

Compatibility of Motion Facilitates Visuomotor Synchronization

Michael J. Hove

Max Planck Institute for Human Cognitive and
Brain Sciences and Cornell University

Michael J. Spivey

University of California, Merced

Carol L. Krumhansl
Cornell University

Prior research indicates that synchronized tapping performance is very poor with flashing visual stimuli compared with auditory stimuli. Three finger-tapping experiments compared flashing visual metronomes with visual metronomes containing a spatial component, either compatible, incompatible, or orthogonal to the tapping action. In Experiment 1, synchronization success rates increased dramatically for spatio-temporal sequences of both geometric and biological forms over flashing sequences. In Experiment 2, synchronization performance was best when target sequences and movements were directionally compatible (i.e., simultaneously down), followed by orthogonal stimuli, and was poorest for incompatible moving stimuli and flashing stimuli. In Experiment 3, synchronization performance was best with auditory sequences, followed by compatible moving stimuli, and was worst for flashing and fading stimuli. Results indicate that visuomotor synchronization improves dramatically with compatible spatial information. However, an auditory advantage in sensorimotor synchronization persists.

Keywords: sensorimotor synchronization, visuomotor synchronization, modality effects, timing, common coding

Sensorimotor synchronization is generally found to be more difficult and more variable with visual rhythms than auditory or tactile rhythms. People rarely synchronize spontaneously with purely visual rhythms, whereas they do with music. The ability to synchronize with music is observed in young children (Eerola, Luck, & Toiviainen, 2006), with a wide range of musical styles (van Noorden & Moelants, 1999), and after hearing only a few beats (Snyder & Krumhansl, 2001). Rhythmic finger tapping has been found to be more variable with flashing visual stimuli than with auditory stimuli (Chen, Repp, & Patel, 2002; Repp & Penel, 2002, 2004), with tactile stimuli intermediate (Kolers & Brewster, 1985). Moreover, reliable synchronization in 1:1 tapping is possible at rates as fast as an interonset interval (IOI) of about 200 ms for auditory sequences (e.g., Fraisse, 1982), contrasted with IOIs around 460 ms for flashing visual sequences (Repp, 2003).¹

The apparent difficulty in synchronizing with visual stimuli has yet to be explained adequately. It may simply stem from the less

frequent occurrence of visual than auditory rhythms in our environment. Another possibility is that it is based on differences in neural connectivity. Fraise (1948) suggested that the action system is more closely linked to the auditory system than to the visual system. More recently, Thaut, Kenyon, Schauer, and McIntosh (1999) proposed a comparatively direct connection between auditory cortex and the reticulospinal neurons used to control movement, which results in increased sensorimotor coupling. Yet another possibility for the observed performance differences stems not from differential connectivity between motor neurons and auditory versus visual pathways, but from the inherent processing styles of those two sensory systems themselves. The auditory system is generally better at resolving temporal variation (e.g., Conway & Christiansen, 2005), whereas the visual system is better at resolving spatial variation (e.g., Posner, Nissen, & Klein, 1976). Under this account, visual information naturally dominates when one is attempting to identify the *spatial location* of a sound (as in the ventriloquism effect; Bertelson & Radeau, 1981), and auditory information naturally dominates when one is attempting to identify or behave contingent on the *temporal incidence* of a sound (as in synchronized tapping tasks; Repp & Penel, 2002, 2004).

The observed difficulties in visuomotor synchronization may appear simply because the vast majority of studies employ purely temporal flashing stimuli devoid of spatial information. This focus on flashing stimuli dates back a century (Dunlap, 1910), and more recently extends into examinations of the neural substrates of visuomotor synchronization in functional MRI (e.g., Jäncke,

This article was published Online First August 9, 2010.

Michael J. Hove, Music Cognition and Action Research Group, Max Planck Institute for Human Cognitive and Brain Sciences, and Department of Psychology, Cornell University; Michael J. Spivey, Cognitive and Information Sciences, University of California, Merced; Carol L. Krumhansl, Department of Psychology, Cornell University.

This work was supported by an American Psychological Association Dissertation Research Award and a Mind and Life Institute Varela Research Grant to Michael J. Hove. We thank Bruno Repp and Peter Keller for comments on a draft of this article.

Correspondence concerning this article should be addressed to Michael J. Hove, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, 04103 Leipzig, Germany. E-mail: michaeljhove@gmail.com

¹ The rate limit in 1:1 tapping seems to reflect the maximum finger frequency. When tapping with every fourth onset, the auditory IOI limit decreases to 100–120 ms (Repp, 2003).

Loose, Lutz, Specht, & Shah, 2000), positron-emission tomography (Penhune, Zatorre, & Evans, 1998), and magnetoencephalography studies (Chen, Ding, & Kelso, 2003). Repp, Patel, and colleagues (Patel, Iversen, Chen, & Repp, 2005; Repp & Penel, 2004) have speculated that different types of visual stimuli, namely those employing a spatial component, might facilitate synchronization.

A great deal of literature explores stimulus–response compatibility effects, but compatibility effects have not yet been demonstrated in the synchronized tapping paradigm. Studies of the synchronization of hand movements with an oscillating visual target found that inphase movements were more stable than antiphase movements (Roerdink, Peper, & Beek, 2005; Wimmers, Beek, & van Wieringen, 1992; for perception–action compatibility effects in bimanual coordination, see, e.g., Bogaerts, Buekers, Zaal, & Swinnen, 2003; Mechsner, Kerzel, Knoblich, & Prinz, 2001). However, it remains unclear how such results would compare with purely temporal flashes or action-neutral moving stimuli. In addition, facilitative effects of compatibility between dynamic visual displays and finger movements have been observed in reaction time tasks (Brass, Bekkering, & Prinz, 2001) and may extend to synchronization performance. Finally, the extrastriate body area in human occipital cortex responds selectively to images of the human body (Downing, Jiang, Shuman, & Kanwisher, 2001) and also is active during limb movements (Astafiev, Stanley, Shulman, & Corbetta, 2004). This suggests a link between perceptual and action-based body representations and might translate into different synchronization performance with human body images versus geometric images. In the following experiments, we investigated whether certain types of spatially varying and compatible visual information enable better synchronization than simple flashing or fading lights.

Experiment 1

Experiment 1 investigated sensorimotor synchronization with visual metronomes that were either purely temporal or contained additional spatial information. Those containing spatial information showed motion of either geometric or biological forms, which might afford different performance. Sequences were presented at two tempi: 500-ms IOI and 400-ms IOI (one above and one below the previously ascertained rate limit for flashing visual stimuli; Repp, 2003). The primary measures of performance were the success rate of synchronizing with the stimuli (i.e., the percentage of trials that had consistent tap-to-target asynchronies as determined by Rayleigh tests for uniform distribution), the serial dependencies between taps (lag-1 autocorrelations), and the mean phase or temporal position of taps relative to the targets.

Method

Participants. Eleven right-handed Cornell students (six women) between the ages of 19 and 23 years participated in the study. They were previously unfamiliar with the tapping task. Musical training ranged from 0 to 12 years ($M = 4.1$ years), although most were no longer active musicians. The duration of musical training did not significantly correlate with tap timing or synchronization success. Participants received course credit or were paid \$6.

Materials. The four stimulus sequences were QuickTime movies produced using the animation software After Effects. They were presented on a computer screen in a 10 cm × 8 cm viewing window. Videos lasted 26 cycles and each cycle consisted of 20 frames. In the *slow* tempo condition (500-ms IOI), videos were played at a frame rate of 40 frames per second (fps), and in the *fast* tempo (400-ms IOI), the same videos were played at 50 fps.² The four visual metronomes are shown as still pictures in Figure 1.

The *Flash* was a 2 cm × 2 cm stationary white square that remained on screen for 2 video frames. The *Up–Down Bar* was a 5 cm × 1 cm white bar that moved down frame-by-frame from its initial position 2.5 cm above the bottom of the viewing window to the bottom of the viewing window and back again. The *Rotating Bar* used the same bar and initial position as the Up–Down Bar, but the bar pivoted from its right edge until the left edge touched the bottom of the viewing window, then returned to its initial horizontal position. The *Finger* used spliced images of a right index finger (5 cm × 1 cm) tapping in the same configuration as the participant’s finger; the Finger’s vertical trajectory matched the Rotating and Up–Down Bars frame-by-frame, thereby isolating the effect of the image of bodily form without introducing the acceleration and deceleration characteristic of biological motion. The target position with which to synchronize taps in all the spatial metronomes was contact with the bottom of the screen; this was displayed for two frames to match the Flash target duration.

Procedure. Participants sat approximately 75 cm in front of a computer monitor that displayed the visual sequences at eye level. They positioned their right hand in front of themselves at approximately waist level parallel to the screen (pointing to the left) and tapped with their right index finger on a Roland Handsonic HPD-15 drum pad. Sequences were presented and taps were recorded using a MAX/Jitter program running on a Macintosh G4.

Participants were instructed to start tapping with the fifth cycle in the 26-cycle sequence; thus, each trial consisted of 22 taps. Each of the eight trial types (2 tempi × 4 metronome types) was presented in random order in a block. The experiment consisted of 20 blocks, including one training block, thus leaving 152 analyzed trials per participant. The entire experiment lasted approximately 45 min.

Analyses. Synchronized tapping experiments are typically analyzed using linear time series methods, wherein taps and targets are aligned on a linear time scale and their respective asynchronies are calculated. However, linear methods can be problematic with variable periodic data sets. For example, if a tap occurs nearly equidistant between two targets, one must decide which target to use for the asynchrony calculation, and a single large asynchrony with linear methods can greatly skew measures of variability. These issues can be circumvented using circular statistical methods (e.g., Fisher, 1993), wherein each tap in a trial is mapped onto a unit circle in terms of its relative phase (0–360°) from the periodic metronome target (always at 0°). Synchronization performance can then be assessed in terms of average phase (when taps tend to occur relative to the target) and the variability of the taps’

² The monitor refresh rate of 85 Hz (11.7 ms) led to slight deviations of frame timing in Experiments 1 and 2. However, no systematic error or drift occurred. Timing perturbations on this order were recently shown to have no effect on variability of intertap interval or tap-to-target asynchrony (Madison & Merker, 2004).

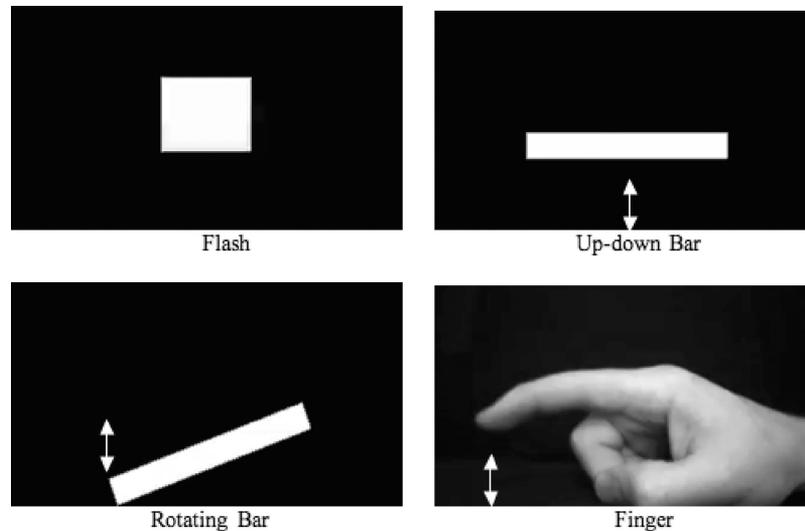


Figure 1. Stills from the four metronomes in Experiment 1; arrows added to depict motion.

relative phases (the stability or degree of tap-to-target synchrony). Visuomotor synchronization, especially at fast tempi, can generate unwieldy data. Thus, circular statistical methods provide a cleaner and more appropriate assessment of synchronization performance for this work (see, e.g., Fisher, 1993; Large & Palmer, 2002; or Kirschner & Tomasello, 2009, for more in-depth discussion of circular analyses).³

Tap timing was calculated in terms of each tap's relative phase (0–360°) from the target, for example, taps slightly after the target have a mean phase of 0–90° and those slightly before the target have a mean phase of 270–360°. Synchronization typically requires a few taps to stabilize; therefore, taps corresponding to the first seven video cycles (the first three taps) were omitted. An indicative measure of synchronization performance is the percentage of trials in which tap-to-target phases never stabilize (Repp, 2003). Each trial's tap-to-target synchronization performance was assessed using a Rayleigh test of uniformity, which tests the null hypothesis of a uniform distribution around the circle of tap-to-target relative phases against the alternative hypothesis of a nonuniform distribution (unimodal) of tap-to-target relative phases. For each trial, the null hypothesis of uniformity was rejected if the Rayleigh test *p* value was less than .05; these trials generally had a unimodal distribution of tap-to-target relative phases and were defined as successful, as opposed to trials with a uniform distribution of taps that would occur with irregular tapping or phase drift. Success rates can then be analyzed for the different metronomes and tempi.

Comparing the mean relative phases, or any circular response in multifactor designs, however, is problematic as “very little statistical methodology is available for analyzing data from experiments in which a circular response has been measured at various levels of two or more explanatory variables” (Fisher, 1993, p. 133). Therefore, relative phase data for tempo and for type of metronome were analyzed in separate one-way circular analyses of variance (ANOVAs; Watson–Williams test) that did not distinguish among participants (nonrepeated measures; Berens, 2009; Watson & Williams, 1956).

In addition, serial dependencies between successive tapping intervals (lag–1 autocorrelations) are reported. Patterns of serial depen-

dency provide clues about underlying processes in timing and are a key component that modelers of sensorimotor synchronization try to capture (e.g., Hary & Moore, 1985; Mates, 1994; Vorberg & Wing, 1996). Negative lag–1 autocorrelations, characterized by alternating between short and long tap intervals, can arise from motor implementation variance and from error correction mechanisms (e.g., Semjen, Schulze, & Vorberg, 2000; Wing & Kristofferson, 1973). Positive lag–1 autocorrelations, characterized by successive short or successive long tap intervals, suggest drift; this drift can be linear (wherein taps speed up or slow down throughout a trial) or nonmonotonic (wherein taps alternate between speeding up and slowing down within a trial; Madison, 2001; Vorberg & Wing, 1996).⁴ Lag–1 autocorrelations on tap intervals were run for each trial and transformed into Fisher *Z* scores. The *Z* scores were compared across metronome and tempo conditions in two-way repeated measures ANOVAs and tested for significant correlations in one-sample *t* tests against zero; Pearson *r* values are reported.

Results and Discussion

The average percentage of successful trials by condition is shown in Figure 2. A 4 (metronome type) × 2 (tempo) repeated

³ Analyses were also run using linear methods, wherein trial success used a threshold criterion based on the standard deviation of tap-to-target asynchronies less than 16.7% of IOI, following Repp's (2003) criterion for visuomotor synchronization. The success rates were very similar and results did not qualitatively differ from circular methods. Rayleigh tests, however, are more lenient than the criterion for auditory synchronization success in Repp (2003; *SD* of asynchrony < 10% of IOI).

⁴ Sensorimotor synchronization is divided between discrete and continuous behaviors (e.g., Spencer, Zelaznik, Diedrichsen, & Ivry, 2003; Zelaznik, Spencer, & Ivry, 2002). Discrete movements, such as the finger-tapping task used here, have a distinct beginning and end and often show negative lag–1 autocorrelation, whereas continuous movements, such as circle drawing, do not have clear endpoints and often show nonnegative lag–1 autocorrelation (Delignières, Torre, & Lemoine, 2008; Torre & Balasubramaniam, 2009).

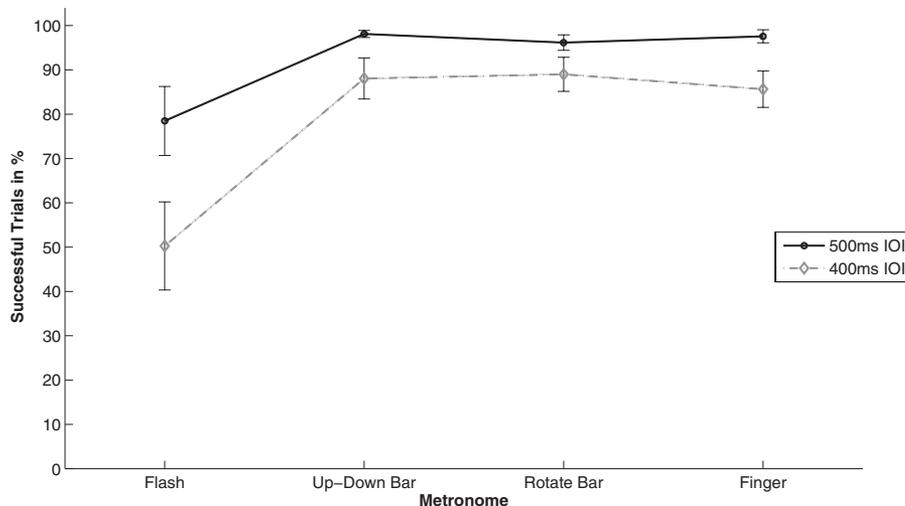


Figure 2. Percentage of successful trials in Experiment 1, as determined by a Rayleigh test of uniformity.

measures ANOVA revealed a main effect of metronome type, where successful trials were more frequent for the spatial metronomes than the Flash metronome, $F(3, 30) = 21.0$, $p < .001$, $\eta_p^2 = .67$. No significant differences among the spatial metronomes were found in pairwise comparisons, $ps > .8$. In addition, there was a main effect of tempo: More successful trials occurred at the slow tempo than at the fast tempo, $F(1, 10) = 18.7$, $p < .01$, $\eta_p^2 = .65$. The Metronome \times Tempo interaction was significant, $F(3, 30) = 17.8$, $p < .01$, $\eta_p^2 = .32$, indicating that the fast tempo was more debilitating for the Flash than for the spatial metronomes.

Participants could not consistently tap with the Flash metronome in the fast (400-ms IOI) tempo condition and were not better than a 50% “synchronization threshold,” consistent with Repp (2003). However, success rates for the three spatial metronomes were well above the 50% synchronization threshold even in the fast tempo condition. This indicates that visuomotor synchronization rate limits for spatially varying metronomes are lower than the previously measured rate limits for flashing metronomes.

Analyses on the serial dependencies between tap intervals showed similar effects of metronome type, $F(3, 30) = 14.5$, $p < .001$, but no effect of tempo ($p > .8$). Lag-1 autocorrelations were significantly lower for the moving metronomes than the Flash metronome ($ps < .01$), and no differences between the moving metronomes were observed in pairwise comparisons. The lag-1 autocorrelations were significantly less than zero for the Up-Down Bar ($r = -.13$), the Rotating Bar ($r = -.16$), and the Finger metronome ($r = -.14$), $ps < .01$, but not for the Flash metronome ($r = -.01$, $p > .4$). This nonnegative autocorrelation in the Flash condition does not simply reflect lower synchronization success rates; autocorrelations in the Flash condition on only successful trials did not differ from zero ($r = -.02$, $p > .5$). Because the motor implementation variance can be assumed to be consistent between conditions, the negative autocorrelations observed for the moving metronomes suggest that error correction is occurring in order to mitigate phase drift and ensure synchrony with the target.

Analyses of tap timing on successful trials showed that taps tended to precede targets, as is commonly found (e.g., Aschersle-

ben, 2002), although mean relative phase or asynchrony is generally not treated as an index of synchronization success. An ANOVA revealed a main effect of tempo: Participants anticipated the target to a greater extent at the slow tempo, $F(1, 20) = 9.0$, $p = .007$. On average, participants tapped 42 ms before the target at the slow 500-ms tempo (mean direction = 328.2°), and 2 ms before the target at the fast 400-ms IOI tempo (mean direction = 358.3°). This decrease in anticipation tendency at faster tempi is common (e.g., Repp, 2003), but the reason for this remains unclear (see Repp, 2005, for a recent review of competing explanations). Metronome type yielded no effect overall, $F(3, 40) = 1.3$, $p > .2$, nor was there an effect of metronome type on mean relative phase at just the slower, more stable tempo ($p = .8$).

In summary, the addition of a spatial component facilitated visuomotor synchronization success, and participants easily synchronized with these metronomes even at the fast 400-ms IOI tempo. No advantage for tapping with biological forms over geometric forms was observed. The anticipation tendency was greater at the slow tempo.

Experiment 2

Experiment 2 investigated whether the moving visual metronomes’ degree of compatibility with the to-be-performed movement influences synchronization performance. Participants synchronized finger tapping with four types of visual metronome. Sequences were presented at three tempi: the two from Experiment 1, plus a faster tempo to examine whether synchronization might be possible at even higher rates.

Method

Participants. Thirteen right-handed Cornell students (eight women) between the ages of 19 and 21 years participated in the study. They were previously unfamiliar with the tapping task. Musical training ranged from 0–10 years ($M = 4.6$), although most were no longer active musicians; musical training did not

significantly correlate with tap timing or synchronization success. Participants received course credit or \$6.

Materials. The QuickTime videos were played at three IOIs: 500 ms (slow), 400 ms (fast), and 300 ms (very fast). The four metronome types are shown in Figure 3. The Flash and Finger videos from Experiment 1 were used. In the Finger video, stimuli and tapping movement were directionally compatible. The Inverted Finger was the Finger video rotated 180 degrees, and the target for tapping occurred when the finger contacted the top of the window; stimuli and tapping movement were incompatible. The left-and-right moving bar (*Side Bar*) was the Up–Down Bar from Experiment 1 rotated 90 degrees counterclockwise, so that the target occurred when the bar struck the viewer window’s right edge; stimuli were orthogonal and action-neutral with respect to the tapping movement. The three spatial videos had identical trajectories and tested the degree of compatibility between stimuli and movement.

Procedure. The procedure was identical to Experiment 1, except that there were 15 blocks containing each of the 12 trial types (3 tempi \times 4 metronome types) in random order. The first block was considered training, thus leaving 168 analyzed trials. The experiment lasted approximately 50 min.

Results and Discussion

The data were preprocessed as before. The percentage of successful trials by condition is shown in Figure 4. A 4 (metronome type) \times 3 (tempo) ANOVA found a main effect of metronome type, $F(3, 36) = 14.9$, $p < .001$, $\eta_p^2 = .55$. Pairwise comparisons showed synchronization performance was better for the compatibly moving Finger than all other metronomes ($ps \leq .001$). The neutrally moving Side Bar yielded better performance than the incompatible Inverted Finger and nonspatial Flash ($ps < .01$). There was no difference between the Flash and the Inverted Finger ($p > .25$). The main effect of tempo was also significant, $F(2, 24) = 50.3$, $p < .001$, $\eta_p^2 = .80$. Pairwise comparisons showed the most failed trials at the very fast (300-ms IOI) tempo ($ps < .001$) and more failed trials at fast (400-ms IOI) than slow (500-ms IOI; $p < .001$). The Metronome \times Tempo interaction was not significant ($p > .4$). Success rates for the compatible Finger metronome were over a 50% “synchronization threshold” even at the fastest 300-ms IOI.

Analyses of dependencies between adjacent tap intervals revealed an effect of metronome type, $F(3, 36) = 11.0$, $p < .001$, but no effect of tempo ($p > .2$). Pairwise comparisons revealed highest lag–1 autocorrelations for the Flash metronome ($ps < .05$), followed by the Inverted Finger ($ps < .05$), and no difference was observed between the Side Bar and the Finger metronomes ($p > .2$). Lag–1 autocorrelations were significantly less than zero for the Finger ($r = -.15$, $p < .001$), the Side Bar ($r = -.13$, $p < .001$), and the Inverted Finger ($r = -.10$, $p < .001$), but not for the Flash ($r = -.04$, $p > .2$) metronomes. Again, the significant negative autocorrelations are indicative of error correction mechanisms mitigating drift. This likelihood for adjacent taps to alternate between long and short taps was observed with the Inverted Finger, but not with the Flash, despite their similar success rates.⁵ This possibly reflects different synchronization processes: The Inverted Finger might show automatic (albeit poor) error correction, whereas the Flash performance might show more local drift.

Analyses of tap timing on successful trials showed that taps occurred earlier at the slower tempi. An ANOVA revealed a main

effect of tempo on mean relative phase, $F(2, 36) = 9.7$, $p < .001$. On average, taps occurred 48 ms before the target in the slow 500-ms IOI condition, 8 ms after the target in the fast 400-ms IOI condition, and 53 ms after the target in the very fast 300-ms IOI condition (converted from mean relative phase). There was no effect of type of metronome on mean relative phase overall, $F(3, 48) = 0.7$, $p > .7$, nor was there an effect of metronome at just the slowest, most stable tempo ($p > .5$).

Experiment 2 demonstrated that the addition of spatial information on its own does not necessarily improve synchronization performance over flashing metronomes, as indicated by the equally poor synchronization for the incompatibly moving Inverted Finger and nonspatial Flash. Incompatible movement direction could have an inhibitory effect that mitigates a spatial facilitation. Synchronization was greatly improved with spatial stimuli moving orthogonally to the to-be-produced movement, and compatible motion further facilitated synchronization performance. An anticipation tendency was evident only at the slowest tempo.

Experiment 3

In light of the facilitation observed in the first two experiments, Experiment 3 probed whether the maximum rate of visuomotor synchronization with compatibly moving targets could approach that of auditory–motor synchronization. Participants synchronized finger tapping with four types of metronomes: (a) auditory beeps; (b) an up-down moving bar compatible with finger movements; (c) a Flash metronome; and (d) a Flash target interspersed with predictably appearing and disappearing “snowflakes,” dubbed *Fade*. The Fade metronome controlled for the possibility that the facilitation observed in moving metronomes stemmed from their continuous and predictable nature rather than their spatial translation (moving metronomes and the Fade metronome have a trackable cycle and, in a sense, provide more warning of their impending impulse point than does the Flash condition). Sequences were presented at even faster tempi: Target IOIs were 400 ms, 300 ms, and 240 ms.

Method

Participants. Ten Cornell students (three women) between the ages of 19 and 32 years participated (including author MJH). Musical training ranged from 0 to 20 years ($M = 6.5$ years) and produced no significant results. Participants received \$4.

Materials. Stimulus sequences were presented with MATLAB’s PsychToolbox running on a 2.4 GHz MacBook Pro (NVIDIA GeForce 9600 video card) with its lid closed and driving an external CRT monitor at a refresh rate of 100 Hz (10 ms). PsychToolbox syncs visual presentation to the refresh rate allowing millisecond-accurate timing for visual as well as auditory stimuli. All sequences lasted 26 cycles. Visual metronomes at the 400-ms and 300-ms target IOIs consisted of 10 images per cycle, with each image lasting 40 ms and 30 ms, respectively; the 240-ms IOI sequences consisted of six images/cycle, with each image presented for 40 ms. Sequences are depicted in Figure 5.

⁵ When examining only the successful trials, the autocorrelations for Flash similarly do not significantly differ from zero ($r = -.03$, $p > .2$). Overall, autocorrelations did not statistically differ between all trials and successful trials only ($p > .5$).

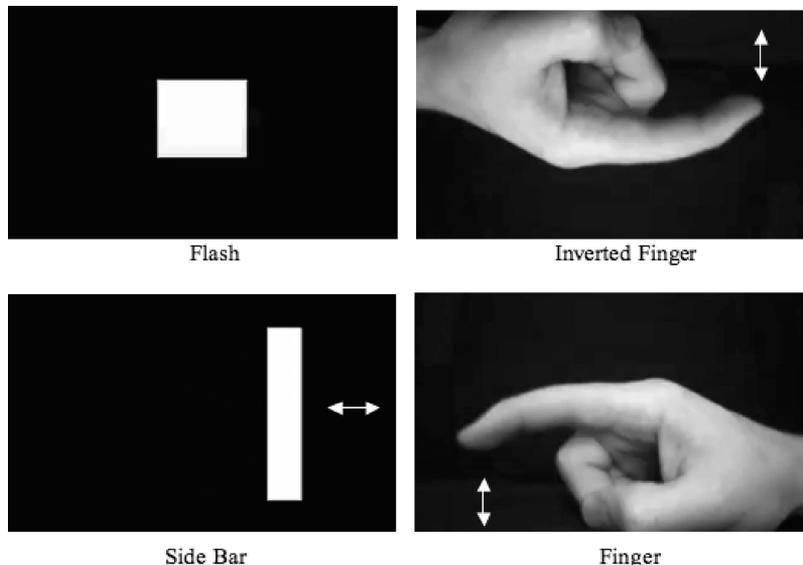


Figure 3. Stills from the four metronomes in Experiment 2; arrows added to depict motion.

Auditory sequences consisted of 40-ms long sine wave beeps at 440 Hz. The up-down Moving Bar was a 5 × 1 cm bar, with a 2.5-cm maximum displacement above the bottom of the screen for the 400- and 300-ms IOI trials and a 1.5-cm maximum displacement for the 240-ms IOI trials. The Flash was a 2 cm × 2 cm white square. The Fade metronome consisted of the Flash target interspersed with dots; it started with a blank screen, then added dots in each successive image until the Flash onset, then subtracted dots until the blank (i.e., blank, 5 dots, 10 dots, 15 dots, 20 dots, FLASH, 20 dots, 15 dots, 10 dots, 5 dots, etc., for 400- and 300-ms IOI trials; and blank, 10 dots, 20 dots, FLASH, 20 dots, 10 dots, etc., for 240-ms IOI trials).

Procedure. Participants sat in front of a computer monitor wearing Beyer Dynamic circumaural headphones and tapped on a light cardboard box fitted with a microphone. On a separate Mac

G4 computer running Audacity at an 8000-Hz sample rate, taps were recorded on one channel, and trial onset markers from the stimulus computer were recorded on the other channel.

All 12 trial types (3 tempi × 4 metronome types) were presented in random order in a block. The experiment consisted of one practice block and 10 experimental blocks (120 analyzed trials). The experiment lasted approximately 25 min.

Results and Discussion

Tapping data were analyzed using circular statistical methods and the taps occurring during the first four targets of each trial were omitted from analyses. The percentage of successful trials by condition is shown in Figure 6. A 4 (metronome type) × 3 (tempo) ANOVA found a main effect of metronome type, $F(3, 27) = 39.0$,

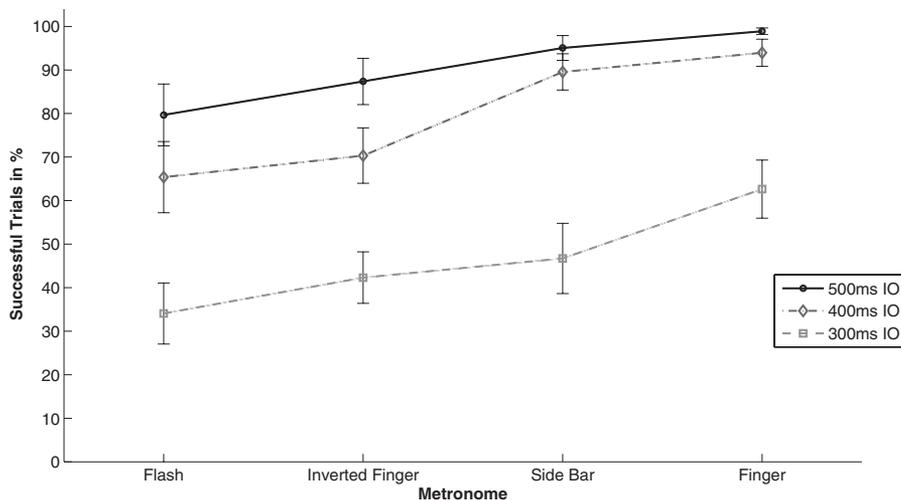


Figure 4. Percentage of successful trials in Experiment 2, as determined by a Rayleigh test of uniformity.

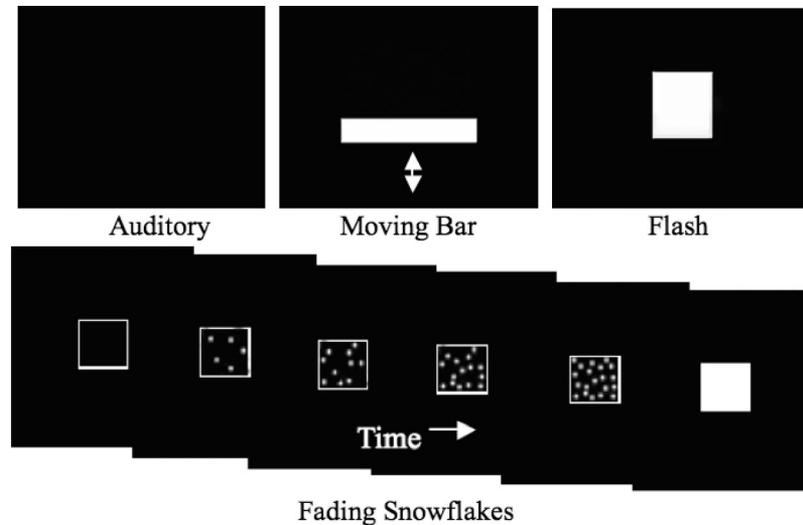


Figure 5. The four metronomes from Experiment 3.

$p < .001$, $\eta_p^2 = .81$. Pairwise comparisons showed that the synchronization performance with the auditory metronomes was better than that with each visual metronome ($ps < .01$). Among the visual metronomes, the Moving Bar yielded better synchronization performance than the Flash or the Fade ($ps < .05$), and no difference was observed between the Flash and the Fade ($p > .1$). Tempo also affected synchronization performance, $F(2, 18) = 23.2$, $p < .001$, $\eta_p^2 = .72$, with pairwise comparisons showing that performance was worse at faster tempi ($ps < .05$). The Tempo \times Metronome interaction, $F(6, 54) = 3.5$, $p < .01$, $\eta_p^2 = .28$, indicates that the fast tempi impede synchronization more for visual than auditory metronomes.

Analyses of the lag-1 autocorrelations revealed an effect of metronome, $F(3, 27) = 10.5$, $p < .001$, but no effect of tempo ($p > .2$). Pairwise comparisons showed that the autocorrelations were lower for the auditory and Moving Bar metronomes than those for the Fade and Flash metronomes ($ps < .01$). The autocorrelations were significantly greater than zero for the Flash metronomes ($r = .11$, $p < .01$) and Fade ($r = .09$, $p < .05$), but did not differ from zero for the Moving Bar ($r = .00$) or the auditory metronomes ($r = -.05$). Two components drive lag-1 autocorrelation values in opposite directions: Error correction decreases values, whereas drift increases values. The overall shift toward positive values indicates more drift in Experiment 3 compared with the other experiments.⁶ The observed shift could stem from a combination of faster tempi, lower overall synchronization success, fewer trials per condition, or different participants; however, the exact reason remains elusive. The positive shift is not simply an artifact of more failed trials, as the autocorrelations did not significantly differ between all trials and only successful trials, $F(1, 9) = 3.0$, $p > .1$. In addition, when examining only the slowest, most stable tempo, autocorrelation values remain positive for the Fade ($r = .08$, $p < .05$) and the Flash ($r = .11$, $p < .05$) metronomes, are negative for the auditory metronome ($r = -.10$, $p < .05$, one-tailed), and did not differ from zero for the Moving Bar ($r = .01$, $p > .7$). These differences between conditions for lag-1 autocorrelation suggest different underlying processes.

Taps tended to precede the targets, especially in the slowest tempo. On average, taps occurred 37 ms before the target in the 400-ms IOI condition, 10 ms after the target in the 300-ms IOI condition, and 1 ms after the target in the 240-ms IOI condition (converted from mean relative phase). The circular ANOVA on tempo did not attain significance because of the highly variable results at the fastest tempo, $F(2, 27) = 2.10$, $p = .14$. However, the difference between the 400-ms and 300-ms tempi was significant, $F(1, 18) = 6.21$, $p = .02$. There was no effect of metronome type on mean relative phase overall, $F(3, 36) = 1.27$, $p > .25$, nor was there an effect of metronome at just the slowest, most stable tempo ($p > .8$).

Experiment 3 demonstrates that synchronization performance is more stable in the auditory domain, even compared with compatibly moving visual stimuli. The Fade metronome did not improve synchronization over the Flash; thus, we can confidently conclude that the facilitation with the Moving Bar is due to its compatible *spatial* component rather than its predictability.

General Discussion

These results demonstrate that compatible spatial information greatly facilitates visuomotor synchronization. In Experiment 1, participants' synchronization performance was dramatically better with the three moving metronomes (similar in amplitude, size, and trajectory) than the flashing metronome. No unique advantage was observed for synchronization with biological forms in particular, despite the representational overlap for seeing bodily forms and for producing movements in the extrastriate body area. Our stimuli did not contain biological motion, but future work should explore synchronization with metronomes containing biological trajectory-

⁶ Supplementary analyses show positive lag-2 autocorrelations for the Flash ($r = .10$, $p < .01$) and Fade ($r = .11$, $p < .01$); this further supports drift rather than second-order error correction based on the asynchrony two cycles before the tap (cf. Pressing, 1998).

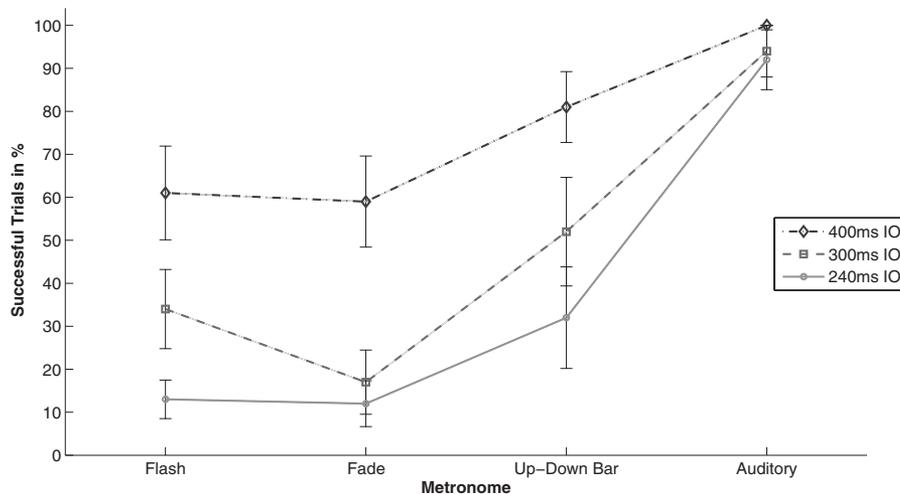


Figure 6. Percentage of successful trials in Experiment 3, as determined by a Rayleigh test of uniformity.

ries of acceleration and deceleration. Experiment 2 demonstrated the importance of directional compatibility between the metronome and body movement. Synchronization performance was equally poor for the action-incompatible Inverted Finger and the Flash. The directional mismatch between target stimulus (upward) and tap response (downward) presumably caused interference, thereby negating the spatial facilitation. Performance improved when tapping with an orthogonal (i.e., noninterfering and noncompatible) sideways moving bar, and best performance occurred with the highly compatible (in direction, amplitude, and size) Finger video.

In Experiment 3 (and much additional pilot work employing stimuli that continuously faded into color bursts), no advantage was observed for a predictable Fade metronome. Thus, rather than predictability, the spatial component appears to be the crucial facilitative factor, probably due to the visual system's proficiency at processing spatial information.

Analyses of serial dependencies between adjacent taps revealed differences between metronomes. Lag-1 autocorrelations for the Flash and Fade metronomes were consistently higher than for the moving metronomes. Positive autocorrelations, characterized by successive long or successive short intervals, are consistent with drift (Madison, 2001) and have been reported with flashing visual stimuli (Chen et al., 2002). For the Flash metronome, we observed positive lag-1 autocorrelation values even for successful trials (which had a unimodal distribution of taps); this suggests non-monotonic drift, wherein taps drift back and forth around the target rather than phase correction leading to a negative dependence between successive tap intervals. The negative autocorrelations with moving metronomes notably intimate different underlying processes and error correction in visuomotor synchronization. Previous neuroimaging work has investigated visuomotor synchronization with flashing stimuli, and future work may uncover different neural regions involved in visuomotor synchronization with moving metronomes.

Rough estimates of a "synchronization threshold" (50% success rates; Repp, 2003) for compatibly moving metronomes in Experiments 2 and 3 converge around 300-ms IOI for the untrained,

novice-tapper participants. Two participants in Experiment 3 easily synchronized at the fastest tempo (i.e., 240-ms IOI); future work could examine effects of training on this uncommon task or examine performance of "hand-eye experts" such as athletes or video gamers. Although this rough threshold is much lower than previously established, the advantage for synchronization in the auditory modality remains, as participants could successfully synchronize with the auditory stimuli, even at the 240-ms IOI tempo.

The compatibility effects support the notion that temporal information for action is not coded in an independent, action-neutral domain (i.e., in a specialized, isolated timing center that extracts temporal information from perceptual systems regardless of orientation, then sends this on to an isolated action planning system). Rather, temporal information here appears to be computed in a task-dependent, action-oriented manner (Ivry & Spencer, 2004). In addition, the compatibility effects between perceived and produced events can be explained in terms of the common-coding theory of perception and action (Hommel, Müssele, Aschersleben, & Prinz, 2001). In this theory, sensory and motor codes share a common representational medium; sensory information of a downward-moving target converges on a shared abstract feature code, which spreads activation to the motor system, prespecifying or biasing it toward downward action. Conversely, an upward-moving sensory target will bias action toward upward movement and interfere with the downward goal.

Theoretical accounts of motor control that employ forward models suggest that motor commands to the muscles are accompanied by an efference copy that predicts sensory consequences of movement (Blakemore, Wolpert, & Frith, 2002; Wolpert & Kawato, 1998). Discrepancies between this prediction and the actual sensory feedback (both direct somatosensory and visual or auditory information indirectly resulting from action) allow rapid adjustment to the movement. In our experiment, the incompatible Inverted Finger stimuli yielded a constant lack of correspondence between visual information and efference copy prediction and somatosensory feedback, which would lead to increased adjustment and hence higher variability and difficulty. Conversely, a tight correspondence between these signals not only evinces good

performance, but also gives us a sense of agency (Farrer & Frith, 2002). When tapping with the compatible Finger video, many participants responded in posttest questioning that they felt that the video was *their* hand or that they were controlling the video. Similar impressions can also occur with auditory signals (Knoblich & Repp, 2009; Repp & Knoblich, 2007). Such a misattribution of ownership or agency might stem from the close correspondence between visual, somatosensory, and motor prediction signals (see also Knoblich & Sebanz, 2005).

To sum, our results demonstrate that visuomotor synchronization performance is greatly facilitated by compatible motion, possibly due to the visual system's proficiency at processing spatial information and the tight linkages between perceptual and action systems; however, an auditory advantage in sensorimotor synchronization persists.

References

- Aschersleben, G. (2002). Temporal control of movements in synchronization. *Brain and Cognition*, *48*, 66–79.
- Astafiev, S. V., Stanley, C. M., Shulman, G. L., & Corbetta, M. (2004). Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nature Neuroscience*, *7*, 542–548.
- Berens, P. (2009). CircStat: A MATLAB toolbox for circular statistics. *Journal of Statistical Software*, *31*, 1–21.
- Bertelson, P., & Radeau, M. (1981). Cross-modal bias and perceptual fusion with auditory–visual spatial discordance. *Perception & Psychophysics*, *29*, 578–584.
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*, 237–242.
- Bogaerts, H., Buekers, M. J., Zaaf, F. T., & Swinnen, S. P. (2003). When visuo-motor incongruence aids motor performance: The effect of perceiving motion structures during transformed visual feedback on bimanual coordination. *Behavioural Brain Research*, *138*, 45–57.
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, *106*, 3–22.
- Chen, Y., Ding, M., & Kelso, J. A. S. (2003). Task-related power and coherence changes in neuromagnetic activity during visuomotor coordination. *Experimental Brain Research*, *148*, 105–116.
- Chen, Y., 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*, 515–532.
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 24–39.
- Delignières, D., Torre, K., & Lemoine, L. (2008). Fractal models for event-based and dynamical timers. *Acta Psychologica*, *127*, 382–397.
- Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001, September 28). A cortical area selective for visual processing of the human body. *Science*, *293*, 2470–2473.
- Dunlap, K. (1910). Reactions to rhythmic stimuli, with attempt to synchronize. *Psychological Review*, *17*, 399–416.
- Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of preschoolers' corporeal synchronization with music. In M. Baroni, A. Addressi, R. Caterina, & M. Costa (Eds.), *Proceedings of the Ninth International Conference on Music Perception and Cognition* (pp. 472–476). Bologna, Italy.
- Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs. another person as being the cause of an action: The neural correlates of the experience of agency. *NeuroImage*, *15*, 596–603.
- Fisher, N. I. (1993). *Statistical analysis of circular data*. Cambridge, England: Cambridge University Press.
- Fraisse, P. (1948). Rythmes auditifs et rythmes visuels [Auditory and visual rhythms]. *L'Année Psychologique*, *49*, 21–41.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The psychology of music* (pp. 149–180). Orlando, FL: Academic Press.
- Hary, D., & Moore, G. P. (1985). Temporal tracking and synchronization strategies. *Human Neurobiology*, *4*, 73–77.
- Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, *24*, 849–878.
- Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology*, *14*, 225–232.
- Jäncke, L., Loose, R., Lutz, K., Specht, K., & Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cognitive Brain Research*, *10*, 51–66.
- Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology*, *102*, 299–314.
- Knoblich, G., & Repp, B. H. (2009). Inferring agency from sound. *Cognition*, *111*, 248–262.
- Knoblich, G., & Sebanz, N. (2005). Agency in the face of error. *Trends in Cognitive Sciences*, *9*, 259–261.
- Kolers, P. A., & Brewster, J. M. (1985). Rhythms and responses. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 150–167.
- Large, E. W., & Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, *26*, 1–37.
- Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 411–422.
- Madison, G., & Merker, B. (2004). Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neuroscience Letters*, *370*, 69–73.
- Mates, J. (1994). A model of synchronization of motor acts to a stimulus sequence: I. Timing and error corrections. *Biological Cybernetics*, *70*, 463–473.
- Mechner, F., Kerzel, D., Knoblich, G., & Prinz, W. (2001). Perceptual basis of bimanual coordination. *Nature*, *414*, 69–73.
- Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricity and modality on synchronization with a beat. *Experimental Brain Research*, *163*, 226–238.
- Penhune, V. B., Zatorre, R. J., & Evans, A. C. (1998). Cerebellar contributions to motor timing: A PET study of auditory and visual rhythm reproduction. *Journal of Cognitive Neuroscience*, *10*, 752–765.
- Posner, M. I., Nissen, M. J., & Klein, R. M. (1976). Visual dominance: An information-processing account of its origins and significance. *Psychological Review*, *83*, 157–71.
- Pressing, J. (1998). Error correction processes in temporal pattern production. *Journal of Mathematical Psychology*, *42*, 63–101.
- Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, *35*, 355–370.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.
- Repp, B. H., & Knoblich, G. (2007). Toward a psychophysics of agency: Detecting gain and loss of control over auditory action effects. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 469–482.
- Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1085–1099.

- Repp, B. H., & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research, 68*, 252–270.
- Roerdink, M., Peper, C. E., & Beek, P. J. (2005). Effects of correct and transformed visual feedback on rhythmic visuo-motor tracking: Tracking performance and visual search behavior. *Human Movement Science, 24*, 379–402.
- Semjen, A., Schulze, H.-H., & Vorberg, D. (2000). Timing precision in continuation and synchronization tapping. *Psychological Research, 63*, 137–147.
- Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception, 18*, 455–489.
- Spencer, R. M. C., Zelaznik, H. N., Diedrichsen, J., & Ivry, R. B. (2003, May 30). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science, 300*, 1437–1439.
- Thaut, M. H., Kenyon, G. P., Schauer, M. L., & McIntosh, G. C. (1999). The connection between rhythmicity and brain function. *IEEE Engineering in Medicine and Biology, 18*, 101–108.
- Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research, 199*, 157–166.
- van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28*, 43–66.
- Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of perception and action* (Vol. 2, pp. 181–262). London: Academic Press.
- Watson, G. S., & Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika, 43*, 344–352.
- Wimmers, R. H., Beek, P. J., & van Wieringen, P. C. W. (1992). Phase transitions in rhythmic tracking movements: A case of unilateral coupling. *Human Movement Science, 11*, 217–226.
- Wing, A. M., & Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Perception & Psychophysics, 14*, 5–12.
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks, 11*, 1317–1329.
- Zelaznik, H. N., Spencer, R. M., & Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology: Human Perception and Performance, 28*, 575–588.

Received April 17, 2009

Revision received November 5, 2009

Accepted December 4, 2009 ■