The Effects of Visual Movement on Beat-Based vs. Duration-Based Temporal Perception

Article *in* Timing & Time Perception · April 2019

DOI: 10.1163/22134468-20191141

4 authors, including:

Nathércia Lima Torres
University of Porto
6 PUBLICATIONS **6** CITATIONS

Carlos Dos Santos Luiz
Instituto Politécnico de Coimbra
22 PUBLICATIONS **61** CITATIONS

São Luís Castro
University of Porto
172 PUBLICATIONS **3,594** CITATIONS

BRILL

Timing & Time Perception 7 (2019) 168–187

Timing
&
Time Perception
brill.com/time

The Effects of Visual Movement on Beat-Based vs. Duration-Based Temporal Perception

Nathércia L. Torres[1,*], Carlos dos Santos Luiz[2], São Luís Castro[1] and Susana Silva[1,*]

[1]Center for Psychology at University of Porto (CPUP), Faculty of Psychology and Education Sciences, University of Porto, Portugal
[2]Polytechnic Institute of Coimbra, Portugal

Received 24 November 2018; accepted 16 March 2019

Abstract

It is known that moving visual stimuli (bouncing balls) have an advantage over static visual ones (flashes) in sensorimotor synchronization, such that the former match auditory beeps in driving synchronization while the latter do not. This occurs in beat-based synchronization but not in beat-based purely perceptual tasks, suggesting that the advantage is action-specific. The main goal of this study was to test the advantage of moving over static visual stimuli in a different perceptual timing system – duration-based perception – to determine whether the advantage is action-specific in a broad sense, i.e., if it excludes both beat-based and duration-based perception. We asked a group of participants to perform different tasks with three stimulus types: auditory beeps, visual bouncing balls (moving) and visual flashes (static). First, participants performed a duration-based perception task in which they judged whether intervals were speeding up or slowing down; then they did a synchronization task with isochronous sequences; finally, they performed a beat-based perception task in which they judged whether sequences sounded right or wrong. Bouncing balls outperformed flashes and matched beeps in synchronization. In the duration-based perceptual task, beeps, balls and flashes were equivalent, but in beat-based perception beeps outperformed balls and flashes. Our findings suggest that the advantage of moving over static visual stimuli is grounded on action rather than perception in a broad sense, in that it is absent in both beat-based and duration-based perception.

Keywords
Audition, beat, timing systems, synchronization, vision

---

[*] To whom correspondence should be addressed. E-mail: limatorres.n@gmail.com / susanamsilva@fpce.up.pt

DOI: 10.1163/22134468-20191141

1. Introduction

From microseconds to circadian rhythms, the perception of time is crucial in human life. The development of different mechanisms to quantify time along a wide range of durations is an essential skill in many types of behavior, such as playing music, speaking, and performing a sport (Merchant & Lafuente, 2008). Studies in timing are also relevant to the field of psychiatry and neurology: atypical temporal processing is often seen in patients who suffer from schizophrenia (Carroll et al., 2009; Peterburs et al., 2013; Thoenes & Oberfeld, 2017) and Parkinson's disease (Biswas et al., 2016; Grahn & Brett, 2009). Increased knowledge on healthy time perception is important to devise new assessment tools and rehabilitation programs for these populations.

Previous research has identified two different timing systems, defining how the underlying timing circuitry of the brain might work (McAuley & Jones, 2003; Pashler, 2001; Yee et al., 1994). The first is *duration-based timing*, which refers to the absolute duration of individual time intervals. Under this mechanism, time is encoded like a stopwatch (Teki et al., 2011). The second system concerns relative time or *beat-based timing*. It engages an implicit regular pulse that marks equally-spaced events in time, providing a reference unit for measuring time (Grahn, 2012). The existence of duration-based and beat-based timing systems is supported by neuroimaging studies. Teki et al. (2011) conducted a functional magnetic resonance imaging (fMRI) experiment in which participants were asked to compare the duration of the last interval with that of the penultimate interval, while varying the rhythmic context of preceding intervals to be regular (beat-based) or irregular (duration-based). The authors observed that the olivocerebellar system mediates duration-based timing, and the striato-thalamo-cortical system assists beat-based timing. The existence of beat-based and duration-based mechanisms is also supported by behavioral studies (McAuley & Jones, 2003; Pashler, 2001).

Modality effects are known to exist in beat-based *perception* as well as in beat-based *sensorimotor synchronization* (synchronization hereafter). Synchronization is a motor task in which movements are executed in time with an external stimulus (Pollok et al., 2009; Repp, 2005), thus requiring the explicit representation of time for motor control that is recruited in event-based timing tasks (Ivry et al., 2002; Zelaznik et al., 2002). Modality effects on beat-based perception and synchronization indicate that beat-based performance is influenced by the modality (auditory vs. visual) in which temporal patterns are presented (Grahn, 2012). In the perceptual domain, several studies have reported that auditory stimuli such as beeps are perceived more accurately than visual ones such as flashing images (Glenburg & Jona, 1991; Grahn, 2012; Grahn et al., 2011; Guttman et al., 2005; Patel et al., 2005; Repp & Panel, 2002; Stauffer et al., 2012), even though it seems possible to extract a beat from visual stimuli (Su, 2014). In line with this, the perceptual processing of visual stimuli may rely on auditory recoding (Grahn et al., 2011), while the reverse is not true (visual recoding of auditory stimuli). Beyond the perceptual domain, the investigation of modality differences in synchronization performance has also shown the advantage of audition over vision (Chen et al., 2002; Hove et al., 2013a; Patel et al., 2005; Pollok et al., 2009). In order to account for these results, it has been suggested that audition would hold an advantage over vision in generating a sense of beat.

Recent studies challenged the idea that visual beat-based processing is always poor, and they did so by showing that visual stimuli with apparent motion (e.g., moving bars, bouncing balls) outperform static visual stimuli (flashes) in synchronization (Gan et al., 2015; Hove et al., 2010; Iversen et al., 2015). By increasing the realism of the visual motion trajectory, both Gan et al. (2015; realism increased by rectified sinusoidal velocity related to gravity) and Silva & Castro (2016, realism increased by ball squashing when hitting the ground) went further and demonstrated that synchronization with a bouncing ball is no less stable than synchronization with an auditory metronome. In other words, this investigation showed that synchronization with moving visual stimuli can be as effective as synchronization with auditory ones, but synchronization with static visual stimuli cannot. The idea of *interaction between modality* (visual vs. auditory) and *continuity* (moving vs. static stimuli) *on synchronization* includes these effects. This interaction has been demonstrated by Hove and colleagues (2013b), who showed that moving (bouncing balls, continuous presence on screen) outperform static (flashes, discrete/discontinuous presence) visual stimuli in driving synchronization, while the reverse occurs in the auditory domain — i.e., discontinuous sounds (beeps) outperform continuous ones (sirens). In the present paper, we do not consider the full 2 × 2 (modality × continuity) interaction: we focus on the comparison between static visual (flashes), moving visual (bouncing balls) and discrete auditory stimuli (beeps), which would be enough to test further the hypothesis that bouncing balls outperform flashes and match beeps in temporal processing.

The reason why visual stimuli need movement (bouncing balls vs. static flashes) to compete with auditory ones (beeps) is not well determined. One possibility is that the lower temporal resolution of vision requires increased reliance on spatial (i.e., stimulus movement) information for equivalent timing accuracy (Hove et al., 2013b). Although this can be partly true, stimulus movement does not seem enough to boost timing performance in all circumstances: the advantage of moving over static visual stimuli (flashes) seems to be synchronization-specific, given the lack of evidence of this advantage in the perceptual domain. Critically, Silva & Castro (2016) demonstrated recently that bouncing balls match beeps in synchronization but not in perception, providing direct evidence that the advantage of moving over static visual (beat-based) stimuli is action/synchronization-specific and remains absent in purely perceptual tasks. Therefore, it seems that human action plays a major role in visual timing. This is in line with findings of improved time perception following overt action (Morillon & Baillet, 2017), and with the more general claim that timing is closely related to the motor system (Comstock et al., 2018).

Although the hypothesis that human action may modulate continuity effects on visual timing seems promising (i.e., moving visual stimuli outperform static ones only under explicit human movement), the advantages of moving over static visual stimuli have not been fully explored. For instance, most studies in timing evaluated only the *beat-based* timing system. Therefore, an interesting question remains to be addressed: we know that the advantage of moving over static visual stimuli is present in beat-based synchronization but absent in beat-based perception. The question is: does this advantage extend to duration-based processing? This question has two facets. One, does the production of durations show the advantage of moving over static visual stimuli, just as beat-based synchronization does? Two, does this occur in duration-based perception, as it does in beat-based perception? In the present study, we took a first step and sought to provide an answer for question two.

The main goal of this study was to compare the effects of stimulus continuity on visual temporal perception across beat-based and duration-based timing systems. Our main motivation was to better understand the advantage of moving over static visual stimuli as it may relate to action vs. perception. In addition — and focusing on differences within perception — the comparison across timing systems would also contribute to test further the hypothesis of distinct sensorimotor networks for beat-based and duration-based timing (Grube et al., 2010; McAuley & Jones, 2003; Pashler, 2001; Teki et al., 2011). Replicating previous findings of continuity effects on visually-driven synchronization but not on beat-based visual perception (Silva & Castro, 2016) was a subsidiary goal.

In order to achieve our goals, we carried out a behavioral study where a single group of participants performed beat-based synchronization, a beat-based perceptual task and a duration-based perceptual task. The two perceptual tasks were designed to engage similar time intervals, processes (interval comparison, see Grondin, 2010) and response mechanisms (forced choice) while addressing the specifics of each perceptual system. The beat-based perceptual task was the same we used in Silva and Castro (2016), which targeted the extraction of an *implicit* regular pulse. Participants viewed/heard standard temporal sequences that combined 600 ms (beats) and 300 ms intervals (half-beats), which kept an integer ratio between them (1:2). Some sequences contained deviant intervals, creating non-integer ratios (e.g., between 600 ms and 433 ms). Participants' task was to discriminate between standard, integer-ratio, 'right' sequences from 'wrong' sequences containing non-integer ratios. The point of including half-beats was to avoid mere isochrony judgements, which could arise from using a constant 600 ms interval (an explicit beat) as reference. The duration-based perception task addressed absolute timing in the sense that no reference unit (beat) could be detected in the stimulus sequences. Participants were shown sequences of two time intervals (ranging between 167 and 733 ms) defining non-integer ratios and were then asked to judge whether each of these sequences seemed to either speed up or slow down. In each of the three tasks (beat synchronization, beat perception and duration perception), participants were tested with optimal/moving (bouncing balls) vs. non-optimal/static visual (flashes) and optimal auditory (beeps) stimuli. Our analysis focused on the comparisons of bouncing balls with beeps, and flashes with beeps.

In the synchronization and beat-based perception tasks, we expected to replicate our previous findings (Silva & Castro, 2016). For synchronization, this means that we expected balls to outperform flashes and to be as effective as beeps (continuity effect in the visual modality). For beat-based perception, we expected to see only a modality effect, with no continuity effects in the visual modality (both balls and flashes underperforming beeps).

Concerning the comparison between beat-based and duration-based perception, there could be two scenarios. First, continuity effects in the visual domain could change according to the perceptual timing system — showing up in duration-based perception but not in beat-based perception. This would suggest that the advantage of moving over static visual stimuli relates to brain networks subtending duration-based processing, and is not action-specific. Second, continuity effects could be absent in both timing systems. In this case, we would have reasons to think that the origins of the advantage of moving over static visual stimuli lie outside the perceptual system in a broad sense, i.e., outside both beat- and duration-based perception.

2. Material and Method

2.1. Participants

Forty-five participants (11 men) took part in the experiment. Ages ranged between 18 and 32 years ($M = 20.62$, SD = ±2.67), and schooling between 14 and 18 years. They had normal or corrected-to-normal vision, and did not report psychiatric, neurological, motor or hearing disorders. All participants but six were right-handed. Sixteen had had formal music training beyond elementary school curricula, but only seven did so for more than three years (three: 5 years; two: 7 years; one: 8 years; one: 10 years). They all signed informed consent according to the Declaration of Helsinki.

2.2. Stimuli

Visual sequences (flashes, balls) consisted of videos at 30 frames per second, and auditory sequences (beeps) of 16-bit mono audio files at 44.1 kHz sampling frequency (Fig. 1A). Stimulus details are next described, for the three levels of stimulus type. *Beeps*: short (67 ms) sinusoidal tones ($F_0 = 450$ Hz); *flashes*: short (67 ms) flashes of a static blue ball (2.1° diameter) centered over a black background;

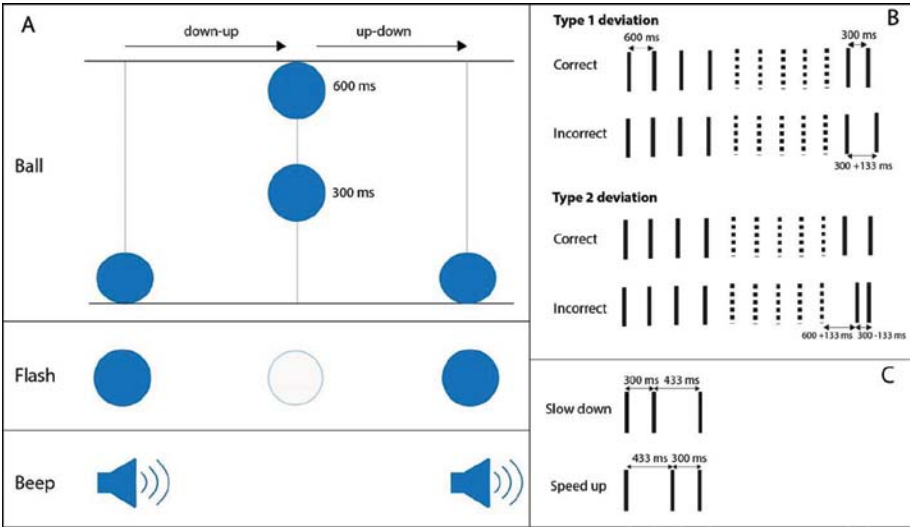

Figure 1. (A) Stimulus types (time on the x-axis): bouncing balls followed a down–up/up–down cycle of continuous movement, while flashes and beeps appeared for 67 ms at the cycle onset (down for balls, start point). In beat-based tasks, shorter cycles (300 ms) were implemented with lower distances from the ground (lower end point). (B) Example stimulus sequences for the beat-based perception task (see text). (C) Example stimulus sequences for the duration-based perception task.

bouncing balls: the same ball bouncing on an imaginary ground, squashing at the lower point of the trajectory. The spatial trajectory of the bouncing ball was linear.

In the duration-based perception task, we tested participants' abilities in judging whether a sequence of intervals seemed to be either speeding up or slowing down. For that purpose, we created 16 sequences of two intervals (three events) in which half of them were speeding up and half slowing down (Fig. 1C). The first interval ranged between 133 and 733 ms. In speed up sequences, the second interval was 566, 300, 167, 133, or 34 ms shorter than the first one (Appendix 1). In slow down sequences, the second interval was 566, 300, 167, 133, or 34 ms larger than the first. Therefore, for each value of interval shortening in speed up versions (e.g., −133 ms) there was an equal enlargement in slow down versions (+133 ms). The reason why we created two modalities (speed up vs. slow down) was to prevent possible biases: for instance, slow down sequences could look more natural for bouncing balls. Therefore, we decided to explore two opposite modes.

In the beat-based synchronization task, sequences were isochronous and included 48 events (beep onset, flash onset, bouncing ball hitting the ground) with inter-onset intervals (IOS) of 600 ms (beat length). Beeps and flashes lasted for 67 ms, thus having discontinuous presence. Bouncing balls were always present on screen and moved up (end point, Fig. 1A) and down (start point) continuously, hitting the ground every 600 ms (1.6 Hz frequency). In the beat-based perception task

(Fig. 1B), we tested participants' abilities in judging whether sequences of isochronous intervals (600 ms) ending with shorter intervals sounded 'right' (correct sequences, ending with 300 ms intervals, half-beat) or 'wrong' (incorrect sequences, ending with 300 ± 133 ms intervals, non-integer fraction). For this task, we created eight correct sequences with length ranging from 4200 ms (7 beats) to 6000 ms (10 beats). Correct sequences presented a series of whole beats ending either with two half-beats (stimuli 1–4, see Appendix 2) or only with one half-beat (stimuli 5–8). In half-beats, the bouncing ball's trajectory was shorter (lower end point, see Fig. 1A). Incorrect sequences were derived from correct ones by adding or subtracting 133 ms to either one or two intervals in the terminal part of the sequence.

Half of the incorrect sequences were designed so that the probe interval started on time and had an incorrect length (300 ± 133 ms *type 1 deviation*); the other half had a probe interval started out of time and an incorrect length (*type 2 deviation,* Fig. 1B). Type 1 deviations included one incorrect interval, while type 2 deviations included two. The four type-1-deviation sequences included two shortened probe intervals (300 − 130 ms) and two enlarged intervals (300 + 130 ms); the same went for type-2-deviation sequences. The reason why we created two deviation types was twofold. First, we needed different sequences for the eight trials and this would be difficult to achieve by varying sequence length only. Second, we wanted to maximize the indices of discrimination across stimulus type levels, and the response to deviation types seemed to be a good approach.

2.3. Procedure

We ran the experiment on E-prime 2 (https://pstnet.com/products/e-prime/). Participants sat 55 cm away from a Samsung Syncmaster 957DF monitor, with a Roland SPD-8 MIDI drum pad sideways (side of the dominant hand). First, they performed the duration-based temporal perception task, then the synchronization task, and lastly the beat-based temporal perception task.

In the duration-based perception task, they were asked to judge whether each of the 16 sequences (8 + 8) was either speeding up or slowing down, by pressing key '1' or '2' on the computer keyboard. We showed them one example of each stimulus type (ball, beep and flash) speeding up and slowing down, and then clarified possible doubts. In the synchronization task, participants were instructed to use a stick for tapping along with the stimulus (beep onset, flash onset, ball hitting the ground) for as long as it lasted. The audio signals generated by tapping the drum pad were recorded in an audio file whose onset was locked to the onset of the stimulus. Participants wore headphones in all tasks – in the auditory ones to listen to the stimuli, and in the visual ones to minimize any noise from outside the room. In the beat-based perception task, they were asked to judge whether each of 16 sequences (8 + 8) was correct or incorrect by pressing key '1' or '2' on the computer keyboard. We told them that correct versions should sound/look like someone was walking and then started to walk faster, while incorrect versions should sound/look like someone who suddenly started to walk with a limp. Sequence presentation was randomized across participants and stimulus types. They went through all stimulus types (ball, beep, flash) in each task before proceeding to the next task.

The three stimulus types were ordered in four different ways: beep–ball–flash, beep–flash–ball, ball–flash–beep and ball–beep–flash. For each of these orders, we created two conditions in the perception task: one in which the left key ('1') meant correct and another where it meant incorrect, and one which the left key ('1') meant speed up and another where it meant slow down. Each participant was assigned to one of these eight conditions (four orders × two keys). Half of the participants performed auditory first, and the other half visual first.

At the end of the experimental session, participants were given a questionnaire on strategies that they might have used, namely relying on recoding of visual stimuli or vice-versa.

2.4. Data Preprocessing and Statistical Analysis

In line with recent studies in the field (Gan et al., 2015; Hove et al., 2013a, b; Iversen, 2015; Silva & Castro, 2016), the tapping time series from the synchronization task were analyzed applying circular statistics method (Fisher, 1993) as implemented in the Circstats toolbox for Matlab (Berens, 2009). The audio files generated by participants in the synchronization task (their taps) were first analyzed with software Praat (http://www.fon.hum.uva.nl/praat/). Tap onsets were detected with the function 'annotate-to text grid (silences)', which determines the onset and offset of silent vs. sounding periods in the audio files. Since synchronization typically requires a few taps to stabilize, the first two seconds of the sequence were discarded from analysis, and so were deviant intervals (longer than 1000 ms and shorter than 200 ms). Synchronization performance was quantified in three ways. In the first one, we computed the *mean asynchronies* of taps relative to beats. Asynchronies were represented in terms of relative phase: each tap was mapped onto a circular unit ranging from $-\pi$ to π. Zero values mean perfect alignment to the beat; negative values (0 to $-\pi$) indicate that the tap preceded the beat, and positive values (0 to $+\pi$) that the tap followed the beat. Mean asynchronies correspond to the mean relative phase per subject (ranging from $-\pi$ to $+\pi$). In the second measure, we assessed *synchronization stability*, which is also based on asynchronies. In circular statistics, synchronization stability is described by the R index. R indexes the regularity of the tap-to-target coordination, and it ranges from 0 (unstable tapping with uniformly distributed relative phases) to 1 (perfect stable tapping with a unimodal distribution of relative phases). Stability correlates inversely with circular variance (CV), such that $R = 1 - \mathrm{CV}$. Finally, we measured the *error correction* for period, as indexed by lag-1 autocorrelation for Inter-Tap-Intervals. The negative value of the lag-1 autocorrelation means that a longer interval tends to be followed by a shorter one, which can be taken as a sign of online error correction (Iversen et al., 2015). These three linear measures were analyzed with repeated-measures ANOVAs with *stimulus type* as within-subjects factor. The ANOVA was followed by cross-stimulus comparisons in case of interaction, using paired-samples *t*-tests Bonferroni-corrected for multiple comparisons.

Performance in the perception tasks (discrimination between speed up and slow down versions; right and wrong) was approached with d' measures (Stanislaw & Todorov, 1999). We used repeated-measures ANOVAs with *timing system* (duration-based vs. beat-based) and *stimulus type* (ball, beep, flash) as factors. As a planned comparison, we then analyzed stimulus type effects on duration-based and beat-based separately, followed by paired-samples *t*-tests comparing stimulus type levels. We added Bayesian analyses (Bayes factors) with non-informative priors to further investigate null results concerning stimulus type effects.

In our previous study (Silva & Castro, 2016), the sensitivity to different deviation types (type I vs. II, see Materials) mirrored the results of discrimination analysis, in that visual stimuli showed increased sensitivity to type II deviations while auditory ones showed no preference. In order to strengthen the possibility that this type of analysis (sensitivity to deviation types) complements discrimination analysis, we examined once again the interaction between *deviation type* (type 1 vs. type 2, see Appendix 2) and stimulus type on accuracy for incorrect targets (correct rejections). In a parallel approach, we analyzed the *change direction* (speed up vs. slow down, see Appendix 1) × stimulus type interaction in the duration-based perception task.

To investigate further the association/dissociation of timing systems (duration and beat-based), we tested the correlation between performance in duration-based and beat-based systems. Lastly, we also tested the correlation of synchronization performance with perceptual discrimination in the two timing systems in order to get additional results concerning the relation between production and perception.

Even though we had few participants with musical training and training was relatively modest, we wanted to rule out any effects of musical experience. Musical experience is known to influence rhythmic performance, specifically synchronization skills (Chen et al., 2008). To that purpose, we tested if the years of musical training correlated either with synchronization or perceptual discrimination.

3. Results

3.1. Synchronization

3.1.1. Mean Direction of Asynchronies

The mean direction of asynchronies (see Sect. 2.4) was negative for all conditions. Negative values indicated that the taps occurred before the beat, and that there was anticipating behavior consistent with beat-based timing (Fig. 2). There was no significant effect of stimulus type, $F(1.45) = 2.969$, $p = 0.074$; $\eta^2_p = 0.063$.

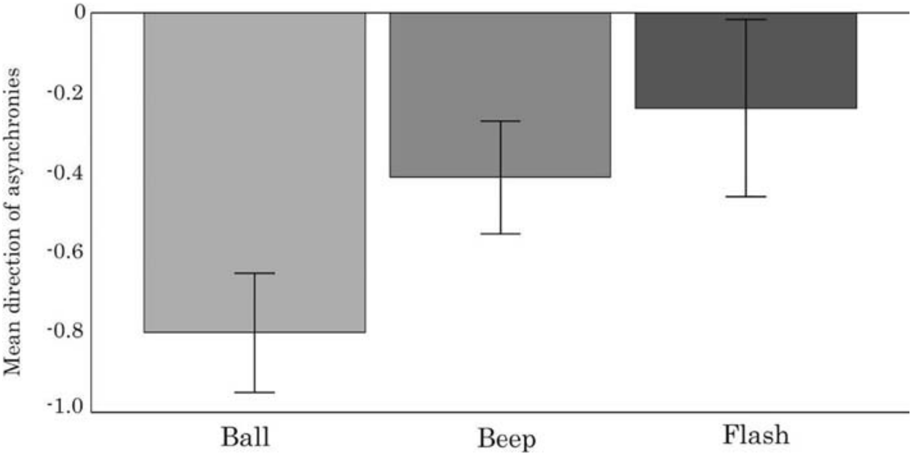

Figure 2. Mean direction of tap–beat asynchronies (represented as relative phases) as a function of stimulus type (moving visual – ball; auditory – beep; static visual – flash). Vertical bars represent the standard error of the mean.

3.1.2. Stability of Synchronization

The ANOVA for stability of synchronization showed a significant effect of stimulus type, $F(2,88) = 24.482$, $p < 0.001$; $\eta^2_p = 0.357$. Cross-stimulus comparisons showed non-significant differences between balls and beeps, $t(44) = 0.310$, $p = 0.756$; the remaining comparisons yielded significant differences between beeps and flashes, $t(44) = 5.771$, $p < 0.001/0.003$ (uncorrected/corrected p), $d = 1.11$, and balls and flashes, $t(44) = 6.102$, $p < 0.001/0.003$, $d = 1.08$. Thus, beeps and balls were equivalent in driving synchronization stability, and both outperformed flashes (Fig. 3).

3.1.3. Lag-1 Autocorrelation

There was a significant effect of stimulus type, $F(2,88) = 7.853$, $p = 0.001$, $\eta^2_p = 0.161$. Beeps and balls presented negative values (larger for balls), showing that subjects corrected their own errors, whereas flashes showed positive ones, suggesting perseverance of the error (Fig. 4). Comparisons across the three conditions indicated a significant difference between balls and flashes, $t(44) = -4.439$, $p < 0.001/0.003$, $d = 0.40$; the remaining comparisons yielded non-significant results [balls vs. beeps: $t(44) = -1.513$, $p = 0.138$; beeps vs. flashes: $t(44) = -2.493$, $p = 0.017/0.051$].

3.2. Perception

3.2.1. Discrimination (*D-prime*)

The d' values differed significantly from zero in all conditions (all $p < 0.003$). The interaction between stimulus type and timing system on d' was marginal,

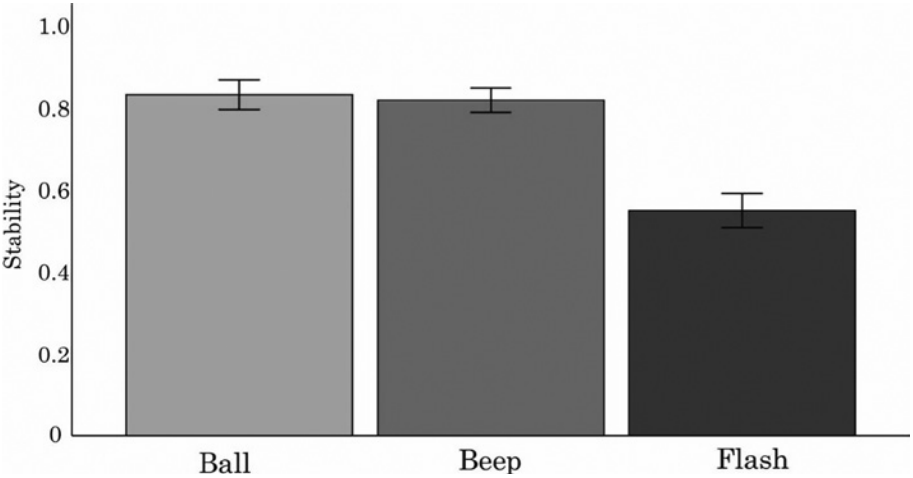

Figure 3. Synchronization stability as a function of stimulus type (moving visual – ball; auditory – beep; static visual – flash). Vertical bars represent the standard error of the mean.

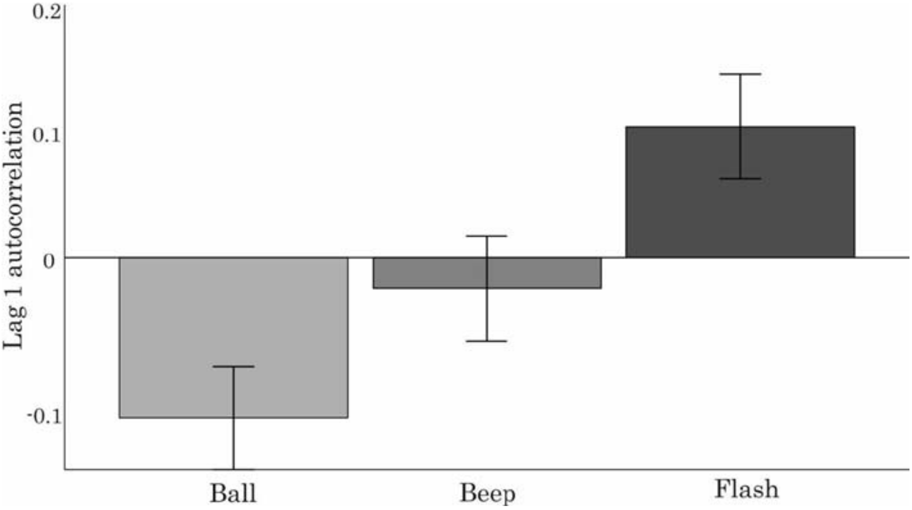

Figure 4. Lag-1 autocorrelation as a function of stimulus type (moving visual – ball; auditory – beep; static visual – flash). Vertical bars represent the standard error of the mean.

$F(2,88) = 2.682$, $p = 0.074$, $\eta^2_p = 0.057$. The effect of stimulus type on duration-based perception was non-significant, $F < 1$, but it was significant for beat-based perception, $F(2,88) = 7.473$, $p = 0.001$, $\eta^2_p = 0.145$ (Fig. 5). In the duration-based condition, there were no significant differences among the three stimuli [balls vs.

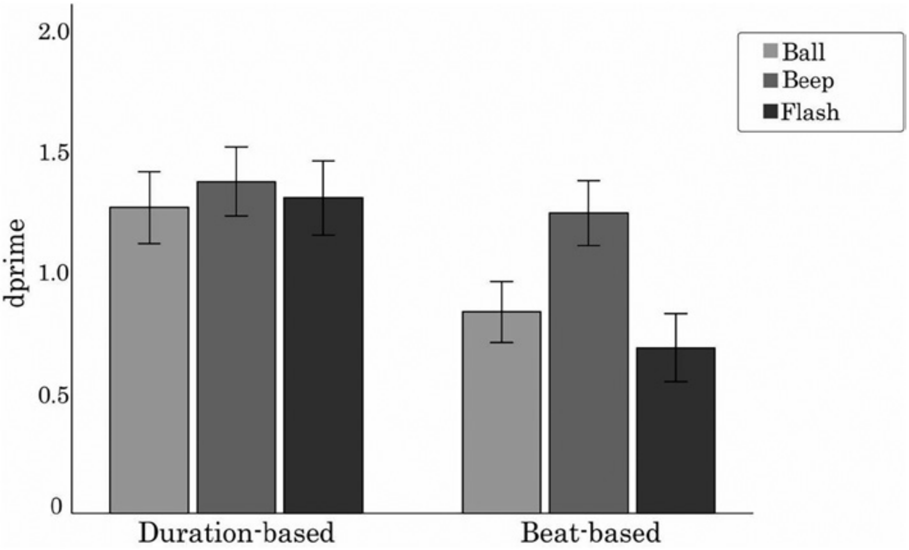

Figure 5. Discrimination between speed up and slow down sequences (duration-based) and correct and incorrect sequences (beat-based) as a function of stimulus type (moving visual – ball; auditory – beep; static visual – flash). Vertical bars represent the standard error of the mean.

beeps, $t(44) = -0.706$, $p = .484$; balls vs. flashes, $t(44) = -0.268$, $p = .790$; beeps vs. flashes, $t(44) = 0.444$, $p = 0.660$], while in the beat-based condition, beeps outperformed balls and flashes, and balls yielded no significant differences compared to flashes [beeps vs. flashes, $t(44) = 3.664$, $p = .001/0.003$, $d = 0.60$; balls vs. beeps, $t(44) = -2.532$, $p = .015/0.045$, $d = 0.46$; balls vs. flashes, $t(44) = 0.813$, $p = .271$ (Fig. 5). Bayes factors for cross-stimulus comparisons in the duration-based task showed substantial evidence in favor of the null hypothesis [balls vs. beeps, 7.02; balls vs. flashes, 8.80; beeps vs. flashes, 8.00].

Comparisons for beeps in beat-based vs. duration-based conditions showed no significant differences [$t(44) = -0.728$, $p = 0.471$]. Bayesian evidence in favor of the null hypothesis was substantial (Bayes factor of 7.96).

3.2.2. *Effects of Change Direction × Stimulus Type on Duration-Based Perception*

The ANOVA showed a significant interaction between change direction (speed up vs. slow down) and stimulus type on accuracy for duration-based perception, $F(2,88) = 5.947$, $p = 0.004$, $\eta^2_p = 0.119$. Comparisons between speed up and slow down across the three stimulus types showed enhanced performance for speed up in balls and flashes [balls, $t(44) = -3.604$, $p = 0.001/0.003$, $d = 0.57$; flashes, $t(44) = -4.135$, $p < 0.001/0.003$, $d = 0.64$], while for beeps there were no differences between speed up and slow down [$t(44) = 0.137$, $p = 0.892/2.676$]. This pattern (beeps vs. balls and flashes) does not match the results of discrimination analysis (cf. Sect. 3.2.1, no differences across the three stimuli), suggesting that sensitivity to change direction is irrelevant for characterizing the efficiency of duration-based perception.

3.2.3. Effects of Deviation Type and Stimulus Type on Beat-Based Perception

The ANOVA on the correct rejections of wrong sequences showed a non-significant main effect of deviation type (type 1 vs. type 2, see Sect. 2.2 and Appendix 2), $F(1,49) = 0.922$, $p = 0.342$, $\eta^2_p = 0.018$, as well as a non-significant interaction between deviation type and stimulus type, $F(2,98) = 2.234$, $p = 0.113$, $\eta^2_p = 0.44$. This indicates that balls, beeps and flashes do not differ in conveying different error structures (incorrect onset and incorrect interval vs. incorrect onset only). This result does not mirror the differences between stimulus types observed in the discrimination analysis (cf. Sect. 3.2.1), thus not confirming our previous findings (Silva and Castro, 2016) and suggesting that the sensitivity to specific deviation types might not work as complementary index of perceptual performance across stimulus types.

3.2.4. Correlation between Beat-Based Timing and Duration-Based Timing

The analysis showed no significant correlation between timing systems (beat-based and duration-based) for any stimulus type level (balls across timing systems: $r = 0.058$, $p > 0.70$; beeps: $r = 0.167$, $p > 0.27$; flashes: $r = 0.264$, $p = 0.080$). This dissociation converges with the differences between timing systems observed in the discrimination analysis (Sect. 3.2.1).

3.3. Correlations between Synchronization and Perception

Flashes in duration-based timing correlated moderately with flashes in synchronization stability ($r = 0.357$, $p = 0.016/0.048$). Beyond this, synchronization stability did not shown any other correlation with the *d*′ values of duration-based (balls: $r = 0.023$, $p > 0.87$; beeps: $r = 0.175$, $p > 0.25$) or beat-based perceptual tasks (balls: $r = 0.235$, $p > 0.12$; beeps: $r = 0.112$, $p > 0.42$; flashes: $r = 0.222$, $p > 0.14$).

3.4. Correlations between Musical Expertise, Synchronization Stability and Perceptual Discrimination

The number of years of musical training did not correlate significantly with beat-based perceptual discrimination (balls: $r = 0.139$, $p > 0.36$; beeps: $r = 0.065$, $p > 0.67$; flashes: $r = 0.018$, $p > 0.90$), duration-based perceptual discrimination (balls: $r = -0.064$, $p > 0.67$; beeps: $r = -0.123$, $p > 0.52$; flashes: $r = -0.035$, $p > 0.81$) or synchronization stability (balls: $r = -0.015$, $p > 0.92$; beeps: $r = 0.008$, $p > 0.95$; flashes: $r = -0.092$, $p > 0.53$). When we considered expertise as the number of years of musical *practice* instead (including, thus, informal musical training), correlations included marginally-significant cases (correlation of practice with duration-based perception – flashes: $r = 0.281$, $p = 0.061$; with synchronization stability – beeps: $r = 0.267$, $p = 0.069$, flashes: $r = 0.243$, $p = 0.100$).

4. Discussion

Recent findings indicated that moving visual stimuli (bouncing balls) can outperform static ones (visual flashes) and be as effective as auditory stimuli (beeps) in sensorimotor synchronization, but not in beat-based purely perceptual tasks. These findings suggested that the advantage of moving visual stimuli over static ones could be action-specific, but one piece of the puzzle was missing: is the advantage of moving visual stimuli also absent in duration-based perceptual tasks? This was the main question we addressed in the present study, and the answer we found was 'yes': the advantage of moving over static visual stimuli seems to be broadly action-specific, in that it is absent in perception, whatever the timing system.

In order to answer our question, we started by checking the replicability of previous findings concerning the beat-based timing system — i.e., that moving visual stimuli outperform static ones in synchronization, but not in beat-based perceptual performance. Beat-based synchronization results replicated previous findings: balls not only outperformed flashes, but they also matched beeps in measures of stability and error correction. The same went for beat-based perception, in that moving visual stimuli (bouncing balls) were as insufficient as static ones (flashes) to facilitate discrimination between correct and incorrect versions, and both were less efficient than beeps.

Concerning duration-based perception, we saw no differences between bouncing balls and flashes. This indicates that the advantage of moving over static visual stimuli that characterizes sensorimotor synchronization, and which was absent in beat-based perception, is also absent in duration-based perception. Therefore — and in response to our main questions — it seems that the origins of such an advantage lie outside the perceptual system in a broad sense, since neither beat-based nor duration-based perception benefits from movement in visual stimuli. At the same time, the finding that both perceptual systems are insensitive to visual movement, in contrast to synchronization, strengthens the hypothesis of dissociation between rhythmic production and rhythmic perception (Fujii & Schlaug, 2013).

Despite similarities concerning the response to visual movement, beat- and duration-based stimuli dissociated in a few points. First, duration-based perception did not show modality effects: in contrast to beat-based perception, visual stimuli (bouncing balls, flashes) and auditory ones (beeps) elicited the same performance levels. This may come as a surprise in light of intrinsic models of timing (see Grondin, 2010, for a review), which postulate modality-specific mechanisms for duration judgements with a disadvantage for vision. However, even these models highlight the importance of timescale when considering modality specificity, and the interval ranges we used (167–733 ms, mean 392 ms) may have been responsible for the lack of modality effects. To be specific, findings in this field have converged on the idea that short intervals are sensitive to modality while long ones are not. The boundary between short and long is not consensual, though. For instance, Fornaciai et al. (2018) found modality effects on the perception of 300 ms intervals, Murai and Yotsumoto (2016) found the same on the reproduction of 400–600 ms intervals. In contrast, Barne et al. (2018) found evidence of a common neural substrate across vision and audition for the perception of 750–1000 ms intervals. Even though it seems that the boundary between short and long is located somewhere between 600 and 750 ms, Rammsayer and colleagues (Rammsayer & Pichelmann, 2018; Rammsayer & Troche, 2014; Rammsayer et al., 2015) have proposed that the area of transition between short and long intervals is located earlier, in the 400–600 ms range. From this viewpoint, it is reasonable to admit that our intervals (167–733 ms) were at least partly approached as long intervals, and this may have been the reason why modality effects were absent. Future research could investigate this possibility further, by contrasting short (50 ms), long (1000 ms) and transition intervals (400 ms) using our paradigm.

Differences between beat-based and duration-based perception concerning modality effects converged with the absence of correlation in performance across the two perceptual systems. Altogether, these findings strengthen the dissociation between beat-based and duration-based perceptual systems that has been highlighted in research (Keele et al., 1989; McAuley & Jones, 2003; Pashler, 2001; Teki et al., 2011).

One limitation of our study relates to the equivalence between our beat-based and duration-based tasks. Given our emphasis on creating stimuli with appropriate structures for each perceptual timing system, the two tasks differed in several aspects, and we cannot completely rule out that the dissociations between the two perceptual timing systems (presence vs. absence of modality effects; lack of correlation) was partly due to these differences. Specifically, it is possible that shorter sequence lengths in the duration-based task (1000–2000 ms, vs. 4000–6000 ms in beat-based sequences) facilitated performance to the point of generating ceiling effects. The equivalence we saw between performance in beat-based beeps and that in duration-based ones challenges this possibility, but further investigations would still be welcome. Another limitation came from insufficient variability in participants' musical experience, which prevented us from exploring possible effects on beat-based vs. duration-based perception. This could be addressed in future studies, using more extreme participant profiles concerning musical experience.

Finally, in this study we wanted to know whether the advantages of visual movement are action-specific. We addressed one missing piece of this puzzle, but we left another piece unanswered: we saw that the advantage of moving over static visual stimuli is absent in perception in a broad sense — i.e., both beat- and duration-based perception — but we did not examine if such advantage was present in production in a broad sense — i.e., we did not examine duration-based time production along with beat synchronization. We made this choice as part of a start-small approach, in which we avoided the difficulties of devising a suitable duration-based production task in this first approach to the main problem. One way of approaching this challenge in future research could be asking participants to hold a button for a given time (e.g., 2 s) while using moving vs. static visual stimuli to either train participants or mark the passing of time. Differences in accuracy across the two stimulus types could reveal whether visual movement makes a difference in duration-based production. An alternative could be the paradigm used by Breska & Ivry (2018), where participants are stimulated with two non-integer ratio intervals and then asked to produce a third interval, equal to the first one.

Our study contributed to expand the description of temporal processing in healthy individuals, allowing future advances in the identification of pathological markers. Another practical implication of our study concerns the possibility of optimizing stimulus selection for specific visual tasks: if we want someone to synchronize with a visual stimulus, it seems preferable to make the stimulus move continuously; if we want him/her to perceive a duration-based visual sequence, movement will not make a difference; if we want her/him to perceive a beat-based visual sequence, maybe we should quit and use an auditory one.

Acknowledgements

We would like to thank two anonymous reviewers for their comments on a previous version of this manuscript. We are grateful to our participants.

Funding

This work was funded by Fundação para a Ciência e a Tecnologia and COMPETE 2020 program under grant CPUP UID/PSI/00050/201.

References

Barne, L. C., Sato, J. R., de Camargo, R. Y., Claessens, P. M. E., Caetano, M. S. and Cravo, A. M. (2018). A common representation of time across visual and auditory modalities, *Neuropsychologia 119*, 223–232.

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics, *J. Stat. Softw.*. **31**, 1–21. doi: 10.18637/jss.vo31.i10.

Biswas, A., Hedge, S., Jhunjhunwala, K., and Pal, P. K. (2016). Two sides of the same coin: Impairment in perception of temporal components of rhythm and cognitive functions in Parkinson's disease, *Basal Ganglia 6*, 63–70.

Breska, A. and Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease, *Proc. Natl Acad. Sci. USA 115*, 12283–12288.

Carroll, C. A., O'Donnell, B. F., Shekhar, A. and Hetrick, W. P. (2009). Timing dysfunctions in schizophrenia span from millisecond to several-second durations, *Brain Cogn. 70*, 181–190.

Chen, Y., Repp, B. H. and Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions, *Hum. Mov. Sci. 21*, 515–532.

Chen, J. L., Penhune, V. B. and Zatorre, R. J. (2008). Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training, *J. Cogn. Neurosci. 20*, 226–239.

Comstock, D. C., Hove, M. J. and Balasubramaniam, R. (2018). Sensorimotor synchronization with auditory and visual modalities: behavioral and neural differences. *Front. Comput. Neurosci. 12*, 53. doi: 10.3389/fncom.2018.00053.

Fisher, N. I. (1993). *Statistical Analysis of Circular Data*. Press Syndicate of the University of Cambridge, Cambridge, UK.

Fornaciai, M., Markouli, E. and Di Luca, M. (2018). Modality-specific temporal constraints for state-dependent interval timing, *Sci. Rep. 8*, 10043. doi: 10.1038/s41598-018-28258-4.

Fujii, S. and Schlaug, G. (2013). The Harvard Beat Assessment Test (H-BAT): a battery for assessing beat perception and production and their dissociation. *Fron. Hum. Neurosci.* 7, 771. doi: 10.33.89/fnhum.2013.00771.

Gan, L., Huang, Y., Zhou, L., Qian, C. and Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory, *Sci. Rep. 5*, 11974. doi: 10.1038/srep11974.

Glenburg, A. M. and Jona, M. (1991). Temporal coding in rhythm tasks revealed by modality effects, *Mem. Cogn. 19*, 514–522.

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Exp. Brain Res. 220*, 51–61.

Grahn, J. A. and Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease, *Cortex 45*, 54–61.

Grahn, J. A., Henry, M. J. and McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: Audition primes vision, but not vice versa. *NeuroImage*, *54*, 1231–1243.

Grondin, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions, *Attent. Percept. Psychophys. 72*, 561–582.

Grube, M., Lee, K. H., Griffiths, T. D., Barker, A. T. and Woodruff, P. W. (2010). Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals, *Front. Psychol. 1*, 171. doi: 10.3389/fpsyg.2010.00171.

Guttman, S. E., Gilroy, L. A. and Blake, R. (2005). Hearing what the eyes see: auditory encoding of visual temporal sequences, *Psychol. Sci. 16*, 228–235.

Hove, M. J., Spivey, M. J. and Krumhansl, C. (2010). Compatibility of motion facilitates visuomotor synchronization, *J. Exp. Psychol. Hum. Percept. Perform. 36*, 1525–1534.

Hove, M. J., Fairhurst, M. T., Kotz, S. A. and Keller, P. E. (2013a). Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness, *NeuroImage*. *67*, 313–321.

Hove, M. J., Iversen, J. R., Zhang, A. and Repp, B. H. (2013b). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome, *Psychol. Res*. *77*, 388–398.

Iversen, J. R., Patel, A. D., Nicodemus, B. and Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals, *Cognition*. *134*, 232–244.

Ivry, R. B., Spencer, R. M., Zelaznik, H. N. Zelaznik, H. N. and Diedrichsen, J. (2002). The cerebellum and event timing, *Ann. N. Y. Acad. Sci. 978*, 302–317.

Keele, S. W., Nicoletti, R., Ivry, R. I. and Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgements? *Psychol. Res. 50*, 251–256.

McAuley, J. D. and Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: a comparison of interval and entrainment approaches to short-interval timing, *J. Exp. Psychol. Hum. Percept. Perform. 29*, 1102–1125. doi: 10.103//0096-1523.29.6.1102.

Merchant, H. and Lafuente, V. (2008). *Neurobiology of interval timing* (1st ed.). New York, NY, USA: Springer.

Morillon, B. and Baillet, S. (2017). Motor origin of temporal predictions in auditory attention, *Proc. Natl Acad. Sci. USA 114*, E8913–E8921. https://doi.org/10.1073/pnas.1705373114.

Murai, Y. and Yotsumoto, Y. (2016). Timescale- and sensory modality-dependency of the central tendency of time perception, *PLoS One 11*, e0158921. https://doi.org/10.1371/journal.pone.0158921.

Pashler, H. (2001). Perception and production of brief durations: beat-based versus interval-based timing, *J. Exp. Psychol. Hum. Percept. Perform. 27*, 485–493.

Patel, A. D., Iversen, J. R., Chen, Y. and Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat, *Exp. Brain Res. 163*, 226–238.

Peterburs, J., Nitsch, A. M., Miltner, W. H. and Straube, T. (2013). Impaired representation of time in schizophrenia is linked to positive symptoms and cognitive demand, *PLoS One 8*, 615. doi: 10.1371/journal.pone.0067615.

Pollok, B., Krause, V., Butz, M. and Schinitzler, A. (2009). Modality specific functional interaction in sensorimotor synchronization, *Hum. Brain Mapp. 30*, 1783–1790.

Rammsayer, T. and Pichelmann, S. (2018). Visual–auditory differences in duration discrimination depend on modality-specific, sensory-automatic temporal processing: Converging evidence for the validity of the Sensory-Automatic Timing Hypothesis, *Q. J. Exp. Psychol. 71*, 2364–2377.

Rammsayer T. H. and Troche S. J. (2014). In search of the internal structure of the processes underlying interval timing in the sub-second and second range: a confirmatory factor analysis approach, *Acta Psychol. 147*, 68–74.

Rammsayer, T. H., Borter, N. and Troche, S. J. (2015). Visual-auditory differences in duration discrimination of intervals in the subsecond and second range. *Front. Psychol. 6*, 1626. doi: 10.3389/fpsyg.2015.01626.

Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev. 12*, 969–992.

Repp, B. H. and Panel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences, *J. Exp. Psychol. Hum. Percept. Perform. 28*, 1085–1099.

Silva, S. and Castro, S. L. (2016). Moving stimuli facilitate synchronization but not temporal perception, *Front. Psychol. 7*, 1798. doi: 10.3389/fpsyg.2016.01798.

Stanislaw, H. and Todorov, N. (1999). Calculation of signal detection theory measures, *Behav. Res. Methods Instrum. Comput.. 31*, 137–149.

Stauffer, C. C., Haldemann, J., Troche, S. J. and Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: evidence from hierarchical structure of modality-specific and modality-independent levels of temporal information processing, *Psychol. Res. 76*, 20–31.

Su, Y.-H. (2014). Audiovisual beat induction in complex auditory rhythms: [point-light figure movement as an effective visual beat, *Acta Psychol.* (*Amst.*) *151*, 40–50.

Teki, S., Grube, M., Kumar, S. and Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing, *J. Neurosci. 31*, 3805–3812.

Thoenes, S. and Oberfeld, D. (2017). Meta-analysis of time perception and temporal processing in schizophrenia: Differential effects on precision and accuracy, *Clin. Psychol. Rev. 54*, 44–64.

Yee, W., Holleran, S. and Jones, M. R. (1994). Sensitivity to event timing in regular and irregular sequences: influences of musical skill, *Percept. Psychophys. 56*, 461–471.

Zelaznik, H. N., Spencer, R. M. and Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements, *J. Exp. Psychol. Hum. Percept. Perform. 28*, 575–588. https://doi.org/10.1037/0096-1523.28.3.575.

Appendix

Appendix 1.

Stimulus sequences for Speed up and Slow down (ms).

	Type	Interval 1	Interval 2	Difference		Type	Interval 1	Interval 2	Difference
1	Slow down	300	433	−133	9	Speed up	433	300	133
2	Speed up	300	167	133	10	Slow down	167	300	−133
3	Speed up	467	433	34	11	Slow down	433	467	−34
4	Speed up	733	167	566	12	Slow down	167	733	−566
5	Slow down	300	467	−167	13	Speed up	467	300	167
6	Speed up	433	133	300	14	Slow down	133	433	−300
7	Speed up	467	300	167	15	Slow down	300	467	−167
8	Speed up	733	433	300	16	Slow down	433	733	−300

Appendix 2.

Stimulus sequences (IOIs) for error detection task.

	Error type	Intervals correct version (ms)	Intervals incorrect versions (ms)
1	Type 1	600–600–600–600–600–600–600–600–300–300–(600)	600–600–600–600–600–600–600–600–*433*–(467)
2	Type 1	600–600–600–600–600–600–300–300–(600)	600–600–600–600–600–600–300–*167*–(733)
3	Type 1	600–600–600–600–600–600–600–300–300–(600)	600–600–600–600–600–600–600–*433*–(467)
4	Type 1	600–600–600–600–600–300–(600)	600–600–600–600–600–*167*–(733)
5	Type 2	600–600–600–600–600–600–600–600–600–300–(300)	600–600–600–600–600–600–600–600–600–*467*–*433*–(300)
6	Type 2	600–600–600–600–600–600–600–600–(300)	600–600–600–600–600–600–600–600–*733*–*167*–(300)
7	Type 2	600–600–600–600–600–600–600–300–(300)	600–600–600–600–600–600–600–*467*–*433*–(300)
8	Type 2	600–600–600–600–600–600–(300)	600–600–600–600–600–600–*733*–*167*–(300)

Numbers in italics indicate intervals that were changed in incorrect versions. The final interval is indicated in parentheses as it has an undefined end point (the end of the stimulus).