Does Movement Amplitude of a Co-performer Affect Individual Performance in Musical Synchronization?

Ian D. Colley⬤, Manuel Varlet, Jennifer MacRitchie⬤
and Peter E. Keller

Abstract

Interpersonal coordination in musical ensembles often involves multisensory cues, with visual information about body movements supplementing co-performers' sounds. Previous research on the influence of movement amplitude of a visual stimulus on basic sensorimotor synchronization has shown mixed results. *Uninstructed* visuomotor synchronization seems to be influenced by amplitude of a visual stimulus, but *instructed* visuomotor synchronization is not. While music performance presents a special case of visually mediated coordination, involving both uninstructed (spontaneously coordinating ancillary body movements with co-performers) and instructed (producing sound on a beat) forms of synchronization, the underlying mechanisms might also support rhythmic interpersonal coordination in the general population. We asked whether visual cue amplitude would affect nonmusicians' synchronization of sound and head movements in a musical drumming task designed to be accessible regardless of musical experience. Given the mixed prior results, we considered two competing hypotheses. H1: higher amplitude visual cues will improve synchronization. H2: different amplitude visual cues will have no effect on synchronization. Participants observed a human-derived motion capture avatar with three levels of movement amplitude, or a still image of the avatar, while drumming along to the beat of tempo-changing music. The moving avatars were always timed to match the music. We measured temporal asynchrony (drumming relative to the music), predictive timing, ancillary movement fluctuation, and cross-spectral coherence of ancillary movements between the participant and avatar. The competing hypotheses were tested using conditional equivalence testing. This method involves using a statistical equivalence test in the event that standard hypothesis tests show no differences. Our results showed no statistical differences across visual cues types. Therefore, we conclude that there is not a strong effect of visual stimulus amplitude on instructed synchronization.

Keywords

Ancillary movements, entrainment, human movement, music cognition, musical movement, musical synchrony, sensorimotor synchronization

Introduction

In ensemble music performance, musicians use multisensory cues to achieve a synchronized sound. Such cues likely include: auditory feedback to reduce asynchronies and asynchrony variability (Chen et al., 2002); intrapersonal somatic cues such as head movements to reinforce a sense of musical meter (Phillips-Silver & Trainor, 2007, 2008); and visual cues to facilitate anticipation of upcoming

The MARCS Institute for Brain, Behaviour and Development, Western
Sydney University, Penrith, NSW, Australia

Corresponding author:
Ian D. Colley, The MARCS Institute for Brain, Behaviour and
Development, Western Sydney University, Locked Bag 1797, Penrith
NSW 2751, Australia.
Email: idcol001@gmail.com

temporal patterns in the music (Colley et al., 2018). Assuming co-performers in a musical environment can see each other, *intra*personal somatic cues may also become *inter*personal visual cues, such that one person's rhythmic body movements might be seen by another person. Indeed, mutual visual access among partners in previous work (a dyadic sensorimotor-synchronization task with musical sequences) was found to improve the synchrony of partners' ancillary head movements, as well as their synchronization with the target auditory stimulus (Colley et al., 2020).

Studies on pure visuomotor synchronization (no audio component) have shown mixed results regarding the effect of amplitude of a periodic visual stimulus on one's ability to synchronize with the stimulus. Participants were found to spontaneously synchronize forearm movements with an oscillating circle better with larger amplitudes of circle movement, even when the period duration was kept the same (Varlet et al., 2012). Additionally, postural sways showed greater phase entrainment with larger environmental stimulus movements (Dijkstra et al., 1994). In both cases, synchronization with the visual stimulus was considered *uninstructed*, meaning participants were spontaneously synchronizing their movements, possibly without awareness. On the other hand, research on *instructed* rhythmic synchronization suggests there is no effect of stimulus amplitude (de Rugy et al., 2008; Peper & Beek, 1998). Similarly, synchronizing finger taps with an image of a finger featuring apparent motion was not affected by the amplitude of the apparent motion (Hove & Keller, 2010). Additionally, synchronization tapping with a virtual conductor was not influenced by the amplitude of conductor gestures (Wöllner et al., 2012).

Regarding music-related ancillary movements, previous studies have demonstrated that ancillary movements generally play a role in communicating a performer's expressive intentions, with larger movements signaling increased expressive intensity (Davidson & Broughton, 2016; Luck et al., 2014; Thompson & Luck, 2011). However, the influence of the size of ancillary movements on co-performers' synchronization abilities has not been tested. This would be difficult to test, as any benefit of a co-performer on a partner's synchronization depends to some extent on the skill and reliability of the co-performers (Pecenka & Keller, 2011) as well as social motives (Lumsden et al., 2012). As such, it would be hard to have a consistent visual cue in the form of ancillary movements.

To further explore the role of visual stimulus amplitude on synchronization we focused on the role of range of motion—or *movement amplitude*—of a high-performing co-performer's movements on one's ability to synchronize with a concurrent musical beat. To address the issue of not having a reliable stimulus, we programmed a virtual co-performer. With this controllable stimulus, we tested whether larger body movements of a very accurate co-performer could improve the synchronization accuracy of an observer. Also, assuming the co-performer's movements were always matched to the musical beat (which we controlled for), then larger movements would produce higher velocities. Velocity has been shown to be an important factor in visually mediated synchronization in earlier work (Luck & Sloboda, 2008, 2009; Luck & Toiviainen, 2006; Varlet et al., 2014). Velocity is also important in conductor gestures such that musicians and nonmusicians synchronize best with movement featuring high rates of vertical velocity change (Colley et al., 2018).

Overall, there is some evidence that uninstructed visuomotor coordination is affected by stimulus amplitude, but there is also evidence that stimulus amplitude has weak to no effects on instructed visuomotor coordination. The aim of the current study was to test whether movement amplitude of a visual stimulus affects one's ability to synchronize in a musical situation, where synchronization among co-performers often involves visual cues. Another interesting aspect of musical synchronization is that synchrony is not necessarily instructed. Certainly the main objective in most music is to match sounds in time, and as such, audio-motor synchronization among performing musicians is instructed. However, occasional apparent visuomotor synchronization may be uninstructed, or, ancillary.

We tested the influence of stimulus amplitude by having research volunteers drum to the beat of specially composed pieces of tempo-changing ensemble music, while observing a virtual co-performer (avatar), whose movements were manipulated to exhibit various amplitudes of motion, but were always matched to the musical beat. We used drumming as opposed to finger-tapping because individuals tend to miss fewer beats when drumming (Madison et al., 2013; Manning et al., 2017), thus yielding higher-quality data. We recorded their drumming in order to measure the asynchrony of their drum strokes, and to quantify their predictive timing, which is the ability to anticipate upcoming beat intervals (Colley et al., 2017, 2018). We also motion capture recorded participants during the drumming task (Colley et al., 2018) to measure the synchrony of their ancillary head movement with the avatar using cross-spectral coherence (Richardson et al., 2005; Schmidt & O'Brien, 1997; Varlet et al., 2015), as well as to quantify the determinism of their ancillary movements using detrended fluctuation analysis (DFA; Wang & Yang, 2012). We used DFA alongside cross-spectral coherence to understand the impact of visual cues on postural sway independent of synchrony with the visual cue. Coherence alone would not capture the structure or rigidity of posture, and previous work has shown that movements associated with postural control while standing tend to default to pink noise type fluctuations (Blázquez et al., 2009), but this is altered by rhythmic visual cues (Colley et al., 2018).

Given the mixed prior research, we had two separate hypotheses regarding the effect of avatar movement amplitude on one's ability to synchronize with a musical beat.

1. Based on work on uninstructed coordination, temporal asynchronies relative to a musical pacing signal will be lower when participants observe an avatar with a large movement amplitude, compared to avatars with relatively small, or no movement amplitude.
2. Based on work on instructed coordination, temporal asynchronies will be lower with a moving avatar compared to a still image, but will not change with different movement amplitudes.

Our other measure from the musical drumming task was predictive timing. Based on the finding that temporally relevant biological motion (compared to temporally relevant non-biological motion) facilitates predictive timing (Colley et al., 2018), we structured our hypothesis in a similar manner to the previous hypothesis.

1. Predictive timing will be higher when participants observe an avatar with a large movement amplitude, compared to avatars with relatively small, or no movement amplitude.
2. Predictive timing will be higher with a moving avatar compared to a still image, but will not change with different movement amplitudes.

Regarding our motion capture measures (cross-spectral coherence and DFA), we also had two possible hypotheses.

1. Coherence (between the participant and avatar) and αDFA will be higher when participants observe avatars with larger movement amplitudes, compared to relatively small movement amplitudes, or no movement.
2. Coherence and DFA will be higher with avatars featuring any movement compared to a still image, but will be the same across movement amplitudes.

To test these hypotheses, we used the method of conditional equivalence testing (Campbell & Gustafson, 2018).

Methods

Participants

Participants ($N = 30$, 23 male, $M_{age} = 19$) were recruited through Western Sydney University's research participation programme, and given course credit for completing the experiment. Participants were accepted regardless of musical experience, as we were interested in synchronization abilities in the general population. However, we assessed musical training with a questionnaire. Three participants had more than 5 years of musical training, and were currently involved in instrumental music performance. Of the remaining 27 participants, 12 people reported having 1 academic year or less of music education, and 15 people reported having no formal music education.

In a previous study, Varlet et al. (2012) report an effect size of partial $\eta^2 = .29$ for the effect of stimulus amplitude on unintended visuomotor synchronization. Based on a post-hoc power analysis, our study with 30 participants had greater than .95 power to detect an effect of movement amplitude on synchrony, if the effect generalizes across dependent measures and to other synchronization tasks.

Design

The main experimental design was repeated measures, with one-factor, which we will call *visual cue* (see Figure 1). The factor *visual cue* refers to magnitude of movements in the visual stimulus, and had four levels: normal movement, movement amplified by 100%, movement amplified by 200%, and no movement (control). As a shorthand, the four conditions will be referred to as Regular, Amp1, Amp2, and Still respectively.

Auditory Stimuli

The music with which participants drummed was made for a previous experiment (Colley et al., 2018) and is described in greater detail in the associated article. The duration of each piece was 2 min (and therefore the trial duration was also 2 min). It was composed using MIDI instruments (xylophone, glockenspiel, harp) with short sound envelopes (150–250 ms) so that notes in the melody would not overlap, thereby avoiding ambiguous beat onsets. The musical texture was homophonic and the harmonies were common in Western voice leading. There was no change in rhythm in any of the three instrument parts, so that the lines of music created a single target pulse stream. The average IOI was 500 ms, but there were tempo changes throughout the music (IOI range: 332–668 ms) in order to assess participants' anticipatory timing abilities. The range of these tempo changes was in the order of those observed in expressive musical performance (e.g., Repp, 1992, 1998). There were six rates of change for the tempo changes: +/- 10, +/- 16, and +/-22 ms per beat. There were three pieces of music. All three were similar in style but featured the tempo changes at different times in the music. It should be noted that the tempo changes were randomly generated for each of the three pieces when the stimuli were made but were not randomly generated at each experimental session. In other words, all participants heard the same music. Further details about the music structure, timing, and composition can be found in a previous study (Colley et al., 2018).

Visual Stimuli

The avatar used in the visual stimuli was made by averaging the motion capture recordings of 10 high-performing participants (i.e., relatively good synchronizers) from a previous experiment (Colley et al., 2018), in which they drummed to the same music used here. Thus there were
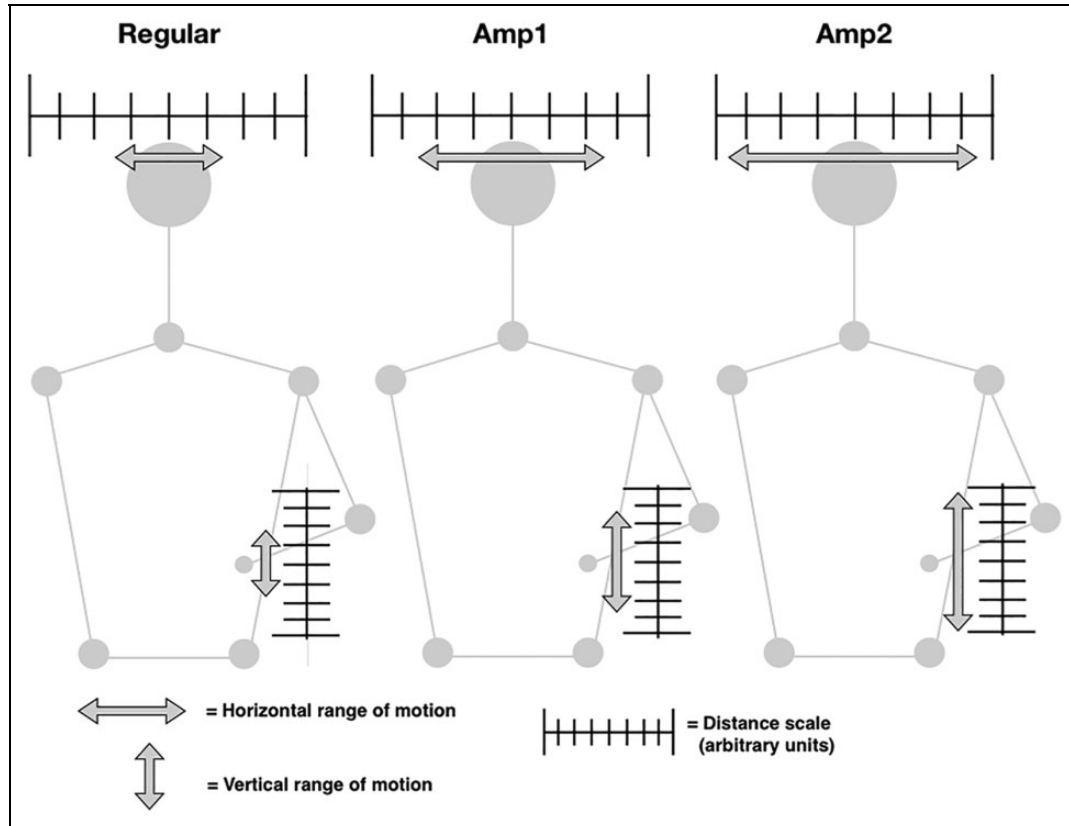
Figure 1. Three images of the three moving visual cues that participants observed during the task. The *Regular* condition (left) was the averaged motion profile of natural movements. *Amp1* (center) increased the range of motion of the *Regular* condition by 100% along the horizontal and vertical planes. *Amp2* (right) increased the range of motion of the *Regular* condition by 200%. The *Still* control condition maintained the image of the avatar shown in this figure for the entire trial (without the scales and arrows). Note that depth of movement was represented by the changing diameters of the circles, but there was very little movement along this axis.

three versions of the avatar, one for each of the three pieces of music. In order for a former participant's data to be included in an avatar, a participant had to be right-handed, have no missed beats, and an average absolute asynchrony below 30 ms for all three pieces of music. With 10 of these participants identified, we reduced the data in their recordings by selecting a subset of motion capture markers that gave the impression of a human body. We removed the left arms from the motion capture recordings used in the visual stimuli, as the former participants tended to exhibit task-irrelevant movements with the left hand (e.g., scratching their head, or resetting a loose marker). Each of these motion capture datasets contained xyz coordinates (represented as distance from an origin point) of the aforementioned markers for each frame of the recording. In brief, these coordinate values were averaged across the 10 model participants. Further details about the averaging procedure used to create the avatars can be found in similar work (Colley et al., 2018).

Once the base avatar was made, we manipulated its movement trajectory to create the other visual cue conditions. The Amp1 condition was made by expanding the range of motion of all markers along all spatial axes (x, y, z) by 100%. In other words, the position coordinates of the base avatar were linearly mapped to fit in between new minimum and maximum values. Thus the timing and relative shape of the avatars stayed the same, but the range of motion increased. The same was done for the Amp2 condition, but the range was increased by 200%. The Still condition (control) was an image of the avatar in its first frame of the animation.

Apparatus

An Alesis Percpad (tapping pad) was used to collect the drumming data in MIDI format. Participants' movements were recorded with a 12-camera Vicon motion capture system at 100 Hz sampling rate, with reflective markers arranged using a custom model with four markers on the head and one marker on each of the following locations: central on the back of the neck, left shoulder, right shoulder, right shoulder blade, right elbow, right inner wrist, and right outer wrist (all participants were right-handed). The motion capture recording and the drum recording were synced by sending a serial trigger signal to Nexus (the motion capture software) at the onset of each

trial. The experimental procedure (avatar animations, stimuli presentation, trigger signals, and data collection) was programmed using C++ in the Xcode coding environment on a 2015 MacBook Pro. Auditory stimuli were sent through stereo speakers, and visual stimuli were presented on a 17" monitor with a 60 Hz refresh rate.

Procedure

Participants received a study information and consent form by email after signing up for the experiment. They were given a paper copy to sign when they arrived for the experiment. Next, with permission from the participant, the experimenter attached motion capture markers to the body parts listed in the *Apparatus* section. While attaching the markers, the experimenter explained the task and answered questions.

Participants were instructed to "drum along to the beat of the music," to "be aware that the speed of the music would sometimes change," and to "always be watching the visuals on the monitor." In an attempt to ensure participants watched the visual cues, we used catch letters, wherein a letter would appear at the center of the screen at pseudo-random timepoints during a trial. Participants were told to say these letters out loud so the experimenter could verify that they were observing the screen and reporting the correct letters. Letter appearances were timestamped to assess whether they had any influence on drumming asynchrony (see *Data Analysis* section). No specific instructions regarding movement were given. Instead, participants were told to stand however they felt comfortable throughout the trial, so long as their feet and eyes were facing the monitor. There were 24 trials of duration 2 min. Participants had one 30 s-long practice trial with no visuals, which they could repeat upon request. There was no electronically generated auditory feedback from the drum pad, though participants could hear and feel the drum stick hitting the drum pad. After the experiment, participants were given a short musical background questionnaire to assess their musical training (if any) and music-listening habits.

Data Analysis

Drumming Analysis

To check for unusual influence by the catch letters on asynchronies we used the seasonal hybrid extreme studentized deviant (SH-ESD) test on the asynchrony time series. SH-ESD detects outliers in seasonal time series data, "seasonal" meaning the time series has periods of fixed length, as in our tempo-changing music. SH-ESD is similar to Grubbs' test for outliers, but is preferred for time series. To check if the catch letters successfully sustained participant attention, the experimenter confirmed that the letter said by the participant matched what appeared on the screen throughout the experiment session.

From our drumming recordings we produced two measurements: asynchrony and predictive timing. Asynchrony was calculated as the average of absolute time differences in ms between the cumulated sequence of musical beat intervals (or inter-onset intervals [IOIs]) and the cumulated sequence of participant drum intervals (or inter-tap intervals [ITIs]). To quantify predictive timing we used the prediction/tracking index (Colley et al., 2017; Pecenka & Keller, 2009). This measure is the ratio of a prediction coefficient over a tracking coefficient. The prediction coefficient represents the strength of the statistical relationship between the ITI and IOI series. The tracking coefficient is the statistical relationship between the ITI series and the lag-1 IOI series. Thus the prediction coefficient is high if participants are anticipating the changing beat intervals and thereby closely matching the intervals, and the tracking coefficient is high if participants are responding to changing beat intervals one beat later, thereby resembling the lagged IOI series. For asynchrony and P/T Index, we used Grubbs' test to identify outliers.

Motion Capture Analysis

All reported analyses of motion capture recordings considered the head movements of participants, as it was found in previous research that this was the part of the body that moved the most (besides the arm, which is considered an instrumental movement, and our hypotheses concern ancillary movements of the body) in these experimental procedures. Also, the head is the most visible part of the body in most musical ensemble contexts, and therefore would presumably serve as the most salient cue with which an individual might synchronize their own ancillary movements. The validity of this assumption is lent support by the fact that recent work on interpersonal coordination in ensembles has focused on head movements (Bishop et al., 2019; Chang et al., 2019).

From our motion capture recordings we produced two measures: cross-spectral coherence and DFA. For both measures, we used the root-sum-square of the raw motion capture data. This produces a directionless signal that incorporates features from all three spatial planes (x, y, z), and we had no specific hypotheses regarding the direction of participant movements. We reduced the motion capture data further by down-sampling to 50 Hz from 100 Hz, and filtering the resulting signal with a 10 Hz low-pass filter.

Cross-spectral coherence measures the consistency of phase relationships among multiple frequencies in a signal. It produces a value between zero (no synchrony) and one (perfect synchrony at all measured frequencies). In this case, we are measuring the phase relationships among different frequencies of movement between participants, and the avatar. As there was no movement in the control

stimulus (a still image), we used a pseudo-pair control. This means that to analyze control trials, we compared the movement of a participant with the movement of the same participant from a different trial featuring the same music. The coherence window size was set at 512, and the overlap size at 50%. The range of measured frequencies was .1 Hz to 8 Hz, and the reported coherence scores are the average of all coherence values from within this range.

DFA quantifies the noise color of a signal. Briefly, signals can exhibit white noise (random values within a narrow range), pink noise (some degree of predictable patterns; some drift), or Brownian noise (highly predictable pattern; heavy drift). Body sway during passive standing tends to exhibit pink noise (Wang & Yang, 2012). If participants entrain to a rhythmic stimulus, we expect DFA to show values above pink noise, as ancillary body movements become more rhythmic and predictable. The output from DFA is α, which typically ranges from 0.5 (white noise) to 1.5 (Brownian noise) with 1.0 (pink noise) in between. For both coherence and DFA we again used Grubbs' test to identify outliers.

Equivalence Test

We used conditional equivalence testing (Campbell & Gustafson, 2018) to address our divergent hypotheses. In traditional hypothesis testing, non-significant test statistics indicate that one should not reject the null hypothesis that two means are equal, but this does not speak to the equivalence of the two or more conditions being compared. In other words, one cannot accept the null hypothesis that two or more means are equal. With conditional equivalence testing, one first uses a standard hypothesis test (in our case, ANOVA). If there are null results in a comparison of two means of interest, and if it is relevant to the hypothesis, one then uses an equivalence test to determine whether the means are statistically equal, or if their relationship is inconclusive with the given data.

The equivalence test we used was the two one-sided test (TOST) method (Lakens et al., 2018). This involves three basic steps.

1. Setting equivalence bounds [-EQ_{low}, EQ_{high}]. The equivalence bounds form the range of difference scores that are not significant and therefore the comparisons are considered equal. The bounds are set to include effect sizes that are considered theoretically equal. If this range is not known or there is no theoretical reason to set a particular set of equivalence bounds, then one uses the smallest detectable effect size given the current data distribution and sample size to set the bounds.
2. Testing whether the difference score of interest falls within the equivalence bounds. This is done by running two one-sided *t*-tests (also called one-tailed tests), with $H_0 1$ that the mean group difference

between conditions is greater than EQ_{high}, and $H_0 2$ that the mean group difference is less than -EQ_{low}. Another way to think of this is as a 90% confidence interval of the estimate of interest (difference scores in this case) that is generated by the two *t*-tests.
3. If both *t*-tests (i.e., the 90% confidence interval of difference score estimates) fall within the equivalence bounds as indicated by two significant *p*-values, then we reject the null hypotheses that the difference score is either greater than the high equivalence bound, or less than the low equivalence bound, and declare equivalence. If one *t*-test is non-significant, the confidence interval will exceed the equivalence bounds, and we declare inconclusive results. If both one-sided *t*-tests are non-significant, then the original ANOVA comparison was significant (this is just a conceptual example; an equivalence test would be unnecessary in this case since the ANOVA was significant).

To set our equivalence bounds we used the data-driven *smallest detectable effect size* method, as we had no theoretical reason to identify a priori non-significant effect sizes for our measures. We considered basing our equivalence bounds for asynchrony on a just noticeable difference (JND) for asynchronous beats, but studies on this topic have had mixed results (Drake & Botte, 1993; Halpern & Darwin, 1982), and a JND for asynchrony would depend on IOI size (Friberg & Sundberg, 1995; Lerens et al., 2014), which is not constant in our stimuli. An asynchrony JND would likely also depend on the acoustical features of a sound (London et al., 2019) and of the room. As such, the *smallest detectable effect size* method of setting equivalence bounds seemed appropriate. We corrected for multiple comparisons using Bonferroni correction. Figure 3 shows the results of the equivalence tests, with the larger of the two *p*-values shown for each test.

In addition to the frequentist statistics, Bayes factors were also used to quantify the evidence in favor of the alternative hypothesis over the null hypothesis (BF_{10}). They are reported alongside *p*-values and are consistent with the results of both the ANOVAs and equivalence tests.

Results

Asynchrony

We first checked whether participants succeeded in the catch-letter task. All participants correctly named all letters, so we believe the task was effective. We then tested for outliers in participants' asynchrony series due to the catch letters. The SH-ESD test showed, on average, 2.6 outlying asynchrony scores for each participant. This is far fewer than the number of letters that appeared in a trial, and only 5 of 78 total outliers across all participants occurred
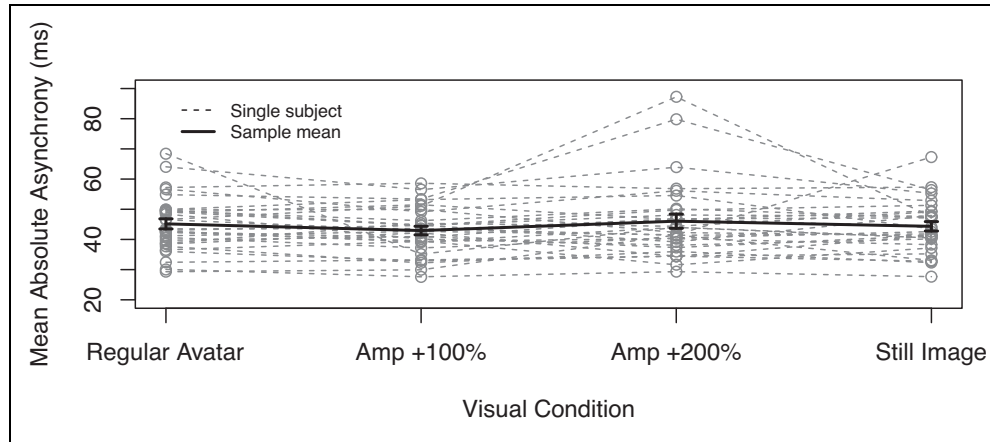
Figure 2. The untransformed mean asynchrony scores expressed in ms. Note that the statistical tests used the log10 transformed data but the untransformed distributions are shown here. Error bars represent standard error of the mean.

within 500 ms after a letter appearing. As such, we have little reason to believe the letters influenced asynchronies.

Prior to the asynchrony ANOVA, we used a log10 transform as the average asynchrony scores were positively skewed in the Regular and Amp2 conditions. No participants were outliers. The ANOVA showed no statistically significant differences between any of the four visual cue condtions (Regular, Amp1, Amp2, and Still), $F(3, 87) = 1.25$, $p = .30$, $\eta^2 = .01$, $BF_{10} < 1$ for all comparisons (see **Figure 2**). Therefore, we used a series of equivalence tests to determine if the different condition comparisons were statistically equal, or inconclusive given the current data. This is best summarized visually in **Figure 3**, top row, which shows the 90% confidence intervals that correspond to each TOST comparison. Intervals within the equivalence bounds are statistically equal. We see that asynchrony was statistically equivalent when comparing the following conditions: Regular to Amp2, Regular to Still, and Amp2 to Still. While only marginally non-significant, the remaining comparisons are considered inconclusive, meaning we cannot conclude a statistical difference or equivalence with the current dataset. The results for individual comparisons were: Reg-Amp1 ($t_{low}(29) = -.003$, $t_{high}(29) = .04$, $p = .04$); Reg-Amp2 ($t_{low}(29) = -.04$, $t_{high}(29) = .03$, $p = .001$); Reg-Still ($t_{low}(29) = -.02$, $t_{high}(29) = .03$, $p = .002$); Amp1-Amp2 ($t_{low}(29) = -.05$, $t_{high}(29) = .003$, $p = .04$); Amp1-Still ($t_{low}(29) = -.03$, $t_{high}(29) = .008$, $p = .01$); Amp2-Still ($t_{low}(29) = -.02$, $t_{high}(29) = .04$, $p = .003$).

P/T Index

The P/T distributions were positively skewed for all conditions so we used a log10 transform on the data. Three participants were removed as outliers after the transform. The ANOVA showed no statistically significant differences between any of the four visual cue condtions (Regular, Amp1, Amp2, and Still), $F(3, 78) = 1.90$, $p = .14$, $\eta^2 = .02$, $BF_{10} < 1$ for all comparisons (see **Figure 4**). The equivalence tests (Figure 3, second row) showed

equivalence for all comparisons: Reg-Amp1 ($t_{low}(26) = -.18$, $t_{high}(26) = .13$, $p < .001$); Reg-Amp2 ($t_{low}(26) = -.22$, $t_{high}(26) = .08$, $p = .02$); Reg-Still ($t_{low}(26) = -.23$, $t_{high}(26) = .01$, $p = .02$); Amp1-Amp2 ($t_{low}(26) = -.23$, $t_{high}(26) = .07$, $p = .04$); Amp1-Still ($t_{low}(26) = -.24$, $t_{high}(26) = .07$, $p = .01$); Amp2-Still ($t_{low}(26) = -.13$, $t_{high}(26) = .12$, $p < .001$).

DFA

DFA distributions were all normal. No participants were identified as outliers. DFA values were generally slightly above 1.0 (**Figure 5**), and within the range observed in previous work on ancillary motion (Colley et al., 2018). The ANOVA did not yield significant effects between any of the four visual cue conditions (Regular, Amp1, Amp2, and Still), $F(3, 87) = 1.90$, $p = .14$, $\eta^2 = .004$, $BF_{10} < 1$ for all comparisons. The equivalence tests (Figure 3) showed the following statistical equivalences: Reg-Amp1 ($t_{low}(29) = -.03$, $t_{high}(29) = .007$, $p = .01$); Reg-Amp2 ($t_{low}(29) = -.03$, $t_{high}(29) = .01$, $p < .001$); Reg-Still ($t_{low}(29) = -.009$, $t_{high}(29) = .03$, $p = .01$); Amp1-Amp2 ($t_{low}(29) = -.02$, $t_{high}(29) = .02$, $p < .001$); Amp2-Still ($t_{low}(29) = -.006$, $t_{high}(29) = .04$, $p = .02$). There was one inconclusive comparison, Amp1-Still ($t_{low}(29) = .002$, $t_{high}(29) = .04$, $p = .09$).

Coherence

The distributions for cross-spectral coherence were normal, and there were no outliers. Coherence values were generally between 0.5 and 0.6 (**Figure 6**), which is in line with previous work (Colley et al., 2020). The ANOVA was significant, $F(3, 87) = 531$, $p < .001$, $\eta^2 = .77$. A Bonferroni-corrected post-hoc test showed that the pseudo-pair control condition showed lower coherence than all other conditions. There were no other statistical differences. The difference between the pseudo-pair and other conditions was reflected in Bayes factor as well,
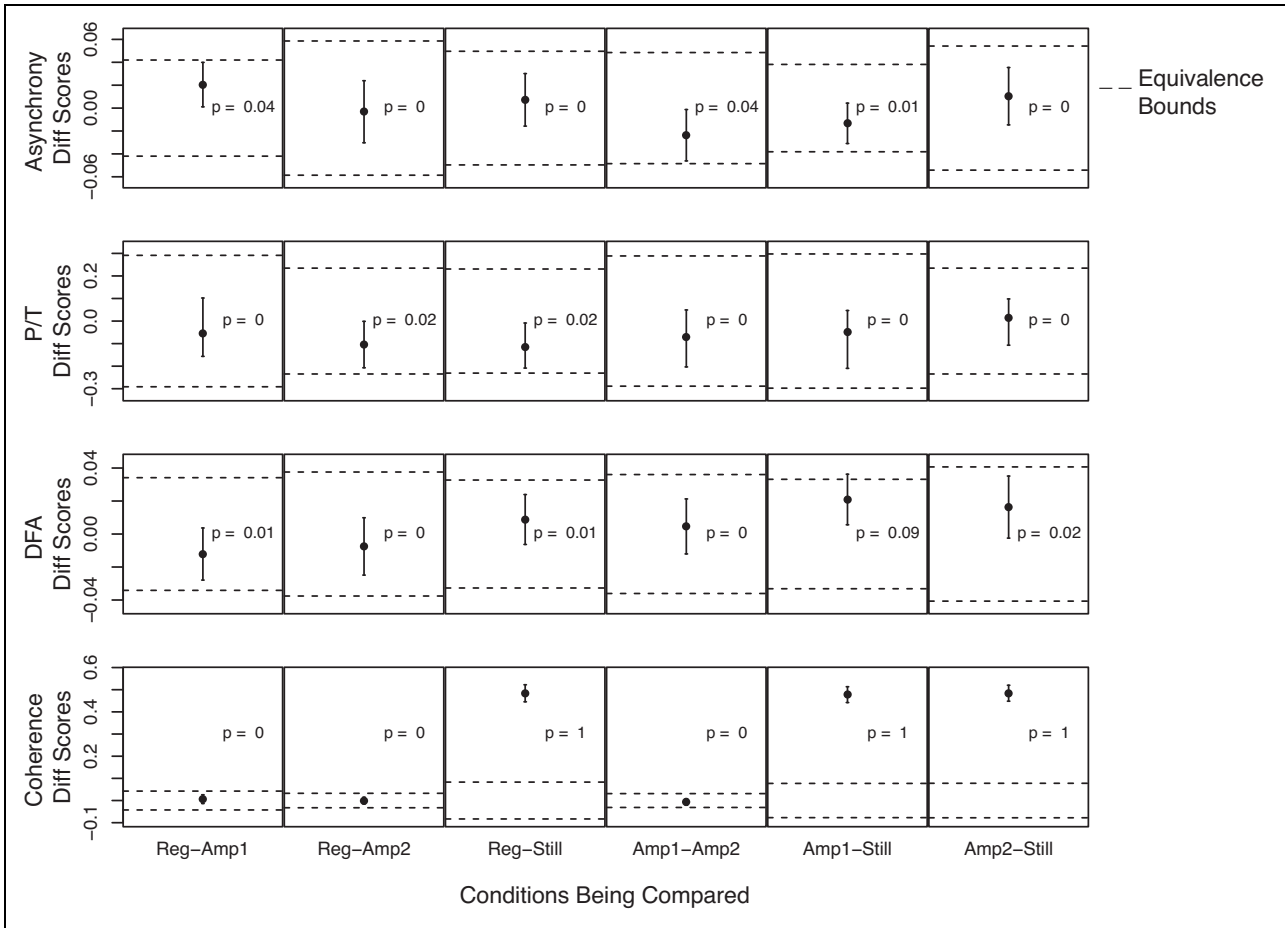
Figure 3. The equivalence bounds and corresponding TOST results to test for statistical equivalence. Each row corresponds to one of our four dependent variables. Each column corresponds to a particular pair-wise comparison of the four conditions. The error bars represent 90% confidence intervals of difference scores.
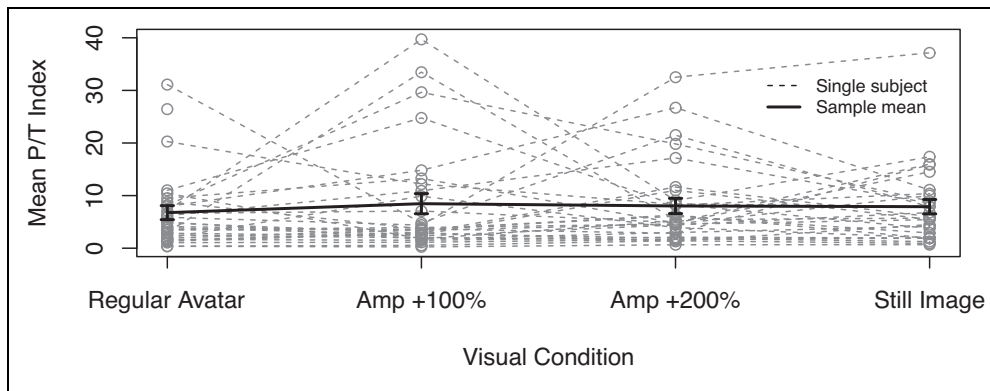

Figure 4. The non-transformed mean P/T Index scores expressed as a ratio of leading/lagging ARMA coefficients (see Methods). Note that the statistical tests used the log10 transformed data, but the natural distributions are shown here. Error bars represent standard error of the mean.

$BF_{10} > 10^{18}$ (all other comparisons had $BF_{10} < 1$). The equivalence test (**Figure 3**) reflected this: there were statistical equivalences for all comparisons of Regular, Amp1, and Amp2: Reg-Amp1 ($t_{low}(29) = $ -.02, $t_{high}(29) = $ .03, $p < .001$); Reg-Amp2 ($t_{low}(29) = $ -.02, $t_{high}(29) = $ .02, $p < .001$); Amp1-Amp2 ($t_{low}(29) = $ -.02, $t_{high}(29) = $ .01, $p < .001$). However, when Regular, Amp1, or Amp2 conditions were compared to the pseudo-pair control, the confidence intervals were well above the equivalence bounds: Reg-Control ($t_{low}(29) = $ .44, $t_{high}(26) = $ .53, $p = .1$);
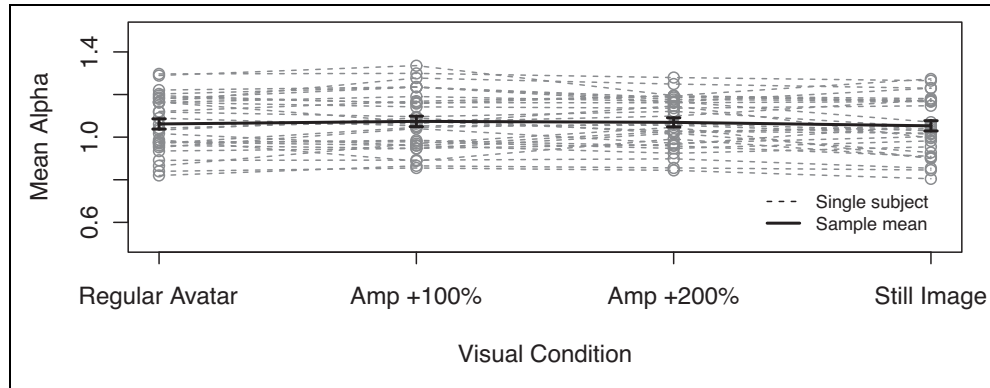
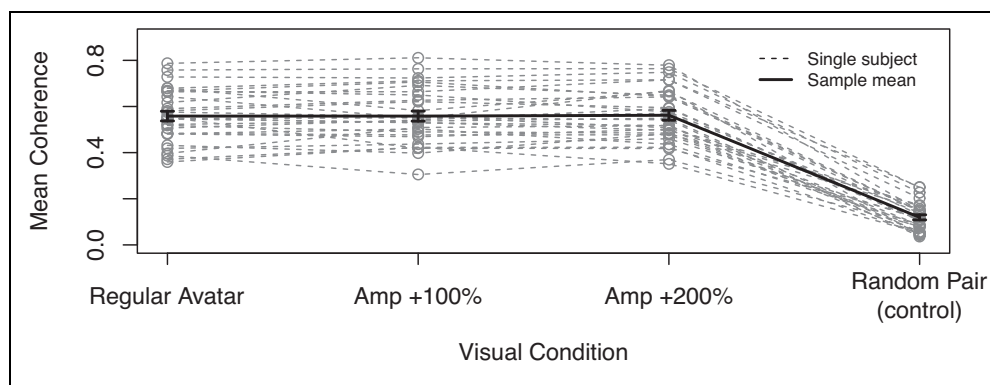**Figure 5.** The mean DFA scores of participants.

Figure 6. The mean cross-spectral coherence scores between participants and the avatar (or a pseudo-pair).

Amp1-Control ($t_{low}(29) = .43$, $t_{high}(29) = .52$, $p = 1$); Amp2-Control ($t_{low}(29) = .44$, $t_{high}(29) = .53$, $p = 1$).

Discussion

This experiment investigated the role of movement amplitude of a visual stimulus in facilitating musical synchronization and influencing ancillary movements. The visual stimulus of which we manipulated the amplitude was a high-performing virtual co-performer (a motion capture avatar). The rationale for this is that a co-performer can be beneficial to a partner if the co-performer is good at the task (Pecenka & Keller, 2011). Additionally, higher amplitudes of movement that are timed to a fixed musical sequence produce higher velocities (by moving more distance in the same time), which have been shown to improve musical synchronization (Colley et al., 2018). Given mixed prior results on movement amplitude and visuomotor synchronization, we advanced two hypotheses: if overall musical synchrony (i.e., instructed and uninstructed movements) is influenced by the amplitude of co-performer movements, then higher amplitudes of stimulus movement will result in lower asynchrony, and higher coherence; alternatively, if musical synchrony is not influenced by the amplitude of a co-performer, then higher amplitudes of stimulus movement will not produce differences in our

dependent measures. We also considered the determinism of ancillary movements (DFA), which is not a measure of synchrony but quantifies the extent to which movements are predictable. If stimulus amplitude influences movements, then we would expect larger amplitudes to produce higher DFA values, as movements linked to the musical structure would be relatively predictable. If stimulus amplitude does not influence movements, then we would expect no difference in DFA values across amplitude conditions.

Overall, our results suggest that there is no reliable effect of movement amplitude of a visual stimulus on synchronization accuracy, predictive timing, ancillary movement fluctuations, or the synchrony of ancillary movements between the participant and the avatar. A number of comparisons between the moving visual stimulus conditions were statistically equivalent, suggesting that our amplitude manipulation produced three effectively identical stimuli (despite physical differences in the visual displays), and so we have greater support for our second set of hypotheses. What is surprising is that the movement conditions were generally no different than the control condition, in which participants observed a still image. The exception to this was the cross-spectral coherence measure, which showed higher coherence between participants' head movements and the moving avatars' head movements, than between participants' head movements and a copy of their

own movements from another trial (a pseudo-pair). This finding, alongside the apparent success of the catch letters, suggests that participants were not ignoring the visual display. If they were not observing the visual cues, then their ancillary coherence in the experimental trials would likely resemble the coherence from the pseudo-pair control. Note that in the non pseudo-pair conditions, mean coherence was around 0.6, which is a moderately high degree of coordination. This is likely because the stimuli were rhythmic, providing some degree of predictability for corresponding body movements.

First, we will discuss the drumming dependent variables: asynchrony and P/T Index. It seems that the instructed synchronization of our participants was not affected by the moving visual cues, even compared to a still image visual cue. This could be due to participants' generally small amount of training in music, which was reflected in the average absolute asynchrony across conditions (about 45 ms, compared to 25 ms for the highly synchronized individuals used in creating the avatar). This is consistent with another synchronization study that tested nonmusicians with similar tempo-changing stimuli (Mills et al., 2015). For example, motor experts (people with experience executing deliberate movements in a given domain) tend to be more perceptually sensitive to gross body movements in their domain. Basketball players predict shot success better than referees, who typically observe but do not play the game (Aglioti et al., 2008). Similarly, violinists predict tone onsets better than musicians of other instruments when observing video of a violinist performing a cueing motion, a movement meant to help observers predict a tone onset (Wöllner & Canal-Bruland, 2010). More recent work has shown that gestures can effectively convey a beat and tempo in musical duos, but only expert musicians were tested, and musicians with more ensemble experience synchronized better (Bishop & Goebl, 2018a). In another study, musicians were generally able to perceive audiovisual asynchronies in musical performance videos, but pianists showed more perceptual sensitivity when observing other pianists (Bishop & Goebl, 2018b). Given the results of these studies, musical expertise may be beneficial for integrating temporal information from a moving body. Furthermore, musicians in one study only looked at the conductor 28% of the time during the performance of a piece of music, and each glance was less than 1 s in duration (Fredrickson, 1994), suggesting they have trained the ability to receive temporal information from brief glances. Only three of our participants had extensive musical training, and only two had ensemble training, meaning the sample was mostly nonmusicians. The three musicians' asynchrony scores were in the lowest four values of the sample, so they were performing well relative to the remaining sample. However, they did not qualify as outliers so we have no reason to treat them as a separate group. Furthermore, removing the three musicians from the sample (resulting in $N = 27$) did not change the significance of the results of the hypothesis tests. As such, the participants may have observed the stimuli as instructed, but may not have been able to extract relevant temporal information from a full upper-body display, which had multiple moving parts. In other words, participants did not have experience watching a complex rhythmic stimulus to form a temporal prediction.

Expanding on this, a previous study showed that a video of a conductor (from the waist up, similar to our avatars) yielded more precise tapping than a video of a metronome for musicians, but not nonmusicians. In the same study, neural activation in the superior frontal gyrus correlated positively with the amount of time spent practicing with a conductor (Ono et al., 2015). Both groups performed the same in the metronome condition, perhaps because the metronome had a single moving part that corresponds directly to the beat. A previous study (Colley et al., 2018) showed that both musicians and nonmusicians benefitted from a virtual conductor, which was presented as a single moving circle. This suggests that visual cues for instructed synchronization are most effective for the general population if they are kept simple (i.e., one moving part). Complex whole-body movements likely require training to analyze in real time. Indeed, it has been shown that body movements exhibit multiple periodicities when dancing (Burger et al., 2014; Su, 2016), and that tracking multiple moving objects simultaneously complicates action prediction (Atmaca et al., 2013). Thus, segments of the body that move in relation to a musical beat might be perceived as individually moving parts rather than as a whole phase-locked system, which in turn might depreciate the value of a visual cue. Future studies might test this explicitly by manipulating the number of visible limbs/moving parts in an avatar, and comparing synchronization performance between ensemble musicians, solo musicians, and nonmusicians.

Our motion capture results reinforced one common finding: individuals tend to entrain their movements to a visual rhythm (Clayton, 2007; Kotz et al., 2014; Schmidt et al., 2007; Schmidt & Turvey, 1994; Varlet et al., 2015). But this uninstructed visuomotor entrainment of the head does not appear to be increased by the amplitude of the visual stimulus, at least in a multisensory context such as music performance. But again, the effect of stimulus amplitude on synchrony may be a matter of expertise, such that experienced ensemble musicians would be more likely to show greater ancillary movement coherence with the amplified avatars, particularly if auditory stimuli were removed (Goebl & Palmer, 2009). Alternatively, an effect of stimulus amplitude on synchrony in a musical task might be more prominent among pairs of live co-performers, without any virtual avatar, as suggested by relevant findings in dyadic synchronization tasks (Colley et al., 2020; Goebl & Palmer, 2009; Keller & Appel, 2010). As for the fluctuations of movements as measured by DFA, there was no difference across conditions. Importantly, participants'

DFA scores for all conditions were centered just above 1.0, suggesting that people tended to move with little more structure than passive standing balance (Blázquez et al., 2009). We expected the amplitude manipulation to increase DFA scores, indicating more rhythmically structured movements of the participants. If our participants were in fact unable to extract temporal information from the avatars, then they may have neglected the visual information entirely as it was deemed unreliable (Elliott et al., 2010).

In general, it is possible that our participants exhibited ceiling effects in their behavior. To address this in future work, task difficulty could be increased by introducing large-scale discontinuous tempo changes and pauses into the musical pacing signal. Studies of musical duo performance have shown that the benefits of visual cues are enhanced in the presence of such features (e.g., Bishop et al., 2019; Kawase, 2014). Larger amplitude movements of an avatar may therefore be beneficial when auditory cues are characterized by greater temporal uncertainty than was the case in our study.

On the topic of movement amplitudes, the amplitude manipulation in this experiment was not natural, such that we edited the visual recordings to exaggerate the movements. As humans are especially sensitive to biological/natural movements (e.g., Ueda et al., 2018), the unnatural manipulation may have reduced processing efficiency or even caused the motion cues not to be processed as behaviorally relevant signals. Previous research has found that perceptual judgments of performer identity in dancing avatars are influenced by variations in movement amplitude induced by asking the models to dance expressively versus unexpressively (Sevdalis & Keller, 2011). Future studies on sensorimotor synchronization could take a similar approach by directly recording model participants who drum at different amplitudes instead of using artificial modulations.

Finally, it should be noted that our sample came from a healthy population. However, an individual's ability to control periodic movements can be impaired if afflicted with a motor disorder such as Parkinson's Disease (Hove et al., 2012; Nombela et al., 2013). Research on rehabilitation in Parkinson's Disease has shown that external rhythmic cues—both auditory and visual—can restore some functionality to patients (Ghai et al., 2018; Hove & Keller, 2015). The moving visual stimuli presented in this experiment might provide some benefit to patients with movement disorders where healthy participants received no advantage relative to the control stimulus.

To conclude, our finding that co-performer movement amplitude did not have reliable effects on instructed or uninstructed synchronization suggests that this specific visual cue might not be functionally relevant to basic aspects of interpersonal timing in musical contexts, at least in samples of individuals with little musical training. We draw this conclusion based not only on statistically non-significant differences, but on several statistically equivalent comparisons as well. Future studies of visuomotor and audio-visuomotor synchronization should consider the influence of expertise, especially in musical synchronization. Other possible variables of interest are the complexity or richness of the musical material (e.g., the potential for expressive variation) and stimulus movement (as measured by the number of moving parts or distinct movement frequencies). Musical expertise and complexity may be influential to the extent that ancillary movements play a greater role in providing cues for flexibly aligning expressive performance parameters than in facilitating strictly synchronized timing (Keller, 2014).

Contributorship

IC conceived the study, prepared the stimuli, ran the experiment, analyzed the data, and drafted the manuscript. PK, MV, and JM were involved in the conception of the study. They also advised on data analysis and interpretation, and reviewed and provided feedback on each version of the manuscript.

Declaration of conflicting interests

Funding

Action editor

Andrew Goldman, Indiana University, Department of Music Theory, Jacobs School of Music.

Peer review

Birgitta Burger, Universität Hamburg, Institut für Systematische Musikwissenschaft.
Two anonymous reviewers.

ORCID iDs

Ian D. Colley https://orcid.org/0000-0003-0735-1746
Jennifer MacRitchie https://orcid.org/0000-0003-4183-6552

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116. https://doi.org/10.1038/nn.2182

Atmaca, S., Stadler, W., Keitel, A., Ott, D. V., Lepsien, J., & Prinz, W. (2013). Prediction processes during multiple object tracking (MOT): Involvement of dorsal and ventral premotor cortices. *Brain and Behavior*, *3*(6), 683–700. https://doi.org/10.1002/brb3.180

Bishop, L., Cancino-Chacón, C., & Goebl, W. (2019). Moving to communicate, moving to interact. *Patterns of Body Motion in Musical Duo Performance*, *37*(1), 1–25. https://doi.org/10.1525/mp.2019.37.1.1

Bishop, L., & Goebl, W. (2018a). Beating time: How ensemble musicians' cueing gestures communicate beat position and tempo. *Psychology of Music*, *46*(1), 84–106.

Bishop, L., & Goebl, W. (2018b). Communication for coordination: Gesture kinematics and conventionality affect synchronization success in piano duos. *Psychological Research*, *82*, 1177–1194.

Blázquez, M. T., Anguiano, M., Saavedra, F. A., Lallena, A. M., & Carpena, P. (2009). Study of human postural control system during quiet standing using detrended fluctuation analysis. *Physica A*, *388*, 1857–1866. https://doi.org/10.1016/j.physa.2009.01.001

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S. H., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement. *Frontiers in Human Neuroscience*, *8*. https://doi.org/ARTN903 10.3389/fnhum.2014.00903

Campbell, H., & Gustafson, P. (2018). Conditional equivalence testing: An alternative remedy for publication bias. *PLoS One*, *13*(4). https://doi.org/e0195145

Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2019). Body sway reflects joint emotional expression in music ensemble performance. *Scientific Reports*, *9*(1), 205. https://doi.org/10.1038/s41598-018-36358-4

Chen, Y. Q., Repp, B. H., & Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: Comparisons between auditory and visual pacing and feedback conditions. *Human Movement Science*, *21*(4), 515–532. https://doi.org/Pii S0167-9457(02)00138-0 https://doi.org/10.1016/S0167-9457(02)00138-0

Clayton, M. R. L. (2007). Observing entrainment in music performance: Video-based observational analysis of Indian musicians' tanpura playing and beat marking. *Musicae Scientiae*, *11*(1), 27–59. https://doi.org/Doi10.1177/102986490701100102

Colley, I. D., Keller, P. E., & Halpern, A. R. (2017). Working memory and auditory imagery predict sensorimotor synchronization with expressively timed music. *Quarterly Journal of Experimental Psychology*, *71*(8), 1781–1796. https://doi.org/10.1080/17470218.2017.1366531

Colley, I. D., Varlet, M., MacRitchie, J., & Keller, P. E. (2018). The influence of visual cues on temporal anticipation and movement synchronization with musical sequences. *Acta Psychologica*, *191*, 190–200.

Colley, I. D., Varlet, M., MacRitchie, J., & Keller, P. E. (2020). The influence of a conductor and co-performer on auditory-motor synchronisation, temporal prediction and ancillary entrainment in a musical drumming task. *Human Movement Science*, *72*, 102653.

Davidson, J., & Broughton, M. C. (2016). Bodily mediated coordination, collaboration, and communication in music performance. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford handbook of music psychology* (pp. 573–595). Oxford University Press.

de Rugy, A., Oullier, O., & Temprado, J. J. (2008). Stability of rhythmic visuo-motor tracking does not depend on relative velocity. *Experimental Brain Research*, *184*, 269–273.

Dijkstra, T. M., Schöner, G., & Gielen, C. C. (1994). Temporal stability of the action-perception cycle for postural control in a moving visual environment. *Experimental Brain Research*, *97*, 477–486.

Drake, C., & Botte, M.-C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics*, *54*, 277–286.

Elliott, M. T., Wing, A. M., & Welchman, A. E. (2010). Multisensory cues improve sensorimotor synchronisation. *European Journal of Neuroscience*, *31*(10), 1828–1835. https://doi.org/10.1111/j.1460-9568.2010.07205.x

Fredrickson, W. E. (1994). Band musicians' performance and eye contact as influenced by loss of a visual and/or aural stimulus. *Journal of Research in Music Education*, *4*, 306–317.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *Journal of the Acoustical Society of America*, *98*(5), 2524–2531.

Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A. O. (2018). Effect of rhythmic auditory cueing on Parkinsonian gait: A systematic review and meta-analysis. *Scientific Reports*, *8*(1), 506. https://doi.org/10.1038/s41598-017-16232-5

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, *26*, 427–438.

Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics*, *31*(1), 86–89.

Hove, M. J., & Keller, P. E. (2010). Spatiotemporal relations and movement trajectories in visuomotor synchronization. *Music Perception*, *28*(1), 15–26. https://doi.org/10.1525/mp.2010.28.1.15

Hove, M. J., & Keller, P. E. (2015). Impaired movement timing in neurological disorders: Rehabilitation and treatment strategies. *Neurosciences and Music V: Cognitive Stimulation and Rehabilitation*, *1337*, 111–117. https://doi.org/10.1111/nyas.12615

Hove, M. J., Suzuki, K., Uchitomi, H., Orimo, S., & Miyake, Y. (2012). Interactive rhythmic auditory stimulation reinstates natural 1/f timing in gait of Parkinson's patients. *PLoS One*, *7*(3), 1–8. https://doi.org/10.1371/journal.pone.0032600

Kawase, S. (2014). Gazing behavior and coordination during piano duo performance. *Attention, Perception, & Psychophysics*, *76*(2), 527-540. https://doi.org/10.3758/s13414-013-0568-0

Keller, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression. In D. Fabian, R. Timmers, & E. Schubert (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures* (Vol. *1*., pp. 260–282). Oxford University Press.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*, 27–46.

Kotz, S. A., Penhune, V., Henry, M., Large, E., Grahn, J. A., & Dalla Bella, S. (2014). Entrainment: A domain general cognitive timing mechanism? *International Conference on Timing*

and *Time Perception*, *126*, 24–26. https://doi.org/10.1016/j.sbspro.2014.02.304

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269.

Lerens, E., Araneda, R., Renier, L., & De Volder, A. G. (2014). Improved beat asynchrony detection in early blind individuals. *Perception*, *43*, 1083–1096. https://doi.org/10.1068/p7789

London, J., Nymoen, K., Langerød, M. T., Thompson, M. R., Code, D. L., & Danielsen, A. (2019). A comparison of methods for investigating the perceptual center of musical sounds. *Attention, Perception, & Psychophysics*, *81*(6), 2088-2101. https://doi.org/10.3758/s13414-019-01747-y

Luck, G., Saarikallio, S., Burger, B., Thompson, M., & Toiviainen, P. (2014). Emotion-driven encoding of music preference and personality in dance. *Musicae Scientiae*, *18*(3), 307–323. https://doi.org/10.1177/1029864914537290

Luck, G., & Sloboda, J. (2008). Exploring the spatio-temporal properties of simple conducting gestures using a synchronization task. *Music Perception: An Interdisciplinary Journal*, *25*(3), 225–239. https://doi.org/10.1525/mp.2008.25.3.225

Luck, G., & Sloboda, J. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception*, *26*(5), 465–473. https://doi.org/10.1525/mp.2009.26.5.465

Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: An automated feature-extraction analysis. *Music Perception*, *24*, 189–200.

Lumsden, J., Miles, L. K., Richardson, M. J., Smith, C. A., & Macrae, C. N. (2012). Who syncs? Social motives and interpersonal coordination. *Journal of Experimental Social Psychology*, *48*(3), 746–751. https://doi.org/10.1016/j.jesp.2011.12.007

Madison, G., Karampela, O., Ullén, F., & Holm, L. (2013). Effects of practice on variability in an isochronous serial interval production task: Asymptotical levels of tapping variability after training are similar to those of musicians. *Acta Psychologica*, *143*, 119–128.

Manning, F. C., Harris, J., & Schutz, M. (2017). Temporal prediction abilities are mediated by motor effector and rhythmic expertise. *Experimental Brain Research*, *235*, 861–871.

Mills, P. F., van der Steen, M. C., Schultz, B. G., & Keller, P. E. (2015). Individual differences in temporal anticipation and adaptation during sensorimotor synchronization. *Timing & Time Perception*, *3*(1-2), 13–31. https://doi.org/10.1163/22134468-03002040

Nombela, C., Hughes, L. E., Owen, A. M., & Grahn, J. A. (2013). Into the groove: Can rhythm influence Parkinson's disease? *Neuroscience & Biobehavioral Reviews*, *37*, 2564–2570.

Ono, K., Nakamura, A., & Maess, B. (2015). Keeping an eye on the conductor: Neural correlates of visuo-motor synchronization and musical experience. *Front Hum Neurosci*, *9*, 154. https://doi.org/10.3389/fnhum.2015.00154

Pecenka, N., & Keller, P. E. (2009). Auditory pitch imagery and its relationship to musical synchronization. *Annals of the New York Academy of Sciences*, *1169*, 282–286. https://doi.org/10.1111/j.1749-6632.2009.04785.x

Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, *211*(3-4), 505–515. https://doi.org/10.1007/s00221-011-2616-0

Peper, C. E., & Beek, P. J. (1998). Are frequency-induced transitions in rhythmic coordination mediated by a drop in amplitude. *Biological Cybernetics*, *79*, 291–300.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition*, *105*, 533–546.

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*(1), 94–102. https://doi.org/10.1016/j.bandc.2007.11.007

Repp, B. H. (1992). Diversity and commonality in music performance: An analysis of timing microstructure in Schumann's "Traumerei". *Journal of the Acoustical Society of America*, *92*(5), 2546–2568.

Repp, B. H. (1998). A microcosm of musical expression. I. Quantitative analysis of pianists' timing in the initial measures of Chopin's Etude in E major. *The Journal of the Acoustical Society of America*, *104*, 1085–1100. https://doi.org/10.1121/1.423325

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology Human Perception and Performance*, *31*(1), 62–79.

Schmidt, R. C., & O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecological Psychology*, *9*(3), 189–206.

Schmidt, R. C., Richardson, M. J., Arsenault, C., & Galantucci, B. (2007). Visual tracking and entrainment to an environmental rhythm. *Journal of Experimental Psychology Human Perception and Performance*, *33*(4), 860–870. https://doi.org/10.1037/0096-1523.33.4.860

Schmidt, R. C., & Turvey, M. T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biological Cybernetics*, *70*(4), 369–376. https://doi.org/10.1007/Bf00200334

Sevdalis, V., & Keller, P. E. (2011). Perceiving performer identity and intended expression intensity in point-light displays of dance. *Psychological Research*, *75*, 423–434.

Su, Y. H. (2016). Sensorimotor synchronization with different metrical levels of point-light dance movements. *Frontiers in Human Neuroscience*, *10*(186). https://doi.org/10.3389/fnhum.2016.00186

Thompson, M. R., & Luck, G. (2011). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*(1), 19–40. https://doi.org/10.1177/1029864911423457

Ueda, H., Yamamoto, K., & Watanabe, K. (2018). Contribution of global and local biological motion information to speed perception and discrimination. *Journal of Vision*, *18*(3), 2–2. https://doi.org/10.1167/18.3.2

Varlet, M., Bucci, C., Richardson, M. J., & Schmidt, R. C. (2015). Informational constraints on spontaneous visuomotor

entrainment. *Human Movement Science*, *41*, 265–281. https://doi.org/10.1016/j.humov.2015.03.011

Varlet, M., Coey, C. A., Schmidt, R. C., Marin, L., Bardy, B. G., & Richardson, M. J. (2014). Influence of stimulus velocity profile on rhythmic visuomotor coordination. *Journal of Experimental Psychology Human Perception and Performance*, *40*(5), 1849–1860. https://doi.org/10.1037/a0037417

Varlet, M., Coey, C. A., Schmidt, R. C., & Richardson, M. J. (2012). Influence of stimulus amplitude on unintended visuomotor entrainment. *Human Movement Science*, *31*(3), 541–552. https://doi.org/10.1016/j.humov.2011.08.002

Wang, C., & Yang, W. (2012). Using detrended fluctuation analysis (DFA) to analyze whether vibratory insoles enhance balance stability for elderly fallers. *Archives of Gerontology and Geriatrics*, *55*, 673–676. https://doi.org/10.1016/j.archger.2011.11.008

Wöllner, C., & Canal-Bruland, R. (2010). Keeping an eye on the violinist: Motor experts show superior timing consistency in a visual perception task. *Psychological Research*, *74*(6), 579–585. https://doi.org/10.1007/s00426-010-0280-9

Wöllner, C., Deconinck, J. F., Parkinson, J., Hove, M. J., & Keller, P. E. (2012). The perception of prototypical motion: Synchronization is enhanced with quantitatively morphed gestures of musical conductors. *Journal of Experimental Psychology Human Perception and Performance*, *38*, 1390–1403. https://doi.org/10.1037/a0028130