



Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness

Michael J. Hove^{a,*}, Merle T. Fairhurst^a, Sonja A. Kotz^a, Peter E. Keller^{a,b}

^a Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^b MARCS Institute, University of Western Sydney, Australia

ARTICLE INFO

Article history:

Accepted 26 November 2012

Available online 30 November 2012

Keywords:

Timing

Rhythm

Modality differences

Sensorimotor synchronization

Basal ganglia

ABSTRACT

Synchronizing movements with auditory beats, compared to visual flashes, yields divergent activation in timing-related brain areas as well as more stable tapping synchronization. The differences in timing-related brain activation could reflect differences in tapping synchronization stability, rather than differences between modality (i.e., audio-motor vs. visuo-motor integration). In the current fMRI study, participants synchronized their finger taps with four types of visual and auditory pacing sequences: *flashes* and a *moving bar*, as well as *beeps* and a *frequency-modulated 'siren'*. Behavioral tapping results showed that visuo-motor synchronization improved with moving targets, whereas audio-motor synchronization degraded with frequency-modulated sirens. Consequently, a modality difference in synchronization occurred between the discrete beeps and flashes, but not between the novel continuous siren and moving bar. Imaging results showed that activation in the putamen, a key timing area, paralleled the behavioral results: putamen activation was highest for beeps, intermediate for the continuous siren and moving bar, and was lowest for the flashes. Putamen activation differed between modalities for beeps and flashes, but not for the novel moving bar and siren. By dissociating synchronization performance from modality, we show that activation in the basal ganglia is associated with sensorimotor synchronization stability rather than modality-specificity in this task. Synchronization stability is apparently contingent upon the modality's processing affinity: discrete auditory and moving visual signals are modality appropriate, and can be encoded reliably for integration with the motor system.

© 2012 Elsevier Inc. All rights reserved.

Introduction

Precise temporal coordination between action and different perceptual systems is crucial for interacting with a dynamic environment. Precise visuo-motor integration is needed to catch a ball (or catch dinner) and audio-motor integration is needed to synchronize movements with music. Empirically, the temporal integration of action and perception is commonly examined in tasks requiring finger tapping to an isochronous pacing sequence. Previous neuroimaging and behavioral studies of tapping have established strong modality differences between audio-motor and visuo-motor synchronization. However, the vast majority of previous studies used only *flashing* visual stimuli, and flashes are known to yield poorer synchronization performance than auditory stimuli (Repp, 2005). Visuo-motor synchronization improves significantly with moving stimuli (Hove and Keller, 2010; Hove et al., 2010; Iversen et al., submitted for publication). In the present study, we tested whether previously observed activation differences reflect modality

per se (i.e., audio-motor vs. visuo-motor integration), or differences in synchronization performance.

Neuroimaging studies have uncovered divergent neural activation patterns for visuo-motor versus audio-motor synchronization. Differences extend well beyond primary sensory areas into regions implicated in the brain's timing networks, including the basal ganglia, supplementary motor areas (SMA), and cerebellum (e.g., Buhusi and Meck, 2005; Coull et al., 2011; Macar et al., 2002; Schwartz et al., 2012). Direct comparisons of audio-motor and visuo-motor synchronization reported activation in different areas of the cerebellum (Jäncke et al., 2000; Penhune et al., 1998). Additionally, audio-motor, but not visuo-motor synchronization, yielded significant activation in the SMA (Jäncke et al., 2000; Penhune et al., 1998). In a meta-analysis on 38 neuroimaging studies of finger-tapping, striking differences between audio- and visuo-motor synchronization were uncovered in the putamen of the basal ganglia; synchronization with auditory, but not visual stimuli, consistently activated the putamen (Witt et al., 2008). The putamen is a key area for beat and rhythm processing (Coull et al., 2011; Grahn and Rowe, 2009; Kotz et al., 2009; Teki et al., 2011; Wiener et al., 2009). A recent study comparing audio and visual beat perception showed more putamen activation and more sensitive beat perception for auditory than for visual stimuli; nevertheless, within the visual condition the degree of putamen activation predicted beat sensitivity (Grahn et al., 2011).

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Research Group "Music Cognition and Action", Stephanstrasse 1a, 04103 Leipzig, Germany. Fax: +49 341 9940 2614.

E-mail address: michaeljhove@gmail.com (M.J. Hove).

Taken together, neural activation differences observed between auditory and visual modalities in synchronization and beat perception have important implications. For example, it has been argued that auditory rhythms induce an internal rhythm that guides movement, whereas visual rhythms do not generate an internal rhythm (Jäncke et al., 2000). Additionally, modality differences have provided evidence that time is represented in a distributed network rooted in sensorimotor processes, rather than subserved by a centralized clock mechanism (Jantzen et al., 2005). Furthermore, differences in neural activations could support an auditory specialization for encoding temporal information (e.g., Welch and Warren, 1980).

Given these activation differences in timing circuits, it is perhaps unsurprising that a strong behavioral advantage has also been observed for audio-motor over visuo-motor synchronization. Rhythmic finger tapping is much more accurate with auditory stimuli than with flashing visual stimuli (e.g., Chen et al., 2002; Dunlap, 1910; Kolers and Brewster, 1985). Stable synchronization is possible at much faster rates with auditory than with visual sequences (Repp, 2003). In a target-distracter paradigm, when auditory beeps and visual flashes are presented in competition with each other, participants' movement timing is dictated by the auditory stimuli, regardless of volition (Repp and Penel, 2004). Finally, the serial dependence between inter-tap intervals, which intimates underlying timing processes (e.g., Vorberg and Wing, 1996), differs between audio- and visuo-motor synchronization: Inter-tap intervals in audio-motor synchronization typically alternate between short and long intervals (a negative lag1 autocorrelation), which suggests active error correction (Semjen et al., 2000); whereas synchronization with flashing visual stimuli typically has a positive or non-negative lag1 autocorrelation, which suggests weak (or absent) tap-to-tap error correction (Chen et al., 2002; Hove and Keller, 2010; Hove et al., 2010). Together these results suggest different underlying processes for synchronizing with audio- versus flashing visual sequences.

However, nearly all imaging and behavioral evidence for differences between visual and auditory synchronization used *flashing* visual stimuli. While flashes may offer the most similar control for auditory beeps in terms of temporal onset/offset (and no additional confounding factors), they lack ecological validity in that the visual system rarely processes or acts upon purely temporal information devoid of spatial translation. The visual system has considerably lower temporal resolution than audition (e.g., Holcombe, 2009), and thus is severely handicapped in synchronizing with discrete temporal stimuli. Vision excels at processing spatial, rather than temporal information (e.g., Bertelson and Radeau, 1981; Welch and Warren, 1980). When a visual stimulus contains spatiotemporal information (rather than purely temporal information), action timing to intercept that moving stimulus can be very precise (Bootsma and van Wieringen, 1990).

In a series of recent finger-tapping studies, we have shown that synchronization timing improves dramatically with spatiotemporal visual stimuli, compared with purely temporal flashing stimuli. In one study, participants tapped along with flashing visual stimuli and with visual images that alternated between a high and low position creating apparent motion. Synchronization was considerably more stable with the apparent motion stimuli than the flashes (Hove and Keller, 2010). In another study, participants tapped along with visual flashes, fading stimuli, and stimuli that moved frame-by-frame at a linear velocity, as well as an auditory metronome. Synchronization with the moving stimuli was much better than with flashing or fading stimuli; however, an auditory advantage was still observed, especially at very fast tempi (300 and 240 ms IOI), (Hove et al., 2010). In both these studies, the moving visual stimuli also yielded *negative* lag1 autocorrelations, suggesting that error correction was occurring. Additionally, in a target-distracter study that presented moving visual stimuli in competition with auditory beeps, the moving visual stimuli attracted movement timing as much as auditory stimuli, thus erasing the auditory dominance previously observed over flashes (Hove et al., *in press*). Together, these studies demonstrate that motion increases

the temporal reliability of visual encoding (cf. Ernst and Bühlhoff, 2004) and thus facilitates precise visuo-motor integration.

The foregoing suggests that the 'modality differences' observed between auditory and flashing visual stimuli should be interpreted cautiously: It is unclear if the previously reported differences in brain activation truly reflect differences between *modality* (i.e., audio-motor vs. visuo-motor integration) or simply the poor performance with flashing visual stimuli due to their less precise temporal encoding. The significant improvement in visual synchronization with moving stimuli encourages the re-examination of established modality differences. Are differences in neural timing circuits substantially reduced with improved visuo-motor synchronization (or degraded audio-motor synchronization)?

In the present fMRI study, participants synchronized their finger taps with four types of visual and auditory pacing sequences: visual stimuli were *flashes* and a *moving bar*, and auditory stimuli were *beeps* and *frequency modulated 'sirens'*. Within the visual modality, synchronization should improve for the moving bar stimuli compared to the flashes due to the spatial processing advantage. Within the auditory modality, synchronization should degrade for the siren compared to the discrete beeps, since the siren's continuous presentation should reduce or blur the neural encoding of its target compared to the beep's discrete target (cf. Barsz et al., 2002). Thus, these stimuli can disentangle synchronization performance from modality. Critically, neural activation in key timing areas such as the putamen should vary with the stability of synchronization performance, rather than being dictated by modality. We expect to replicate previously observed modality differences only for the discrete stimuli (beeps versus flashes), whereas the modality differences should be substantially less pronounced with the continuous stimuli (siren versus moving bar). Accordingly, we anticipate an interaction between modality and discrete/continuous stimulus structure for both behavioral synchronization performance and neural activation in time-sensitive areas such as the putamen.

Material and methods

Participants

Fourteen right-handed volunteers (7 women) aged 24 to 34 years ($M = 27.7 \pm 3.0$ years) participated in the experiment.¹ Participants were paid for their participation and gave informed consent. Participants had a range of musical training ($M = 7.9$ years; $SD = 9.5$); this did not affect tapping performance ($ps > .5$).

2.2. Experimental protocol

Participants lay supine in the fMRI scanner and tapped their right index finger on hard plastic surface embedded in a custom MR-compatible air-pressure response device. Participants were instructed to synchronize their finger taps along with four different types of isochronous pacing sequences at two tempi. The study employed a 2 (modality: auditory, visual) \times 2 (style: discrete, continuous) \times 2 (tempo: slow 600 ms IOI, fast 400 ms IOI) within-subjects design (see Fig. 1). Including two tempi ensured that participants attended to the stimuli and did not simply tap along with one memorized tempo. Auditory pacing sequences were presented over headphones (MR Confon, Magdeburg, Germany) and visual pacing sequences were presented via a projector (SANYO PLC-XP50L). Taps were recorded and pacing sequences were presented via a PC running Presentation software (Neurobehavioral Systems). The presentation program syncs to the refresh rate and allows consistent timing in

¹ Two additional volunteers participated in the experiment, but failed to synchronize with the pacing sequences in more than half of the trials in the scanner; therefore they were excluded and their imaging data were not analyzed.

the visual and auditory modalities; this was corroborated with a photo-diode test.

The *auditory-discrete* sequences consisted of “beeps” (50 ms duration at 1350 Hz) every 400 or 600 ms. The *auditory-continuous* sequences (“siren”) consisted of frequency modulated pitch sweeps that started at the peak 1350 Hz and decreased in 150 Hz increments every 50 ms, then increased back to the peak. Step-wise frequency modulation was employed to match the frame-by-frame presentation in the visual condition. The slow condition consisted of seven frequency steps from 1350 to 450 Hz and back, and the fast condition consisted of five frequency steps from 1350 to 750 Hz and back. Participants were instructed to tap with the peak frequency of the pitch sweep; peaks are more salient and previous work with pitch sweeps established that participants naturally synchronize with the peaks (McAnally, 2002).

The *visual-discrete* sequences (“flash”) consisted of a white bar flashed for 50 ms on a black background every 400 or 600 ms. The *visual-continuous* sequences (“moving bar”) consisted of a bar moving up and down. The bar started at the bottom of the screen and moved 0.8 cm every 50 ms; the slow condition consisted of seven steps in amplitude, and the fast condition consisted of five steps in amplitude. Participants were instructed to tap when the bar hit the bottom of its trajectory; thus the finger and visual stimulus moved in a directionally compatible manner. During auditory trials, the stationary bar remained on screen in a fixed-position, essentially serving as a fixation point. All trials lasted 19.2 s; Fast trials (400 ms IOI) contained 48 cycles, and slow trials (600 ms IOI) contained 32 cycles. Participants completed a short practice session prior to scanning.

The entire experiment consisted of 10 blocks containing each of the 8 conditions in permuted order, for a total of 80 trials (split into two separate fMRI runs). Trials were initiated automatically by the computer and separated by a variable inter-trial-interval of 9.4–12.2 s. Participants were in the scanner for approximately 40 min.

Imaging data acquisition and analysis

Functional imaging data were collected on a 3 T Siemens Trio system using continuous sampling. Scans used echo planar imaging (EPI) with a repetition time of 2 s; TE = 24 ms; 36 axial oblique slices; 1 mm gap; and voxel size = 3 × 3 × 3 mm³. All participants had previously participated in an MRI experiment, and their high resolution, T1 weighted structural scans (64 slices at 1 × 1 × 1 mm³ voxel size) were obtained from the database. Neuroimaging data were analyzed using FEAT (FMRIB Expert Analysis Tool) Version 5.63, part of FSL (FMRIB’s Software Library, www.fmrib.ox.ac.uk/fsl). Pre-statistic processing included: motion correction using MCFLIRT (Motion Correction FMRIB’s Linear Image Registration Tool; Jenkinson and Smith, 2001); non-brain removal using BET (Smith, 2002); spatial smoothing using a Gaussian Kernel of 5 mm full width at half-maximum; and non-linear high pass temporal filtering (Gaussian-weighted least-squares straight line fitting, with sigma = 60 s). Functional scans were co-registered

onto each individual’s T1 high-resolution structural image and then registered onto a standard brain (Montreal Neurological Institute MNI 152 brain).

Statistical analysis at the individual subject level was carried out using a general linear modeling (GLM) approach (Friston et al., 1994). Each stimulus condition was modeled as a unique explanatory variable (the boxcar functions lasted for the full trial duration of 19.2 s); the regressors were formed by convolving the boxcar function with a hemodynamic response function. Time-series statistical analysis used FILM (FMRIB’s Improved Linear Model) with local autocorrelation correction (Woolrich et al., 2001). First-level contrasts directly compared the four conditions of auditory- and visually-paced synchronization (with collapsed tempi), as well as against implicit baseline. Second-level analysis grouped the first-level analysis of each subject’s two scanning blocks. Group statistics analyses used FEAT, employing FLAME (FMRIB’s Local Analysis of Mixed Effects). This analysis method incorporates the variance within session and across time (fixed effects), as well as cross-session variances (random effects). Cluster thresholding was performed with a voxel-level Z-threshold of 2.3 and a corrected cluster p-value < 0.05 with a cluster-based correction for multiple comparisons using Gaussian Random Field Theory (Friston et al., 1994; Worsley et al., 1992).

As a follow-up to the whole-brain voxel-wise analyses, and based on previous work highlighting its importance in temporal processing (e.g., Grahn and Rowe, 2009), a region of interest (ROI) analysis was used to compare putamen activation for the four stimulus conditions. The ROI was functionally defined as an 8 mm sphere centered in the left putamen at x = −18, y = 14, z = −8. Time-course analyses and individual parameter estimates were extracted from this ROI using PEATE – Perl Event-related Average Time-course Extraction software (<http://www.jonaskaplan.com/peate/>). These parameter estimates were analyzed in an additional ANOVA.

Behavioral analyses

Tap timing data from the response pad were analyzed to assess synchronization performance for each trial. We first calculated the relative phase, a measure of the asynchrony between a tap and the target. Synchronization typically requires a few taps to stabilize; therefore, the first two seconds of each trial were omitted from analyses.

Tap-to-target synchronization stability was assessed in terms of R, the resultant length of tap-to-target relative phases on a unit circle. R indexes the stability of tap-to-target coordination on a scale from 0 (unstable tapping with relative phases distributed uniformly around the unit circle) to 1 (perfectly stable tapping with a unimodal distribution of relative phases). R equals (1 – circular variance), (see Fisher, 1993, for more information on circular statistical treatments). We report the lag1 autocorrelations of the inter-tap intervals in the supplementary materials.

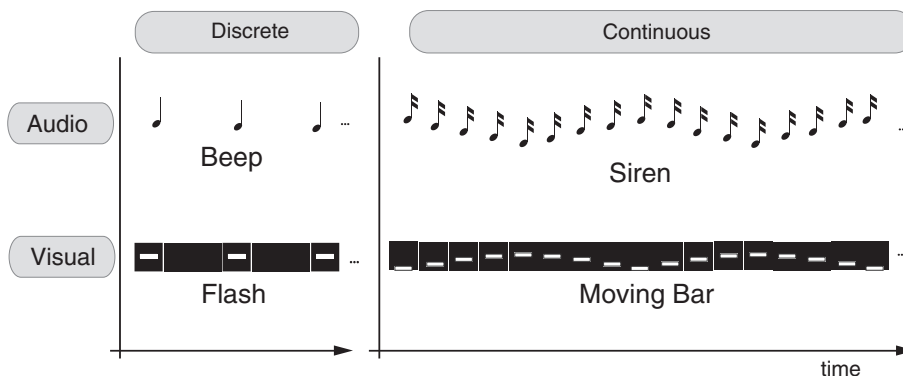


Fig. 1. Schematic of the four pacing sequence conditions.

Results

Behavioral results

Tap-to-target synchronization stability (R) was analyzed in a 2 (modality: auditory, visual) \times 2 (temporal structure: discrete, continuous) \times 2 (tempo: slow, fast) repeated measures Analysis of Variance (ANOVA); see Table 1. Overall, participants synchronized better with the auditory than the visual sequences, as indicated by a main effect of modality, $F(1,13) = 53.72$, $p < .001$, $\eta_p^2 = .805$. There was no main effect of discrete vs. continuous temporal structure, $F(1,13) = 2.04$, $p = .18$, $\eta_p^2 = .135$.

In the critical test, we observed a highly significant interaction of modality and temporal structure, $F(1,13) = 54.88$, $p < .001$, $\eta_p^2 = .808$; see Fig. 2A. Post-hoc comparisons (corrected) showed that auditory synchronization was more stable with discrete beeps than with the continuous siren ($p = .009$); Visual synchronization was more stable with the continuous moving bar than with the discrete flashes ($p < .001$). As previously established, a large modality advantage was observed for discrete beeps over flashes ($p < .001$). The discrete beeps also outperformed the moving bar ($p < .01$). However with the novel continuous stimuli, synchronization stability did not significantly differ between modalities for the siren and moving bar ($p = .093$).

Tap-to-target synchronization was more stable at the slow tempo as indicated by a main effect of tempo, $F(1,13) = 75.88$, $p < .001$, $\eta_p^2 = .854$. Tempo interacted with modality, $F(1,13) = 30.43$, $p < .001$, $\eta_p^2 = .701$, indicating that the fast tempo affected synchronization stability more for the visual stimuli. Synchronization deteriorated at the fast tempo for all four conditions as indicated by separate paired t -tests (all $ps < .05$), signifying that tempo effects are more quantitative than qualitative. The three-way interaction was not significant ($p > .1$). Thus, the primary fMRI analyses will investigate modality differences for discrete and continuous stimuli with collapsed tempi, which was the main experimental motivation. Tempo effects are reported in supplementary materials.

Imaging results

Synchronizing finger tapping with pacing sequences across conditions (relative to baseline) activated regions similar to other finger-tapping studies including sensorimotor cortices, premotor cortex, R cerebellum lobule V, and L putamen.

Between modality contrasts

First we examined modality differences for the discrete stimuli, as have been examined in previous studies: auditory beeps and visual flashes. In the (beep > flash) contrast, tapping with beeps yielded significantly more activation in the left putamen, in addition to the auditory cortex, and mid-line structures including the precuneus, cingulate, and medial prefrontal cortex (mPFC) (Fig. 2B, left panel). See Table 2 for more detailed description and coordinates. The reverse contrast (flash > beep) yielded significantly more activation in right premotor cortex (dorsal and ventral).

We next examined modality differences for the novel continuous stimuli: auditory sirens and visual moving bars (Table 3). In the (siren > moving bar) contrast, tapping with auditory sirens yielded more activation in auditory cortex, precuneus, cerebellum, and bilateral

Table 1

Tapping synchronization stability, R , for each of the four conditions by tempo, \pm standard error.

	Auditory beep	Visual flash	Auditory siren	Visual moving bar
Slow tempo	.942 \pm .009	.733 \pm .044	.873 \pm .018	.879 \pm .018
Fast tempo	.923 \pm .009	.315 \pm .057	.757 \pm .054	.526 \pm .063

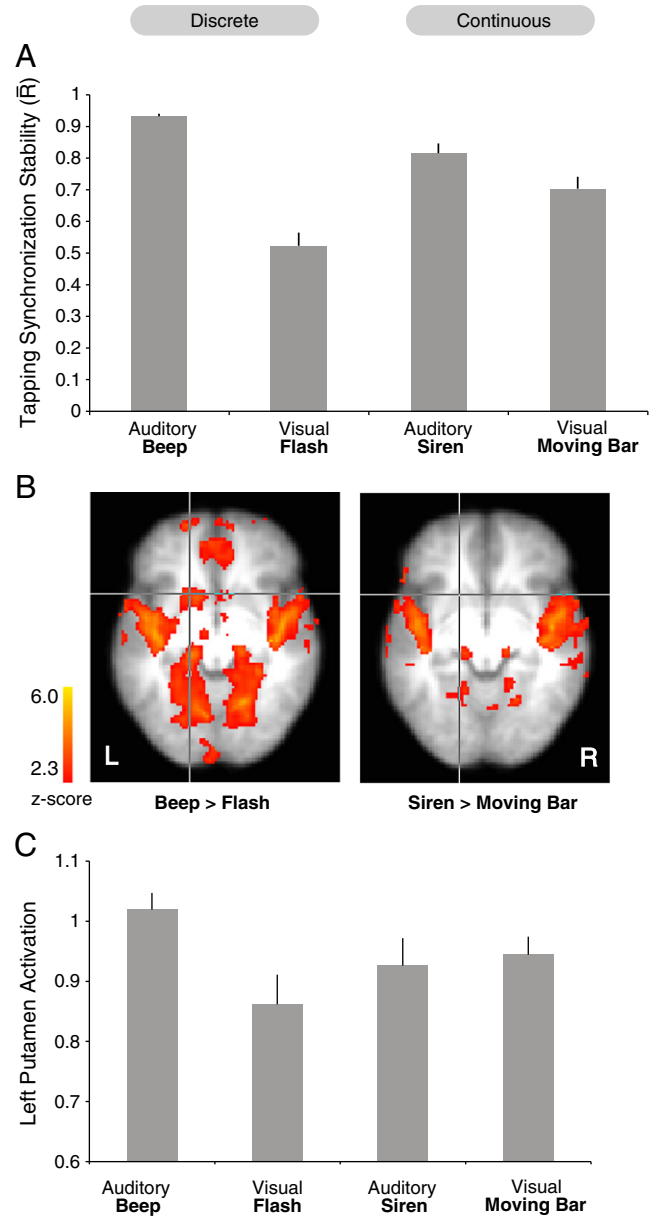


Fig. 2. A) Mean tap-to-target synchronization stability, R , for each of the four stimulus conditions. R quantifies tap-to-target stability as the resultant length of tap-to-target relative phases (i.e., $1 - \text{circular variance}$). B) Modality contrasts for discrete stimuli (beep > flash) on left, and for continuous stimuli (siren > moving bar) on right. Cross-hairs on the left putamen ($-18, 14, -8$ in MNI space) highlight a significant modality difference for the discrete stimuli (whole brain corrected), whereas no modality difference is observed for the continuous stimuli. C) Region of interest analysis for the left putamen activation (in arbitrary units). Error bars represent 1 standard error.

dorsolateral prefrontal cortex (dlPFC); no difference occurred in the putamen (Fig. 2B, right panel). In the reverse contrast (moving bar > siren) tapping with the moving bar yielded more activation in visual cortex and right premotor cortex.

In order to substantiate the differences observed between these two modality contrasts, we looked at the modality \times temporal structure interaction. Most importantly, this interaction [(beep > flash) > (siren > moving bar)] revealed a significant cluster in the left putamen [max activation, $Z = 3.42$ at MNI coordinate $x = -14$; $y = 10$; $z = -10$].

Within modality contrasts

We next examined differences between the discrete and continuous stimuli within each modality. For the auditory stimuli (Table 4), the (beep > siren) contrast yielded significantly more activation in the

Table 2

Modality contrasts of discrete stimuli, auditory beeps and visual flashes, showing local maxima of significant clusters at $p < .05$ (corrected).

Brain region	MNI coordinates			Z score
	x	y	z	
<i>Auditory beep > visual flash</i>				
L putamen	−18	14	−10	3.23
R cerebellum VIII	18	−58	−52	3.69
R cerebellum V/VI	26	−50	−26	3.62
L auditory cortex	−48	−14	0	6.74
R auditory cortex	44	−20	8	5.65
Precuneus	0	−46	50	4.94
Medial prefrontal cortex	2	52	0	4
Cingulate	−4	−10	46	4.42
<i>Visual flash > auditory beep</i>				
R premotor cortex	50	2	52	4.77
R premotor cortex	48	6	30	4.07

putamen, as well as in the precuneus, mPFC, and the angular gyrus. The reverse contrast (siren > beep) yielded more activation in the auditory cortex, the supramarginal gyrus, and the supplementary motor area (SMA).

For the visual stimuli (Table 5), the (flash > moving bar) contrast yielded a significant difference only in the left posterior cerebellum (Crus II). The (moving bar > flash) contrast yielded more activity in the visual cortex, premotor areas, and the supramarginal gyrus.

Region of interest analysis

To further compare activation of the left putamen between the four conditions, we performed an ROI analysis. Parameter estimates extracted from the left putamen sphere were subjected to a 2 (modality: audio, visual) \times 2 (temporal structure: discrete, continuous) repeated-measures ANOVA. Putamen activation results (Fig. 2C) corresponded to the tapping synchronization stability results (Fig. 2A). Overall, the left putamen was more active when participants tapped with auditory stimuli compared with visual stimuli, as indicated by a main effect of modality, $F(1, 13) = 8.03$, $p = .014$, $\eta_p^2 = .382$. No main effect of temporal structure was observed between discrete and continuous stimuli, $F(1, 13) = 0.02$, $p > .8$. In the critical test, a significant interaction of modality and temporal structure, $F(1, 13) = 10.52$, $p = .006$, $\eta_p^2 = .447$, indicates that putamen activation differs between modalities more for discrete stimuli (beeps vs. flashes) than for continuous stimuli (siren vs. moving bar). Post-hoc comparisons between modalities show that the putamen is considerably more active when synchronizing with

Table 3

Modality contrasts of continuous stimuli, auditory sirens and visual moving bar, showing local maxima of significant clusters at $p < .05$ (corrected).

Brain region	MNI coordinates			Z score
	x	y	z	
<i>Auditory siren > visual moving bar</i>				
L auditory cortex	−48	−18	6	6.15
R auditory cortex	50	−8	−2	6.07
Precuneus	10	−80	32	3.99
R cerebellum V	18	54	−16	3.24
R lingual gyrus	22	−58	−2	3.8
L cerebellum V	−12	−52	−14	3.32
L lingual gyrus	−18	−66	2	3.39
L cerebellum crus II	−28	−82	−40	3.64
R inferior frontal gyrus	48	18	16	3.75
L inferior frontal gyrus	−36	10	26	3.24
R frontal pole	32	58	4	3.33
<i>Visual moving bar > auditory siren</i>				
R occipital cortex	44	−68	0	5.15
L occipital cortex	−44	−68	6	4.84
R premotor cortex	34	−6	48	4
L superior parietal lobule	−24	−56	52	4.15

Table 4

Auditory contrasts within modality, auditory beeps and auditory sirens, showing local maxima of significant clusters at $p < .05$ (corrected).

Brain region	MNI coordinates			Z score
	x	y	z	
<i>Beep > siren</i>				
L putamen	−16	10	−10	3.14
Precuneus	−12	−54	6	4.33
Medial prefrontal cortex	4	62	10	4.04
L angular gyrus	−46	−72	36	4.05
R angular gyrus	36	−78	36	3.54
<i>Siren > beep</i>				
R auditory cortex	52	−14	2	5.36
L auditory cortex	−58	−16	6	5.51
R supramarginal gyrus	44	−40	38	3.31
Supplementary motor area	−4	0	56	3.64

beeps compared to flashes ($p = .002$); whereas no difference in putamen activation occurred between the siren and moving bar ($p = .567$). Within modality contrasts show more putamen activation for beeps than sirens ($p = .013$), and a trend (albeit non-significant) for more activation for the moving bar than flash ($p = .140$).

An additional ANOVA on putamen activation that included the additional factor of tempo revealed no main effect of tempo ($p > .2$), nor did tempo interact with modality ($p > .5$), temporal structure ($p > .1$) or the modality \times temporal structure interaction ($p > .7$). Thus, the effects of putamen activation were similar for the slow and fast tempi. This discounts the possibility that putamen activation was driven by task difficulty, which was presumably greater in the fast tempo.

In order to further examine the relation between synchronization stability and putamen activation, we ran a linear mixed-model analysis. The model included: synchronization stability (R) as the dependent variable; modality (audio/visual), temporal structure (discrete/continuous), and their interaction as fixed factors; putamen activation as a covariate; and participant as a random factor. Synchronization stability was significantly influenced by modality ($F = 58.2$, $p < .001$) and the modality \times temporal structure interaction ($F = 14.87$, $p < .001$), which corresponds to the ANOVA results reported above. Importantly, putamen activation was an additional significant predictor of synchronization stability ($F = 6.47$, $p = .014$). This result further substantiates the relationship between rhythmic synchronization stability and putamen activation. A participant's relative synchronization performance corresponds to his or her relative putamen activation. Our hypotheses are less concerned with individual differences, and more participants would be needed to test this question adequately. However, as an exploratory analysis, we ran correlations between synchronization stability and putamen activation separately for each of the four pacing sequences. Synchronization stability significantly correlated with putamen activation in the flash condition ($r = .54$, $p = .048$), but correlations in the other conditions were not significant ($ps > .3$).

Table 5

Visual contrasts within modality, visual flashes and visual moving bar, showing local maxima of significant clusters at $p < .05$ (corrected).

Brain region	MNI coordinates			Z score
	x	y	z	
<i>Flash > moving bar</i>				
L cerebellum crus II	−22	−84	−42	3.25
<i>Moving bar > flash</i>				
R occipital cortex	16	−92	−6	5.82
L occipital cortex	−24	−94	−2	4.85
L premotor cortex	−26	−6	52	4.35
R premotor cortex	32	−4	50	3.91
L supramarginal gyrus	−50	−28	42	3.39

Discussion

Behavioral effects

Visuo-motor synchronization was more stable with continuous moving targets than with discrete flashes; whereas audio-motor synchronization was more stable with discrete beeps than continuous pitch-modulated sirens. A modality difference in synchronization was observed between discrete beeps over flashes, but not between the continuous siren and moving bar. These results indicate that synchronization is not dictated by modality, but instead depends on the nature of the stimulus. Stable synchronization requires a precise and reliable encoding of the stimulus timing. Here the stimuli that afforded stable synchronization were well suited to each modality's processing affinity, and thus were likely encoded more precisely.

Difficulties in timing and synchronization with flashing visual stimuli are well established (e.g., Dunlap, 1910; Kolers and Brewster, 1985). Flashing stimuli provide only temporal information, and the visual system has relatively low temporal resolution (on the order of tens of milliseconds; Holcombe, 2009). Poor temporal resolution might partially stem from the slow transduction and processing times in visual system, which is tens of milliseconds slower than in the auditory system (Arrighi et al., 2006; King and Palmer, 1985); and longer processing should generate higher variability. The neural encoding of a flashing stimulus would thus have a relatively variable, diffuse spike-timing pattern, rather than a clear, precise encoding needed for perception-action integration.

Perceptual encoding of moving visual stimuli benefits from the visual system's high spatial resolution and motion processing, and results in improved visuo-motor synchronization. When a stimulus moves, the visual system's extensive motion-sensitive processing networks can track the temporal dynamics of the stimulus. Motion processing, along with anticipatory mechanisms, enable highly accurate interception of moving visual targets (Bootsma and van Wieringen, 1990; Zago et al., 2009) as required in visuo-motor synchronization. Previous work has shown that tapping synchronization improves with moving stimuli (Hove and Keller, 2010; Hove et al., 2010; Iversen et al., submitted for publication), and arm movements stably synchronize with compatibly moving stimuli (Buekers et al., 2000; Schmidt et al., 2007). In addition to improved perceptual encoding of visual motion, synchronization with moving stimuli likely benefited from the directional compatibility of the stimulus and finger movement. Strong stimulus-response compatibility assists sensorimotor integration (e.g., Hommel et al., 2001), and previous work has shown that tapping with incompatibly moving stimuli created little benefit over flashing stimuli (Hove et al., 2010).

Within the auditory modality, synchronization was more stable with beeps than frequency-modulated sirens. Pitch-modulated sequences are relatively unexplored in timing studies. In one previous study, tapping appeared less stable with frequency-modulated stimuli than with discrete clicks (McAnally, 2002).² Discrete auditory targets are likely encoded more precisely than pitch-modulated targets. The onset of a discrete auditory stimulus is known to elicit a strong phasic neural response (e.g., Barsz et al., 2002), and this phasic burst increases target salience. Conversely, when one tone closely follows another tone (as in the siren condition), the neural encoding of the target is considerably reduced (Barsz et al., 2002). Additionally, although the beep and siren

conditions had identical targets (50 ms tone at 1350 Hz), the siren's target was likely obscured by the surrounding tones through forward- and backward-masking. Together, a reduced neural response and masking would essentially 'blur' the siren's target, providing a less reliable encoding of target time, which ultimately results in lower synchronization stability.

In sum, synchronization performance is not dictated by modality, but is contingent upon the reliability and precision of target encoding. Stimulus encoding improves with motion in vision, and with a discrete target in audition. More reliable stimulus encoding yields more stable sensorimotor synchronization.

The importance of perceptual reliability, as opposed to strict modality dominance, has emerged in work on multimodal sensory integration. According to the Bayesian optimal integration hypothesis, multimodal information is integrated based on its relative precision or reliability (e.g., Alais and Burr, 2004; Ernst and Bühlhoff, 2004; Körding and Wolpert, 2004). For example, while audition typically dominates vision in timing tasks such as finger tapping with multimodal stimuli, when the visual stimulus moves (and is more reliable), the auditory dominance disappears (Hove et al., in press).

In sensorimotor synchronization, timing information from sensory and motor systems is integrated, and the degree of integration is likely based on relative reliability. When the sensory target is clearly encoded, it can be integrated into action timing and yield stable sensorimotor synchronization; whereas if the target is unreliable or 'blurry,' the diffuse spike timing pattern will not integrate with action timing, and the motor timing will tend to maintain its current intrinsic tempo (cf. Repp, 2005; Van Holst, 1939/1973). Interestingly, the reliability of encoding is also essential on the motor side of sensori-motor integration: when tactile feedback from *movement* is removed in finger tapping, synchronization degrades, and when tactile feedback is added in circle drawing, synchronization improves (Studenka and Zelaznik, 2011). Thus, sensorimotor synchronization requires a clear distinguishable encoding of both the pacing signal and the action. Finally, sensorimotor integration benefits from the representational overlap of perception and action codes (Hommel et al., 2001). The preferred stimulus structure for each modality (discrete for auditory and continuous for visual) overlaps greatly with the finger movement's perceptual effect in that modality. In general, the auditory feedback from finger tapping is discrete (hearing a 'thud' upon surface contact), and the visual feedback is continuous (seeing the finger move); the compatibility between the action's perceptual effect and the preferred stimulus style could promote the ease of perception-action integration.

Imaging results

Putamen activation

We manipulated the modality and structure of the pacing sequence in order to dissociate modality from synchronization performance. The most robust activation differences between conditions were observed in the putamen. Putamen activation was highest when tapping with beeps, intermediate with the moving bar and siren, and was lowest with flashes. Previous work had indicated that the putamen was consistently active in auditory, but not visual tapping tasks (e.g., the meta-analysis of Witt et al., 2008). However here, the nearly identical activation between modalities for continuous stimuli indicates that putamen activation is not dictated by modality. Instead, putamen activation corresponded to synchronization stability, with higher activation associated with more stable synchronization.

The putamen's association with rhythmic synchronization is consistent with its well-documented activation in rhythm and timing tasks (e.g., Coull et al., 2011; Grahn and Rowe, 2009; Teki et al., 2011; Wiener et al., 2009). While we did not directly measure or assess task difficulty, the conditions yielding poorer performance were also presumably more difficult. Differences in putamen activation could

² In another study, arm movements were more coordinated with a linearly rising pitch compared to a discrete beat at slow tempi (Rodger and Craig, 2011). However this effect occurred at tempi (2.5 and 4 s intervals) that were slower than established synchronization threshold rates (~2 s, Repp, 2005); thus the task is considerably different from the sub-second synchronization task here and in most synchronization studies.

stem from differences in task difficulty or general mental effort, rather than timing. However, this interpretation is unlikely. First, the effects for putamen activation were consistent across tempi, and were thus not driven by the fast (and presumably more difficult) tempo. Additionally, many studies that carefully control for task difficulty demonstrate that the putamen is primarily associated with timing demands, rather than task difficulty (Coull et al., 2011). For example, when performance is matched between a timing task and a control task such as color discrimination, many areas presumed to be part of the timing network do not survive the comparison; but the putamen remains an integral region for temporal processing after controlling for difficulty (Coull et al., 2004; Livesey et al., 2007). In this study, we manipulated modality and structure in order to influence synchronization performance (and presumably difficulty, since there is no obvious way to dissociate synchronization stability and difficulty). Instead of a strict modality difference in activation of subcortical timing networks, we observed that increased putamen activation occurs when the auditory or visual system provides reliable timing cues that afford stable synchronization.

Some beat perception studies have shown similar putamen activation between modalities when performance is similar. For example, in a rhythm discrimination task with auditory and moving visual stimuli, task performance did not differ between modalities, nor did respective putamen response, suggesting a common timing network (Schubotz et al., 2000). In another study, auditory beat perception was more sensitive and produced higher putamen activation than visual beats; but within the visual condition, the degree of putamen activity predicted beat sensitivity (Grahn et al., 2011). Thus putamen activity seems to reflect the relative beat strength.

As outlined in the previous section, relative beat strength is associated with clear sensory encoding. A phasic burst response in the early sensory pathway and thalamus, as occurs with beeps (Barsz et al., 2002) and moving visual stimuli (Sherman and Guillery, 2005), could provide precise timing information to the SMA and other cortical areas (Schwartz et al., 2012). The SMA is involved not only in motor output timing, but is also an important input structure to the basal ganglia. The successive inhibitory links in the basal ganglia's direct pathway to the thalamus can serve to select and amplify cortical excitability for motor timing in a winner-take-all fashion (Berns and Sejnowski, 1998). This amplified timing activity could be sent back to motor cortex thereby completing a thalamo–cortical–striatal loop, with boosted excitability at target times, and increased inhibition at other times in the rhythmic cycle.

Increased putamen activation during precise rhythmic timing could stem from monitoring oscillations in thalamo–cortical–striatal circuits (Buhusi and Meck, 2005). In this striatal beat frequency model (Matell and Meck, 2004), the striatum detects coincident firing of oscillatory patterns. A clear event onset might implement a 'start gun' that synchronizes the cortical oscillations (Buhusi and Meck, 2005), whereas an unclear signal would not. In sum, the putamen's similar response across modalities and its sensitivity to variations in performance support its key role in timing and rhythm. At the same time, it is part of a distributed timing system involving the coordination of large-scale networks and is tightly coupled with perceptual and action systems (cf. Meck et al., 2008).

Cortical activation

Early sensory areas responded largely as expected: In general, auditory cortex was more activated during auditory conditions, and visual cortex was more activated during visual conditions. The continuous stimuli produced higher activation in respective sensory areas than discrete stimuli, as would be expected from the higher degree of dynamic input.

We manipulated the modality and style of the pacing sequence and observed differences between conditions in respective sensory streams. The superior parietal lobule (SPL) was more activated during the moving bar condition than the siren, suggesting a role in visuo-motor processing. This is consistent with other work showing that the SPL is part of

the dorsal visual stream involved in sensorimotor transformations for visually guided action (Goodale and Milner, 1992) and is often activated in visuo-motor synchronization tasks (e.g., Jantzen et al., 2005). The opposite contrast (siren > moving bar) yielded more activation in the inferior frontal gyrus (IFG), suggesting a role in audio-motor processing, which is consistent with the IFG's proposed role in auditory-motor mappings (Rauschecker and Scott, 2009). The angular gyrus of the inferior parietal lobule was more active during the discrete than continuous auditory conditions, potentially reflecting tighter auditory-motor coupling, and is considered another key region in the auditory-motor interface (Rauschecker and Scott, 2009).

Activation differences were observed in the premotor cortex (PMC), an area involved in sensorimotor transformations (e.g., Chen et al., 2008; Kornysheva et al., 2011; Pollok et al., 2009; Zatorre et al., 2007). Visuo-motor compared to audio-motor synchronization yielded greater activation in the right dorsal PMC (discrete and continuous contrasts), and in the right ventral PMC (discrete contrast only). Greater activation of ventral PMC with visual flashes than auditory targets has also been reported in other studies (Jantzen et al., 2005; Pollok et al., 2009), whereas another study observed relatively more activity with auditory targets (Jäncke et al., 2000). Increased dorsal PMC activation has been reported when synchronizing with auditory beeps compared with visual flashes (Pollok et al., 2009). These effects are typically reported in the contralateral motor control network (left lateralized), whereas all the modality differences that we observed in the PMC were ipsilateral to the finger movement (right lateralized). It is possible that visual synchronization increasingly recruits the ipsilateral PMC. Additionally, a number of studies have shown that PMC responds more when a sensory signal is temporally complex or difficult (Chen et al., 2008; Grahn and Rowe, 2009; Lewis et al., 2004). Thus the higher activation in PMC here might reflect a relative difficulty in extracting temporal features from the visual sequences.

Higher activation was observed in a number of midline structures, including the precuneus, the cingulate, and the medial prefrontal cortex, during more stable tapping conditions. These midline areas are part of the so-called task-negative or default mode network, and tend to be more active when task demands are low (e.g., Fairhurst et al., in press; Fox and Raichle, 2007). Relatively high activation was observed most clearly in the beep condition compared to flash condition, which is consistent with the relative ease of tapping with beeps and the increased task demands of tapping with the flash.

Cerebellar activation

Cerebellar involvement in sensorimotor timing in the subsecond range has been established in previous neuroimaging and neuropsychological work (e.g., Penhune et al., 1998; Rao et al., 1997; Spencer and Ivry, 2012). Here, the right anterior lobe (lobule V) was more activated during audio- compared to visuo-motor synchronization for both discrete and continuous contrasts, suggesting a possible auditory specialization. Additionally, the right inferior lobule VIII was more activated during audio-motor synchronization only for the discrete contrast (beep > flash), possibly reflecting more precise timing.

Right lateralized activity during finger tapping with the right hand is consistent with the cerebellum's ipsilateral somatotopy. The cerebellum's anterior lobe V and lobule VIII are important for sensorimotor function (Stoodley et al., 2010), and lobules V–VIII are often implicated in time processing (Spencer and Ivry, 2012). Evidence from TMS supports distinct cerebellar circuits for processing auditory and visual cues (Del Olmo et al., 2007). However, other studies have shown that audio-motor and visuo-motor synchronization yielded similar activation and concordance in the anterior cerebellum (Witt et al., 2008). In general, we interpret cerebellum activation cautiously, as cerebellar activation patterns have proven difficult to link to functional hypotheses (Spencer and Ivry, 2012).

Conclusion

In conclusion, this study presents evidence that sensorimotor synchronization is largely contingent upon the stimuli's suitability to the processing style of each modality. In the auditory modality, a discrete beep enables a clear encoding of the target timing; and in the visual modality, a continuously moving target can be more clearly encoded due to its spatiotemporal dynamics. After this early encoding, it can serve to coordinate action timing via thalamo-cortical-striatal loops, and differences in rhythmic synchronization performance are especially apparent in the putamen.

The current findings have implications for clinical application. Differences in basal ganglia activation are relevant for Parkinson's gait rehabilitation. Parkinsonian gait improves with auditory rhythmic cueing (e.g., Hove et al., 2012), but flashing lights in goggles are far less effective (Arias and Cudeiro, 2008; Rochester et al., 2005). Spatial visual cues (i.e. stripes painted on the ground) can improve Parkinsonian gait (e.g., Morris et al., 1996), and future work could integrate rhythmicity and spatial visual cueing by employing dynamic moving cues over goggles. Thus, optimizing the design of cues according to modality-specific processing affinities could boost brain function in sensorimotor rehabilitation most effectively.

Acknowledgments

This research was supported by The Max Planck Society. We thank Jan Bergmann for technical assistance and Maria Bader for help in data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.neuroimage.2012.11.032>.

References

- Alais, D., Burr, D., 2004. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.
- Arias, P., Cudeiro, J., 2008. Effects of rhythmic sensory stimulation (auditory, visual) on gait in Parkinson's disease patients. *Exp. Brain Res.* 186, 589–601.
- Arrighi, R., Alais, D., Burr, D., 2006. Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *J. Vis.* 6, 260–268.
- Barsz, K., Ison, J.R., Snell, K.B., Walton, J.P., 2002. Behavioral and neural measures of auditory temporal acuity in aging humans and mice. *Neurobiol. Aging* 23, 565–578.
- Berns, G.S., Sejnowski, T.J., 1998. A computational model of how the basal ganglia produce sequences. *J. Cogn. Neurosci.* 10, 108–121.
- Bertelson, P., Radeau, M., 1981. Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.* 29, 578–584.
- Bootsma, R.J., Van Wieringen, P.C., 1990. Timing an attacking forehand drive in table tennis. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 21–29.
- Buekers, M.J., Bogaerts, H.P., Swinnen, S.P., Helsen, W.F., 2000. The synchronization of human arm movements to external events. *Neurosci. Lett.* 290, 181–184.
- Buhusi, C.T., Meck, W.H., 2005. What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765.
- Chen, J.L., Penhune, V.B., 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.
- 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. *Hum. Mov. Sci.* 21, 515–532.
- Coull, J.T., Cheng, R.-K., Meck, W.H., 2011. Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacol. Rev.* 36, 3–25.
- Coull, J.T., Vidal, F., Nazarian, B., Macar, F., 2004. Functional anatomy of the attentional modulation of time estimation. *Science* 303, 1506–1508.
- Del Olmo, M.F., Cheeran, B., Koch, G., Rothwell, J.C., 2007. Role of the cerebellum in externally paced rhythmic finger movements. *J. Neurophysiol.* 98, 145–152.
- Dunlap, K., 1910. Reactions to rhythmic stimuli, with attempt to synchronize. *Psychol. Rev.* 17, 399–416.
- Ernst, M.O., Bühlhoff, H.H., 2004. Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169.
- Fairhurst, M.F., Janata, P., Keller, P.E., in press. Being and feeling in sync with an adaptive virtual partner: Brain mechanisms underlying dynamic cooperativity. *Cereb. Cortex*.
- Fisher, N.J., 1993. *Statistical Analysis of Circular Data*. Cambridge University Press, Cambridge. <http://dx.doi.org/10.1093/cercor/bhs243>.
- Fox, M.D., Raichle, M.E., 2007. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat. Rev. Neurosci.* 8, 700–711.
- Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. *Hum. Brain Mapp.* 1, 210–220.
- Goodale, M.A., Milner, A.D., 1992. Separate visual pathways for perception and action. *Trends Neurosci.* 15, 20–25.
- Grahn, J.A., Henry, M.J., McAuley, J.D., 2011. fMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *NeuroImage* 54, 1231–1243.
- Grahn, J.A., Rowe, J.B., 2009. Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548.
- Holcombe, A.O., 2009. Seeing slow and seeing fast: two limits on perception. *Trends Cogn. Sci.* 13, 213–221.
- Hommel, B., Müsseler, J., Aschersleben, G., Prinz, W., 2001. The theory of event coding (TEC): a framework for perception and action planning. *Behav. Brain Sci.* 24, 849–937.
- Hove, M.J., Iversen, J.R., Zhang, A., Repp, B.H., in press. Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychol. Res.* <http://dx.doi.org/10.1007/s00426-012-0441-0>.
- Hove, M.J., Keller, P.E., 2010. Spatiotemporal relations and movement trajectories in visuomotor synchronization. *Music Percept.* 28, 15–26.
- Hove, M.J., Spivey, M.J., Krumhansl, C.L., 2010. Compatibility of motion facilitates visuomotor synchronization. *J. Exp. Psychol. Hum. Percept. Perform.* 36, 1525–1534.
- Hove, M.J., Suzuki, K., Uchitomi, H., Orimo, S., Miyake, Y., 2012. Interactive rhythmic auditory stimulation reinstates natural 1/f timing in gait of Parkinson's patients. *PLoS One* 7, e32600.
- Iversen, J.R., Patel, A.D., Nicodemus, B., Emmorey, K., submitted for publication. Synchronization to auditory and visual rhythms in hearing and deaf individuals.
- Jäncke, L., Loose, R., Lutz, K., Specht, K., Shah, N.J., 2000. Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cogn. Brain Res.* 10, 51–66.
- Jantzen, K.J., Steinberg, F.L., Kelso, J.A.S., 2005. Functional MRI reveals the existence of modality and coordination-dependent timing networks. *NeuroImage* 25, 1031–1042.
- Jenkinson, M., Smith, S., 2001. A global optimisation method for robust affine registration of brain images. *Med. Image Anal.* 5, 143–156.
- King, A.J., Palmer, A.R., 1985. Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp. Brain Res.* 60, 492–500.
- Kolers, P.A., Brewster, J.M., 1985. Rhythms and responses. *J. Exp. Psychol. Hum. Percept. Perform.* 11, 150–167.
- Körding, K.P., Wolpert, D.M., 2004. Bayesian integration in sensorimotor learning. *Nature* 427, 244–247.
- Kornysheva, K., von Anshelm-Schiffer, A.M., Schubotz, R.I., 2011. Inhibitory stimulation of the ventral premotor cortex temporarily interferes with musical beat rate preference. *Hum. Brain Mapp.* 32, 1300–1310.
- Kotz, S.A., Schwartz, M., Schmidt-Kassow, M., 2009. Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language. *Cortex* 45, 982–990.
- Lewis, P.A., Wing, A.M., Pope, P.A., Praamstra, P., Miall, R.C., 2004. Brain activity correlates differentially with increasing temporal complexity of rhythms during initialisation, synchronisation, and continuation phases of paced finger tapping. *Neuropsychologia* 42, 1301–1312.
- Livesey, A.C., Wall, M.B., Smith, A.T., 2007. Time perception: manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia* 45, 321–331.
- Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., Maquet, P., 2002. Activation of the supplementary motor area and of attentional networks during temporal processing. *Exp. Brain Res.* 142, 475–485.
- Matell, M.S., Meck, W.H., 2004. Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cogn. Brain Res.* 21, 139–170.
- McAnally, K.L., 2002. Timing of finger tapping to frequency modulated acoustic stimuli. *Acta Psychol.* 109, 331–338.
- Meck, W.H., Penney, T.B., Pouthas, V., 2008. Cortico-striatal representation of time in animals and humans. *Curr. Opin. Neurobiol.* 2, 145–152.
- Morris, M.E., Iansek, R., Matyas, T., Summers, J.J., 1996. Stride length regulation in Parkinson's disease. Normalization strategies and underlying mechanisms. *Brain* 119, 551–568.
- Penhune, V.B., Zatorre, R.J., Evans, A.C., 1998. Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J. Cogn. Neurosci.* 10, 752–765.
- Pollak, B., Krause, V., Butz, M., Schnitzler, A., 2009. Modality specific functional interaction in sensorimotor synchronization. *Hum. Brain Mapp.* 30.
- Rao, S.M., Harrington, D.L., Haaland, K.Y., Bobholz, J.A., Cox, R.W., Binder, J.R., 1997. Distributed neural systems underlying the timing of movements. *J. Neurosci.* 17, 5528–5535.
- Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724.
- 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. *J. Mot. Behav.* 35, 355–370.
- Repp, B.H., 2005. Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992.
- Repp, B.H., Penel, A., 2004. Rhythmic movement is attracted more strongly to auditory than visual rhythms. *Psychol. Res.* 68, 252–270.
- Rochester, L., Hetherington, V., Jones, D., Nieuwboer, A., Willems, A.M., Kwakkel, G., van Wegen, E., 2005. The effect of external rhythmic cues (auditory and visual) on walking during a functional task in homes of people with Parkinson's disease. *Arch. Phys. Med. Rehabil.* 86, 999–1006.

- Rodger, M.W.M., Craig, C.M., 2011. Timing movements to interval durations specified by discrete or continuous sounds. *Exp. Brain Res.* 214, 393–402.
- Schmidt, R.C., Richardson, M.J., Arsénault, C., Galantucci, B., 2007. Visual tracking and entrainment to an environmental rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 860–870.
- Schubotz, R.I., Friederici, A.D., von Cramon, D.Y., 2000. Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *NeuroImage* 11, 1–12.
- Schwartz, M., Tavano, A., Schröger, E., Kotz, S.A., 2012. Temporal aspects of prediction in audition: cortical and subcortical neural mechanisms. *Int. J. Psychophysiol.* 83, 200–207.
- Semjen, A., Schulze, H.-H., Vorberg, D., 2000. Timing precision in continuation and synchronization tapping. *Psychol. Res.* 63, 137–147.
- Sherman, S.M., Guillery, R.W., 2005. *Exploring the Thalamus and Its Role in Cortical Function*. MIT Press, London and Cambridge.
- Smith, S.M., 2002. Fast robust automated brain extraction. *Hum. Brain Mapp.* 17, 143–155.
- Spencer, R.M.C., Ivry, R.B., 2012. Cerebellum and timing. In: Manto, M., Gruol, D., Schmahmann, J., Koibuchi, N., Rossi, F. (Eds.), *Handbook of the Cerebellum and Cerebellar Disorders*. Springer Press.
- Stoodley, C.J., Valera, E.M., Schmahmann, J.D., 2010. An fMRI study of intra-individual functional topography in the human cerebellum. *Behav. Neurol.* 23, 65–79.
- Studenka, B.E., Zelaznik, H.N., 2011. Synchronization in repetitive smooth movement requires perceptible events. *Acta Psychol.* 136, 432–441.
- Teki, S., Grube, M., Kumar, S., Griffiths, T.D., 2011. Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* 31, 3805–3812.
- Van Holst, E., 1939/1973. Relative coordination as a phenomenon and as a method of analysis of central nervous functions. *The Collected Papers of Erich von Holst: Vol. 1. The Behavioural Physiology of Animals and Man*. University of Miami Press, Coral Gables, FL, pp. 33–135.
- Vorberg, D., Wing, A.M., 1996. Modeling variability and dependence in timing. In: Heuer, H., Keele, S.W. (Eds.), *Handbook of Perception and Action*. Academic Press, London, pp. 181–262.
- Welch, R.B., Warren, D.H., 1980. Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667.
- Wiener, M., Turkeltaub, P., Coslett, H.B., 2009. The image of time: a voxel-wise meta-analysis. *NeuroImage* 49, 1728–1740.
- Witt, S.T., Laird, A.R., Meyerand, M.E., 2008. Functional neuroimaging correlates of finger-tapping task variations: an ALE meta-analysis. *NeuroImage* 42, 343–356.
- Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage* 14, 1370–1386.
- Worsley, K.J., Evans, A.C.S.M., Neelin, P., 1992. A three-dimensional statistical analysis for CBF activation studies in human brain. *J. Cereb. Blood Flow Metab.* 12, 900–918.
- Zago, M., McIntyre, J., Senot, P., Lacquaniti, F., 2009. Visuo-motor coordination and internal models for object interception. *Exp. Brain Res.* 192, 571–604.
- Zatorre, R.J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558.