240) = -5.75, p_{adj} < .001$, and a significant positive quadratic slope, $t(240) = 2.42, p_{adj} = .016$, matching the hook-like shape of the lower-register data in Fig. 3B. Stimulus IOI did not affect percent asynchrony, $F(1, 24) = 0.29, p = .593$, with participants tapping at a similar phase when synchronizing to both 400 ms and 600 ms intervals. The interaction between IOI and pitch was also nonsignificant, $F(2, 240) = 0.16, p = .854$, suggesting that pitch affected percent asynchrony similarly across tempos. Higher hearing thresholds again predicted slightly later tapping on average ($\beta = 0.046, SE = 0.046$), but not significantly so, $F(1, 240) = 1.01, p = .317$.

Among participants assigned to the upper register in Experiment 2, pitch also significantly affected percent asynchrony, $F(2, 240) = 12.77, p < .001$. The effect of pitch was characterized by both a significant positive linear slope, $t(240) = 4.38, p_{adj} < .001$, and a significant positive quadratic slope, $t(240) = 2.82, p_{adj} = .005$, consistent with the check-mark shape of the upper-register data in Fig. 3B. IOI did not significantly affect percent asynchrony, $F(1, 24) = 1.37, p = .253$, and the interaction between IOI and pitch was not significant, $F(2, 240) = 0.43, p = .651$. Higher hearing thresholds significantly predicted later tapping in the upper register ($\beta = 0.092, SE = 0.031$), $F(1, 240) = 8.88, p = .003$.

In isolation, hearing thresholds also significantly predicted percent asynchrony across both experiments combined, $F(1, 435) = 7.76, p = .006$. Specifically, participants tended to tap with less negative asynchrony when synchronizing to pitches for which they had less sensitive hearing.

Continuation tapping

Figure 4 shows average relative inter-tap intervals during continuation tapping, following synchronization to tones of each pitch and IOI. Figure 4A presents the data from Experiment 1, and Fig. 4B presents the data from Experiment 2. Smaller relative inter-tap intervals indicate a faster tapping tempo, and values less than one indicate that the participant tapped faster than the target tempo. Qualitatively, the data suggest that participants tapped faster than the target tempo in all conditions across both experiments, and sped up more when asked to tap at 600 ms intervals than at 400 ms ones. In Experiment 1, participants tapped at the slowest tempo after synchronizing to A2 (110 Hz) and A7 (3520 Hz), and fastest after synchronizing to pitches between A4 (440 Hz) and A6 (1760 Hz). In Experiment 2, participants in the lower register generally tapped at a faster tempo after synchronizing to higher pitches, whereas those in the upper register generally tapped slower after synchronizing to higher pitches, although this pattern was clearer at a 400 ms IOI than at 600 ms.

Our mixed modeling analysis for Experiment 1 confirmed that pitch height significantly affected relative inter-tap intervals during continuation tapping, $F(2, 390) = 6.27, p = .002$. Similar to our asynchrony results, the effect of pitch was characterized by a significant positive quadratic slope, $t(390) = 3.53, p_{adj} < .001$, and a nonsignificant linear slope $t(390) = 0.18, p_{adj} = .861$. Participants tapped slowest after synchronizing to low and extremely high pitches. IOI also significantly affected relative inter-tap interval, $F(1, 39) = 48.51, p < .001$. Participants tapped faster than the synchronization tones at both tempos, but underestimated the IOI more when asked to match the slower, 600 ms IOI. The interaction between IOI and pitch was not significant, $F(2, 390) = 0.16, p = .855$. Higher hearing thresholds predicted slower tapping on average ($\beta = 2.4 \times 10^{-4}, SE = 1.6 \times 10^{-4}$), but not significantly so, $F(1, 390) = 2.36, p = .126$.

Among participants assigned to the lower register in Experiment 2, our analysis indicated that pitch again significantly affected relative inter-tap intervals during continuation tapping, $F(2, 240) = 6.57, p = .002$. The effect of pitch was characterized by a significant negative linear slope, $t(240) = -3.94, p_{adj} = .002$, and a nonsignificant positive quadratic slope $t(240) = 1.48, p_{adj} = .140$. Within the lower register, participants tapped at a significantly faster tempo after synchronizing to higher pitches. IOI also significantly affected relative inter-tap interval, $F(1, 24) = 18.82, p < .001$, with participants underestimating 600 ms IOIs to a greater extent than 400 ms ones. Although Fig. 4B suggests a slightly stronger effect of pitch at 400 ms than at 600 ms, the interaction between tempo and pitch did not reach significance, $F(2, 240) = 2.51, p = .083$. Lastly, hearing thresholds significantly predicted inter-tap intervals ($\beta = -5.8 \times 10^{-4}, SE = 2.3 \times 10^{-4}$), $F(1, 240) = 6.42, p = .012$, such that participants assigned to the lower register tended to tap at a faster tempo after synchronizing with pitches to which they were less sensitive.

Among participants assigned to the upper register in Experiment 2, pitch also significantly affected relative inter-tap intervals, $F(2, 240) = 5.87, p = .003$. The effect of pitch was characterized by a significant positive linear slope, $t(240) = 3.00, p_{adj} = .006$, and a nonsignificant positive quadratic slope $t(240) = 1.86, p_{adj} = .064$. Opposite to the pattern observed among participants in the lower register, participants in the upper register tapped slower after synchronizing to higher pitches. Stimulus IOI again significantly affected relative inter-tap interval, $F(1, 24) = 50.51, p < .001$, in a pattern consistent with that of participants in the lower

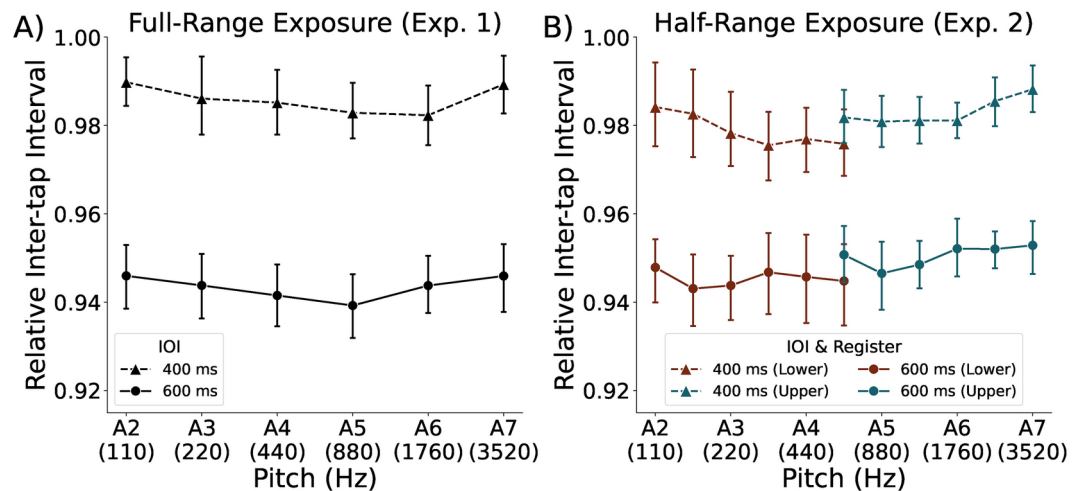


Figure 4. Effect of pitch on continuation tapping rate. Markers indicate participants' average inter-tap interval relative to the target interonset interval (IOI), during continuation tapping following synchronization to a pacing signal of a given pitch and IOI. Smaller values indicate faster tapping rates, and values less than one indicate that participants tapped faster than the target interval. Pitch height exerted a U-shaped effect on relative inter-tap interval across the frequency range of 110–3520 Hz, regardless of whether participants heard (A) the full five-octave range in Experiment 1 or (B) tones from only one half of the full range in Experiment 2. Participants tapped at a relatively slow tempo after synchronizing to A2 or A7, and tapped fastest after synchronizing to pitches in the middle of the range. Error bars indicate within-subject 95% confidence intervals³⁸.

register. The interaction between tempo and pitch was nonsignificant, $F(2, 240) = 0.39$, $p = .680$. Hearing thresholds did not significantly predict relative inter-tap intervals among participants in the upper register ($\beta = 2.4 \times 10^{-5}$, $SE = 1.2 \times 10^{-4}$), $F(1, 240) = 0.04$, $p = .846$.

In isolation, hearing thresholds also did not significantly predict relative inter-tap intervals across both experiments combined, $F(1, 435) = 0.23$, $p = .629$.

Discussion

Using two synchronization–continuation tapping experiments, we investigated how the pitch of a rhythmic auditory stimulus influences sensorimotor timing across a wide range of fundamental frequencies from 110–3520 Hz. To differentiate effects of absolute pitch versus relative pitch within a context, we exposed some participants to tones from the full five-octave range (Experiment 1) and others to only the lower or upper half of the full range (Experiment 2). We observed a U-shaped relation between pitch and mean asynchrony during synchronization tapping (Fig. 3), as well as between absolute pitch and inter-tap intervals during continuation tapping (Fig. 4). Participants tapped with the most negative asynchronies and shortest inter-onset intervals when synchronizing to moderately high pitches between A4 (440 Hz) and A6 (1760 Hz), and tapped with the least negative asynchronies and longest inter-onset intervals when synchronizing to A2 (110 Hz) and A7 (3520 Hz). This pattern was observed regardless of whether participants heard tones drawn from the full five-octave range, or from only one half of the full range, suggesting that the U-shaped effect depends on absolute pitch height rather than relative pitch. Furthermore, pitch caused a similar percent change in sensorimotor timing regardless of the true interonset interval. Importantly, the reversal in timing at upper octaves could not be entirely explained by high-frequency hearing loss in the present sample. Although higher hearing thresholds did predict less negative asynchronies when pooling across both experiments, pitch remained a significant predictor of sensorimotor timing after accounting for hearing thresholds as a covariate. For a discussion of why pitch may have affected mean asynchrony more strongly than relative inter-tap interval in the present study, see the Supplementary Information.

One limitation of the present study in isolation is that our sensorimotor task cannot differentiate slower timing from more accurate timing. This confound arises because humans exhibit a general tendency to tap with a negative mean asynchrony and speed up during continuation tapping^{39,45}, and these same patterns appear in our data. Therefore, later synchronization tapping also means that participants' taps landed closer to the stimulus onsets. Similarly, slower continuation tapping rates mean that participants deviated less from the target tempo. There has been considerable attention in previous literature regarding a possible superiority of lower pitches for rhythm perception^{50–53} and for inducing spontaneous entrained movement or the urge to move^{54–56}. If this superiority exists, slower timing for A2 (110 Hz) in our study could be due to improved timing accuracy rather than biased timing, although some researchers have argued that the low-pitch superiority may itself originate from a timing bias⁵⁷.

We favor the interpretation of our results as a timing bias because a similar U-shaped relation was recently observed in subjective tempo ratings for isochronous, repeating tones spanning the same 110–3520 Hz

range (Fig. 1A)¹⁸. Participants in the present study tended to tap at an earlier phase and faster tempo when synchronizing to pitches that were rated as subjectively faster during these prior experiments. The present replication of this U-shaped effect during a sensorimotor task—in the absence of explicit decision-making—supports the hypothesis that pitch influences timing on-line during perception, rather than post hoc during decision-making.

However, our Experiment 2 results deviate substantially from previous subjective tempo ratings collected from participants who were similarly exposed to only the lower half (110–622.3 Hz) or upper half (622.3–3520 Hz) of the full frequency range¹⁸. In the present study, participants who only heard tones from the upper register tapped as if higher pitches were played at a slower tempo. In contrast, upper-register participants in Pazdera and Trainor, Experiment 5¹⁸, rated higher pitches as subjectively faster (Fig. 1B). One possible explanation is that the U-shaped relation between absolute pitch and perceived tempo has distinct origins from the “higher = faster” bias observed by Pazdera and Trainor, Experiment 5¹⁸, that seemingly depended on relative pitch within a context.

If true, it may be the case that one effect dominates the other under specific conditions, possibly depending on the salience of relative pitch across trials. We note two factors that may have collectively increased the salience of relative pitch in Pazdera and Trainor, Experiment 5¹⁸, above the several experiments that have produced U-shaped curves. First, participants heard the same pitch repeated substantially more times in a row in the present study's two synchronization–continuation experiments (24 tones) than in the subjective rating experiments (five tones). Second, the use of only a single pitch chroma (A) in our Experiment 1 and Pazdera and Trainor, Experiments 1–3¹⁸, may have reduced the salience of pitch changes between trials. Together, the combination of short trials and chroma variability may have produced a stronger sense of relative pitch across trials in Pazdera and Trainor, Experiment 5¹⁸, than in those experiments that produced U-shaped effects. However, further research will be necessary to conclusively demonstrate that the salience of relative pitch determines which effect dominates.

Based on the present study, the U-shaped effect of pitch on perceived tempo likely depends on absolute pitch and arises on-line during perception. The question remains as to whether this effect is innate—perhaps due to the biophysics of the auditory periphery—or learned from real-world correlations²⁹. The dependence on absolute pitch supports the former. Furthermore, the real-world correlation between pitch and timing in speech and music is generally assumed to be a higher–faster relation^{16,17,33–35}, not a U-shaped one, although future cross-cultural corpus analyses will be necessary in order to substantiate this claim. If the U-shaped effect of pitch is innate, it may relate to other known quadratic effects that depend on absolute frequency and/or pitch. For example, simulations have shown sub-cortical neural synchronization strength to be weakest for sounds in the 500–1000 Hz range and strongest below 250 Hz and above 2000 Hz⁵⁸—closely resembling the trajectory of our data in Figs. 3 and 4. Alternatively, it may relate to pitch salience, which peaks at 700 Hz and falls off at higher and lower fundamental frequencies⁵⁹. Stimulus salience has been previously linked to perceived duration⁶⁰, although future research would need to test whether it also influences perceived tempo.

In contrast, the relation between higher relative pitch and faster perceived timing in Pazdera and Trainor, Experiment 5¹⁸, aligns better with the probable correlation between pitch and timing in human vocalization and instrumental music^{16,17,33–35}. It is therefore more likely learned from real-world correlations, possibly helping the brain to anticipate trajectories of pitch and timing during speech and music perception^{29,30}, thereby facilitating communication. In order to properly differentiate a learned tempo bias from an intrinsic property of the auditory system, future research might test whether new correlations between pitch and timing can be learned through novel musical exposure. Previous research suggests that humans implicitly learn both the melodic^{61–64} and rhythmic structure^{65,66} underlying their native musical culture, and that they can learn a new musical grammar even within a short period of exposure^{67–69}. However, the learning of cross-dimensional associations has received little investigation. Therefore, it would be useful to perform a direct test of whether people can learn—and have their timing biased by—novel pitch–timing correlations.

To summarize, we hypothesize two effects—an innate, U-shaped effect of absolute pitch in which perceived tempo peaks somewhere between A4 (440 Hz) and A6 (1760 Hz) and declines at higher and lower octaves, and a learned association between higher pitches and faster timing. Although many previous studies have observed faster perceived tempo for higher pitched or ascending tones^{17,25,26,34,70}, speech^{31,33}, and music^{16,24}, it is uncertain which of these two hypothesized effects were observed by which studies, due to the lack of stimuli exceeding 1000 Hz. Below this level, both effects of pitch are expected to produce similar patterns of results. To disambiguate the two, future research on pitch and timing integration should include pitch conditions spanning low (< 250 Hz), moderately high (500–1000 Hz), and extremely high (> 2000 Hz) frequencies. Regardless of whether the effect of pitch on perceived timing derives from a learned perceptual bias or innate biophysics, the present study demonstrates that pitch biases sensorimotor synchronization, and models of rhythm perception should be extended to account for this influence.

Data availability

We have made all data, code, and stimuli from both experiments publicly available on the Open Science Framework (<https://osf.io/7bptg/>) and GitHub (<https://github.com/jpazdera/IllusoryMotor>).

Received: 22 August 2024; Accepted: 30 April 2025

Published online: 16 May 2025

References

1. Large, E.W. et al. Dynamic models for musical rhythm perception and coordination. *Front. Comput. Neurosci.* **17**. <https://doi.org/10.3389/fncom.2023.1151895> (2023).

2. Palmer, C. & Demos, A. P. Are we in time? How predictive coding and dynamical systems explain musical synchrony. *Curr. Dir. Psychol. Sci.* **31**, 147–153. <https://doi.org/10.1177/09637214211053635> (2022).
3. Gámez, J. et al. Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Ann. N. Y. Acad. Sci.* **1423**, 396–414. <https://doi.org/10.1111/nyas.13671> (2018).
4. Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I. Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* **19**, 827–830. <https://doi.org/10.1016/j.cub.2009.03.038> (2009).
5. Patel, A. D., Iversen, J. R., Bregman, M. R. & Schulz, I. Studying synchronization to a musical beat in nonhuman animals. *Ann. N. Y. Acad. Sci.* **1169**, 459–469. <https://doi.org/10.1111/j.1749-6632.2009.04581.x> (2009).
6. Large, E. W. & Snyder, J. S. Pulse and meter as neural resonance. *Ann. N. Y. Acad. Sci.* **1169**, 46–57. <https://doi.org/10.1111/j.1749-6632.2009.04550.x> (2009).
7. Large, E. W., Almonte, F. V. & Velasco, M. J. A canonical model for gradient frequency neural networks. *Phys. D* **239**, 905–911. <https://doi.org/10.1016/j.physd.2009.11.015> (2010).
8. Large, E. W., Herrera, J. A. & Velasco, M. J. Neural networks for beat perception in musical rhythm. *Front. Syst. Neurosci.* **9**, 159. <https://doi.org/10.3389/fnsys.2015.00159> (2015).
9. Roman, I. R., Roman, A. S., Kim, J. C. & Large, E. W. Hebbian learning with elasticity explains how the spontaneous motor tempo affects music performance synchronization. *PLoS Comput. Biol.* **19**, e1011154. <https://doi.org/10.1371/journal.pcbi.1011154> (2023).
10. Cannon, J. Expectancy-based rhythmic entrainment as continuous Bayesian inference. *PLoS Comput. Biol.* **17**, e1009025. <https://doi.org/10.1371/journal.pcbi.1009025> (2021).
11. Friston, K. The free-energy principle: A unified brain theory?. *Nat. Rev. Neurosci.* **11**, 127–138. <https://doi.org/10.1038/nrn2787> (2010).
12. Koelsch, S., Vuust, P. & Friston, K. Predictive processes and the peculiar case of music. *Trends Cogn. Sci.* **23**, 63–77. <https://doi.org/10.1016/j.tics.2018.10.006> (2019).
13. Vuust, P. & Witek, M. A. G. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Front. Psychol.* **5**. <https://doi.org/10.3389/fpsyg.2014.01111> (2014).
14. Bose, A., Byrne, A. & Rinzel, J. A neuromechanistic model for rhythmic beat generation. *PLoS Comput. Biol.* **15**, e1006450. <https://doi.org/10.1371/journal.pcbi.1006450> (2019).
15. Egger, S. W., Le, N. M. & Jazayeri, M. A neural circuit model for human sensorimotor timing. *Nat. Commun.* **11**, 1–14. <https://doi.org/10.1038/s41467-020-16999-8> (2020).
16. Boltz, M. G. Illusory tempo changes due to musical characteristics. *Music. Percept.* **28**, 367–386. <https://doi.org/10.1525/mp.2011.28.4.367> (2011).
17. Collier, W. G. & Hubbard, T. L. Judgments of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicol. J. Res. Music Cogn.* **17**, 36–55. <https://doi.org/10.1037/h0094060> (1998).
18. Pazdera, J. K. & Trainor, L. J. Pitch-induced illusory percepts of time. *Attention Percept. Psychophys.* **87**, 545–564. <https://doi.org/10.3758/s13414-024-02982-8> (2024).
19. Ammirante, P., Thompson, W. F. & Russo, F. A. Ideomotor effects of pitch on continuation tapping. *Q. J. Exp. Psychol.* **64**, 381–393. <https://doi.org/10.1080/17470218.2010.495408> (2011).
20. Ammirante, P. & Thompson, W. F. Continuation tapping to triggered melodies: Motor resonance effects of melodic motion. *Exp. Brain Res.* **216**, 51–60. <https://doi.org/10.1007/s00221-011-2907-5> (2012).
21. Boasson, A. & Granot, R. Melodic direction's effect on tapping. In *Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music*, 110–119 (Thessaloniki, Greece, 2012).
22. Boltz, M. G. Tempo discrimination of musical patterns: Effects due to pitch and rhythmic structure. *Percept. Psychophys.* **60**, 1357–1373. <https://doi.org/10.3758/bf03207998> (1998).
23. Boasson, A. & Granot, R. Short latency effects of auditory frequency change on human motor behavior. *Audit. Percept. Cogn.* **2**, 98–128. <https://doi.org/10.1080/25742442.2019.1698264> (2019).
24. Gordon, M. S. & Ataucusi, A. Continuous sliding frequency shifts produce an illusory tempo drift. *JASA Express Lett.* **1**, 053202. <https://doi.org/10.1121/10.0005001> (2021).
25. Herrmann, B., Henry, M. J., Grigutsch, M. & Obleser, J. Oscillatory phase dynamics in neural entrainment underpin illusory percepts of time. *J. Neurosci.* **33**, 15799–15809. <https://doi.org/10.1523/jneurosci.1434-13.2013> (2013).
26. Herrmann, B., Henry, M. J., Scharinger, M. & Obleser, J. Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *Neuroimage* **101**, 370–379. <https://doi.org/10.1016/j.neuroimage.2014.07.026> (2014).
27. Cohen, J., Hansel, C. & Sylvester, J. D. A new phenomenon in time judgment. *Nature* **172**, 901–901. <https://doi.org/10.1038/172901a0> (1953).
28. Cohen, J., Hansel, C. & Sylvester, J. D. Interdependence of temporal and auditory judgments. *Nature* **174**, 642–644. <https://doi.org/10.1038/174642a0> (1954).
29. Jones, M. R. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychol. Rev.* **83**, 323–355. <https://doi.org/10.1037/0033-295x.83.5.323> (1976).
30. Henry, M. J. & McAuley, J. D. Perceptual distortions in pitch and time reveal active prediction and support for an auditory pitch-motion hypothesis. *PLoS ONE* **8**, e70646. <https://doi.org/10.1371/journal.pone.0070646> (2013).
31. Boltz, M. G. Memory for vocal tempo and pitch. *Memory* **25**, 1309–1326. <https://doi.org/10.1080/09658211.2017.1298808> (2017).
32. Pazdera, J. K. & Trainor, L. J. Bidirectional interactions of pitch and time. In *Proceedings of the 17th International Conference on Music Perception and Cognition*, 301–306 (Tokyo, Japan, 2023).
33. Feldstein, S. & Bond, R. N. Perception of speech rate as a function of vocal intensity and frequency. *Lang. Speech* **24**, 387–394. <https://doi.org/10.1177/002383098102400408> (1981).
34. Lake, J. I., LaBar, K. S. & Meck, W. H. Hear it playing low and slow: How pitch level differentially influences time perception. *Acta Physiol. (Oxf)* **149**, 169–177. <https://doi.org/10.1016/j.actpsy.2014.03.010> (2014).
35. Broze, Y. & Huron, D. Is higher music faster? Pitch-speed relationships in Western compositions. *Music. Percept.* **31**, 19–31. <https://doi.org/10.1525/mp.2013.31.1.19> (2013).
36. Black, J. W. Relationships among fundamental frequency, vocal sound pressure, and rate of speaking. *Lang. Speech* **4**, 196–199. <https://doi.org/10.1177/002383096100400402> (1961).
37. Kuhn, T. L., Wachhaus, G., Moore, R. S. & Pantle, J. E. Undergraduate nonmusic major vocal ranges. *J. Res. Music Educ.* **27**, 68–75. <https://doi.org/10.2307/3344893> (1979).
38. Loftus, G. R. & Masson, M. Using confidence intervals in within-subject designs. *Psychon. Bull. Rev.* **1**, 476–490. <https://doi.org/10.3758/bf03210951> (1994).
39. Repp, B. H. Sensorimotor synchronization: A review of the tapping literature. *Psychon. Bull. Rev.* **12**, 969–992. <https://doi.org/10.3758/bf03206433> (2005).
40. Repp, B. H. & Su, Y.-H. Sensorimotor synchronization: A review of recent research (2006–2012). *Psychon. Bull. Rev.* **20**, 403–452. <https://doi.org/10.3758/s13423-012-0371-2> (2013).
41. Ammirante, P. & Thompson, W. F. Melodic accent as an emergent property of tonal motion. *Empir. Musicol. Rev.* **5**, 94–107. <https://doi.org/10.18061/1811/47559> (2010).

42. Michon, J. A. *Timing in temporal tracking* (Institute for Perception RVO-TNO, Soesterberg, Netherlands, 1967).
43. Stevens, L. T. On the time-sense. *Mind* **43**, 393–404 (1886).
44. Wing, A. & Kristofferson, A. The timing of interresponse intervals. *Percept. Psychophys.* **13**, 455–460. <https://doi.org/10.3758/bf03205802> (1973).
45. Flach, R. The transition from synchronization to continuation tapping. *Hum. Mov. Sci.* **24**, 465–483. <https://doi.org/10.1016/j.humov.2005.09.005> (2005).
46. Schutz, M. & Vaisberg, J. M. Surveying the temporal structure of sounds used in Music Perception. *Mus. Percept.* **31**, 288–296. <https://doi.org/10.1525/mp.2014.31.3.288> (2012).
47. Schultz, B. G. & van Vugt, F. T. Tap Arduino: An Arduino microcontroller for low-latency auditory feedback in sensorimotor synchronization experiments. *Behav. Res. Methods* **48**, 1591–1607. <https://doi.org/10.3758/s13428-015-0671-3> (2016).
48. Peirce, J. et al. PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods* **51**, 195–203. <https://doi.org/10.3758/s13428-018-01193-y> (2019).
49. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softwa.* **82**. <https://doi.org/10.18637/jss.v082.i13> (2017).
50. Hove, M. J., Keller, P. E. & Krumhansl, C. L. Sensorimotor synchronization with chords containing tone-onset asynchronies. *Perception Psychophys.* **69**, 699–708. <https://doi.org/10.3758/bf03193772> (2007).
51. Hove, M. J., Marie, C., Bruce, I. C. & Trainor, L. J. Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proc. Natl. Acad. Sci.* **111**, 10383–10388. <https://doi.org/10.1073/pnas.1402039111> (2014).
52. Lenc, T., Keller, P. E., Varlet, M. & Nozaradan, S. Neural tracking of the musical beat is enhanced by low-frequency sounds. *Proc. Natl. Acad. Sci.* **115**, 8221–8226. <https://doi.org/10.1073/pnas.1801421115> (2018).
53. Repp, B. H. Phase attraction in sensorimotor synchronization with auditory sequences: Effects of single and periodic distractors on synchronization accuracy. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 290–309. <https://doi.org/10.1037/0096-1523.29.2.290> (2003).
54. Cameron, D. J. et al. Undetectable very-low frequency sound increases dancing at a live concert. *Curr. Biol.* **32**, R1222–R1223. <https://doi.org/10.1016/j.cub.2022.09.035> (2022).
55. Stupacher, J., Hove, M. J. & Janata, P. Audio features underlying perceived groove and sensorimotor synchronization in music. *Mus. Percept.* **33**, 571–589. <https://doi.org/10.1525/mp.2016.33.5.571> (2016).
56. Varlet, M., Williams, R. & Keller, P. E. Effects of pitch and tempo of auditory rhythms on spontaneous movement entrainment and stabilisation. *Psychol. Res.* **84**, 568–584. <https://doi.org/10.1007/s00426-018-1074-8> (2020).
57. Wojtczak, M., Mehta, A. H. & Oxenham, A. J. Rhythm judgments reveal a frequency asymmetry in the perception and neural coding of sound synchrony. *Proc. Natl. Acad. Sci.* **114**, 1201–1206. <https://doi.org/10.1073/pnas.1615669114> (2017).
58. Zuk, N. J., Carney, L. H. & Lalor, E. C. Preferred tempo and low-audio-frequency bias emerge from simulated sub-cortical processing of sounds with a musical beat. *Front. Neurosci.* **12**, 1. <https://doi.org/10.3389/fnins.2018.00349> (2018).
59. Terhardt, E., Stoll, G. & Seewann, M. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *J. Acoust. Soc. Am.* **71**, 679–688. <https://doi.org/10.1121/1.387544> (1982).
60. Matthews, W. J. & Meck, W. H. Temporal cognition: Connecting subjective time to perception, attention, and memory. *Psychol. Bull.* **142**, 865–907. <https://doi.org/10.1037/bul0000045> (2016).
61. Bharucha, J. J. Music cognition and perceptual facilitation: A connectionist framework. *Mus. Percept.* **5**, 1–30. <https://doi.org/10.2307/40285384> (1987).
62. Krumhansl, C. L. & Kessler, E. J. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychol. Rev.* **89**, 334–368. <https://doi.org/10.1037/0033-295x.89.4.334> (1982).
63. Trainor, L. J. & Trehub, S. E. A comparison of infants' and adults' sensitivity to Western musical structure. *J. Exp. Psychol. Hum. Percept. Perform.* **18**, 394–402. <https://doi.org/10.1037/0096-1523.18.2.394> (1992).
64. Trainor, L. J. & Trehub, S. E. Key membership and implied harmony in Western tonal music: Developmental perspectives. *Percept. Psychophys.* **56**, 125–132. <https://doi.org/10.3758/bf03213891> (1994).
65. Jacoby, N. & McDermott, J. H. Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* **27**, 359–370. <https://doi.org/10.1016/j.cub.2016.12.031> (2017).
66. Jacoby, N. et al. Commonality and variation in mental representations of music revealed by a cross-cultural comparison of rhythm priors in 15 countries. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-023-01800-9> (2024).
67. Loui, P., Wessel, D. L. & Kam, C. Humans rapidly learn grammatical structure in a new musical scale. *Mus. Percept.* **27**, 377–388. <https://doi.org/10.1525/mp.2010.27.5.377> (2010).
68. Loui, P. Learning and liking of melody and harmony: Further studies in artificial grammar learning. *Top. Cogn. Sci.* **4**, 554–567. <https://doi.org/10.1111/j.1756-8765.2012.01208.x> (2012).
69. Rohrmeier, M., Rebuschat, P. & Cross, I. Incidental and online learning of melodic structure. *Conscious. Cogn.* **20**, 214–222. <https://doi.org/10.1016/j.concog.2010.07.004> (2011).
70. Herrmann, B. & Johnsrude, I. S. Attentional state modulates the effect of an irrelevant stimulus dimension on perception. *J. Exp. Psychol. Hum. Percept. Perform.* **44**, 89–105. <https://doi.org/10.1037/xhp0000432> (2018).

Acknowledgements

The authors would like to acknowledge Sharanya Badalera, Natalie Jean-Marie, Emily Melo, and Olive Rinaldi for their assistance with data collection, Dave Thompson for his assistance with assembling the Arduino, Martín Miguel for his identification of the Arduino's pre-stimulus “blind spot”, and Teagan Lauriente for insightful discussion regarding Experiment 2. This research was funded by grants to LJT from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-05416), the Canadian Institutes of Health Research (MOP 153130), and the Canadian Institute for Advanced Research (CIFAR).

Author contributions

JKP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – original draft, Writing – review & editing. LJT: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-00827-4>

[0.1038/s41598-025-00827-4](https://doi.org/10.1038/s41598-025-00827-4).

Correspondence and requests for materials should be addressed to J.K.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025