

Sensory and Motor Systems

Perceptual Oscillation of Audiovisual Time Simultaneity

Alessandro Benedetto,¹ David Charles Burr,^{2,3} and Maria Concetta Morrone^{1,4}DOI:<http://dx.doi.org/10.1523/ENEURO.0047-18.2018>

¹Dipartimento Di Ricerca Traslazionale e Delle Nuove Tecnologie in Medicina e Chirurgia, Università Di Pisa, via San Zeno 31, Pisa, 56123, Italy, ²Istituto Di Neuroscienze, Consiglio Nazionale Delle Ricerche (CNR), via Moruzzi 1, Pisa, 56124, Italy, ³Dipartimento Di Psicologia, Farmacologia e Salute Del Bambino (NEUROFARBA), Università Di Firenze, via San Salvi 12, Firenze, 50135, Italy, and ⁴IRCCS Stella-Maris, Viale Del Tirreno 331, Pisa, 56018, Italy Calambrone

Abstract

Action and perception are tightly coupled systems requiring coordination and synchronization over time. How the brain achieves synchronization is still a matter of debate, but recent experiments suggest that brain oscillations may play an important role in this process. Brain oscillations have been also proposed to be fundamental in determining time perception. Here, we had subjects perform an audiovisual temporal order judgment task to investigate the fine dynamics of temporal bias and sensitivity before and after the execution of voluntary hand movement (button-press). The reported order of the audiovisual sequence was rhythmically biased as a function of delay from hand action execution. Importantly, we found that it oscillated at a theta range frequency, starting ~500 ms before and persisting ~250 ms after the button-press, with consistent phase-locking across participants. Our results show that the perception of cross-sensory simultaneity oscillates rhythmically in synchrony with the programming phase of a voluntary action, demonstrating a link between action preparation and bias in temporal perceptual judgments.

Key words: Audiovisual; Behavioral Oscillations; Decision Bias; Simultaneity; Theta; Time Perception

Significance Statement

Judgments of the temporal order of a visual and auditory stimulus are not constant over time but fluctuate rhythmically in the theta range (~7 Hz) in synchrony with the execution of a voluntary movement. Interestingly, these oscillations precede the action execution by about half a second, indicating active synchronization between temporal criterion and the intention to move. Overall, the results point to the presence of a single shared clock for perception and action.

Introduction

Action and perception are tightly coupled systems requiring strong coordination and synchronization over time. How-

ever, how the brain achieves synchronization is still a matter of debate. Several electrophysiological studies have shown that neural oscillations preceding sensory stimulation are causally linked to perception (VanRullen and Koch, 2003; Linkenkaer-Hansen et al., 2004; Busch et al., 2009; de Lange et al., 2013; Hanslmayr et al., 2013; Sherman et al., 2016; VanRullen, 2016), and that endog-

Received January 29, 2018; accepted March 17, 2018; First published May 21, 2018.

The authors declare no competing financial interests.

Author contributions: A.B. performed research. All authors designed research, analyzed data, and wrote the paper.

This work was supported by the European Research Council (FP7 ECS-PLAIN Grant 338866 to M.C.M.).

Correspondence should be addressed to Maria Concetta Morrone, Department of Translational Research on New Technologies in Medicines and Surgery, University of Pisa, via San Zeno 31, 56123 Pisa, Italy. E-mail: concetta@in.cnr.it.

DOI:<http://dx.doi.org/10.1523/ENEURO.0047-18.2018>

Copyright © 2018 Benedetto et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

ogenous oscillations effectively shape perceptual processing. Stimulus-locked behavioral oscillations have been demonstrated in perceptual performance (Fiebelkorn et al., 2011; Landau and Fries, 2012; Romei et al., 2012; Ho et al., 2017), confirming that the system can lock activity to relevant environmental cues. A growing body of scientific literature has shown that voluntary actions can also synchronize perceptual rhythms: for up to a second before and after executing a voluntary action, some visual properties, including visual contrast sensitivity (Tomassini et al., 2015; Benedetto et al., 2016; Benedetto and Morrone, 2017), visual attention (Hogendoorn, 2016), or temporal integration/segregation (Wutz et al., 2016), oscillate at slow frequencies phase-locked with the action execution. Conversely, it has been shown that the phase of neural oscillations can predict reaction time to perceptual events (Lansing, 1957; Surwillo, 1961; Drewes and VanRullen, 2011), and that a transient visual response in the central nervous system can reset the phase of low-frequency tremor oscillations in peripheral muscles (Wood et al., 2015). Despite all this evidence connecting the intrinsic oscillatory nature of perception with action, there is no clear consensus about the basic mechanisms and function of rhythmic modulation on early brain function (Engel and Singer, 2001; Fries, 2005; Klimesch et al., 2007). One fascinating idea is that temporal mechanisms act through synchronization of endogenous oscillations, achieving in this way the temporal binding of independent perceptual, motor, and cognitive phenomena.

This suggestion is supported by recent electrophysiological experiments showing that the visual temporal windows of integration/segregation oscillate rhythmically within delta, theta, and alpha ranges (Ronconi et al., 2017), and similar effects have been shown for (multi-)sensory temporal resolution (Varela et al., 1981; Samaha and Postle, 2015; Cecere et al., 2015; Milton and Pleydell-Pearce, 2016; Benedetto et al., 2017). Moreover, EEG experiments have shown that brain rhythms in the delta/theta range can modulate both the predictability of event timing (Stefanics et al., 2010) and the encoding of memory sequences of events (Heusser et al., 2016).

It is well established that action and perceptual timing are strongly interconnected. Perceived visual duration can be strongly influenced by action, being compressed and/or dilated for perisaccadic stimuli, and also during hand movements (Haggard et al., 2002; Park et al., 2003; Morrone et al., 2005; Binda et al., 2009; Hagura et al., 2012; Tomassini and Morrone, 2016). This action-induced temporal modulation has been observed also for other sensory domains, such as tactile perception, suggesting that it is a general mechanism (Yarrow and Rothwell, 2003; Tomassini et al., 2012, 2014). To test directly if the temporal resolution of perceptual bias and sensitivity also fluctuate rhythmically, we measured the perceptual temporal order of auditory and visual targets before and after a voluntary action execution, hypothesizing a rhythmic modulation of perceptual bias, synchronized with the time of action.

Materials and Methods

Participants

Nine subjects (age mean \pm SEM: 27 ± 0.6 , 3 females) took part in the experiment. All had normal or corrected-to-normal vision and normal audition. All participants provided informed consent, and the experiment was approved by the ethics committee (Comitato Etico Pediatrico Regionale, Azienda Ospedaliero-Universitaria Meyer, Firenze, Italy).

Apparatus

The visual stimuli were generated by the VisaGe (Cambridge Research System) controlled via CRS Toolbox for Matlab (Matlab r2007a, The Mathworks, inc.) and displayed at 57 cm on a gamma-calibrated CRT monitor (Barco Calibrator Line) with a resolution of 800×600 pixels, a refresh rate of 120 Hz, and mean luminance of $38.5 \text{ cd} \cdot \text{m}^{-2}$. Auditory stimuli were generated by the ViSaGe through its trigger output port, synchronized with the refreshing rate of the monitor, and amplified by a speaker positioned centrally, below the monitor. In this way, we ensured a perfect temporal synchronization between the visual and the auditory stimuli, verified by direct measurements. The responses were recorded with an infrared CB6 Response Box (Cambridge Research System) controlled via CRS Toolbox for Matlab.

Stimulus and procedure

Participants maintained fixation on a red square (0.25°) presented in the center of screen on a gray background. The fixation appeared at the beginning of the session and lasted until the end. Stimuli consisted of a visual stimulus, preceded or followed by an auditory stimulus. The visual stimulus consisted of a black Gaussian-blob (contrast of 50%, sigma of 5°) presented in the center of the screen. The auditory stimulus was a short suprathreshold noise burst presented from a central speaker positioned below the monitor. Both visual and auditory stimuli were 8-ms duration and were synchronized with the refreshing rate of the monitor.

The task was a two-alternative forced-choice audiovisual temporal-order judgment (TOJ). Volunteers initiated the trial sequence by pressing a start button, and after a random delay from the button-press, the audiovisual sequence was presented. The stimulus-onset asynchrony (SOA) between the audiovisual stimulus was random, within 0 and ± 500 ms (mean \pm SEM trials per SOA: 74 ± 8), with a higher proportion of trials presented around the time of the physical simultaneity of the audiovisual stimuli (SOA between ± 16 ms: 206 ± 57). The delay between the first stimulus of the pair and the button-press was randomly selected from 15 values: 0, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 300, 350, 450, and 500 ms. Note that this sampling created a time resolution of 25 ms for the first 250 ms from button-press, decreasing to 50 ms for later latencies. Subjects had to report, via button press, whether the sound appeared to precede or succeed the Gaussian blob. The timing of the start of each trial and the responses were voluntarily paced, but had to be continuously spaced within 1.5–3.0 s, resulting in a continuous slow rhythmical tapping within 0.33 and 0.66 Hz. The trials

that did not fit this timing were discarded from further analysis. To investigate the dynamics of audiovisual integration before action execution, in ~20% of the trials the stimulus was unpredictably presented before the start signal. We asked participants to maintain their start/response pacing, even when the start action was not causally related to stimulus onset, but to continue to pay the same attention to these trials. Several sessions were recorded over different days. Overall, we collected 10,776 trials (1197 ± 257 trials per participant, mean \pm SEM).

Data analysis

Analyses were conducted both on individual data and also after pooling all data together into a single dataset (hereafter termed the aggregate-observer). To investigate the bias and sensitivity dynamics, aggregate-observer responses were fitted with a cumulative Gaussian. The responses were reported as proportion of vision leading, modeled as a function of the signed SOAs of the auditory stimulus with respect to the visual. Perceptual bias was evaluated by measuring the point of subjective simultaneity (PSS) of the psychometric function (the point where the proportion of vision-leading is equal to 50%). Sensitivity was evaluated by the slope of the psychometric function, which corresponds to the just-noticeable difference (JND). This is a standard procedure following the Green and Swets signal detection theory (Green and Swets, 1966). An example of this fitting is shown in Fig. 1 for each individual observer. If not otherwise stated, the time from action refers to the temporal delay between the action onset and the first stimulus presented.

We computed a psychometric function on the aggregate observer for each time lag from -275 to 500 ms from action execution (binning the pre-action trials in nonoverlapping bins of 25 ms) and extracted the PSS and the JND from each estimated model (Figs. 2 and 3). For both PSS and JND dynamics, we fitted the best sinusoidal model for each dataset at lags between -275 and 250 ms from action execution. We selected the nearest bins falling around the time of a peak or trough of the best oscillation and computed a psychometric function pooling together all trials falling the peak or trough bins. A bootstrap *t* test was run to test differences in PSS and JND between the peaks and the troughs of the sinusoidal model. The best-fit oscillatory models from the PSS and the JND were statistically evaluated with a bootstrap procedure on surrogate data obtained by randomly shuffling the responses of each trial (using the same time stamps), then performing the standard binning procedure (1000 reiterations). To control for multiple comparisons, the surrogate data were fitted with the best sinusoidal wave form, with frequency, amplitude, and phase as free parameters (Benedetto and Morrone, 2017). A one-tailed nonparametric bootstrap *t* test was run to assess whether the *R*-squared values of the best fit of the data were statistically higher than the 95% of the *R*-squared distribution obtained from the surrogate dataset.

To test the spectral content of the vision leading dynamics (as an approximation of response bias), we ran a single-trial multivariate generalized linear model (glm) analy-

sis (Tomassini et al., 2017) on the aggregate observer for lags within -450 and 250 ms from action execution, including all trials with an SOA less than ± 100 ms. No bias was present in the distribution of the SOAs across subjects and time-from-action execution (mean SOA \pm interquartile range: 0.26 ± 3.05 ms). We fitted a linear regression model including as predictors a sine and a cosine for a given frequency of interest f_i . The probability model behind this analysis can be written as

$$\hat{Y}_n = \beta_0 + \beta_1 \sin(\omega t_n) + \beta_2 \cos(\omega t_n),$$

where t is the time lag of the single trial n ; β_0 , β_1 , and β_2 are the fixed-effect linear regression parameter; \hat{Y}_n is the predicted behavioral performance; and ω is the angular frequency ($\omega = 2\pi f_i$). The fixed-effect linear regression parameters were estimated using standard least square method (LSM) as

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = (X^t X)^{-1} X^t Y,$$

where Y is the vector of the single trial responses (0 or 1 for audition or vision leading responses, respectively). The matrix X has the form:

$$X = \begin{bmatrix} 1 \sin(\omega t_1) \cos(\omega t_1) \\ \dots \dots \dots \\ 1 \sin(\omega t_n) \cos(\omega t_n) \end{bmatrix}.$$

Statistics were computed via permutation obtained by shuffling the responses (10,000 repetitions) to create an empirical noise distribution. We adopted the maximum statistics to correct for multiple comparison: for each permutation, we selected the maximal amplitude of the beta coefficients across all frequencies and ran a bootstrap *t* test between the distribution of the maximal amplitudes and the amplitude obtained from the actual dataset for each frequency. The spectral amplitude (A) was computed as the Euclidean length of the β_1 and β_2 coefficients as

$$A = \sqrt{\beta_1^2 + \beta_2^2}.$$

The amplitude error was estimated by implementing a jackknife resampling procedure to explore the weight of each participant in determining the actual result.

The advantage of this analysis over more traditional approaches (e.g., Fourier analysis or fitting) is that there is no need to bin the data, which can potentially give rise to spurious results given the fixed constant sample time. The analysis also works well for nonuniform and sparse samples, and for this reason is optimal not only for aggregate observer analysis, but also for group mean statistics. We investigated a frequency range between 3 and 15 Hz (with a resolution of 0.01 Hz). Furthermore, we selected the peak of significance from the aggregate observer analysis to investigate single-subject variability. For the selected frequency, we tested against zero the average of the participant-specific beta coefficients (expressed as am-

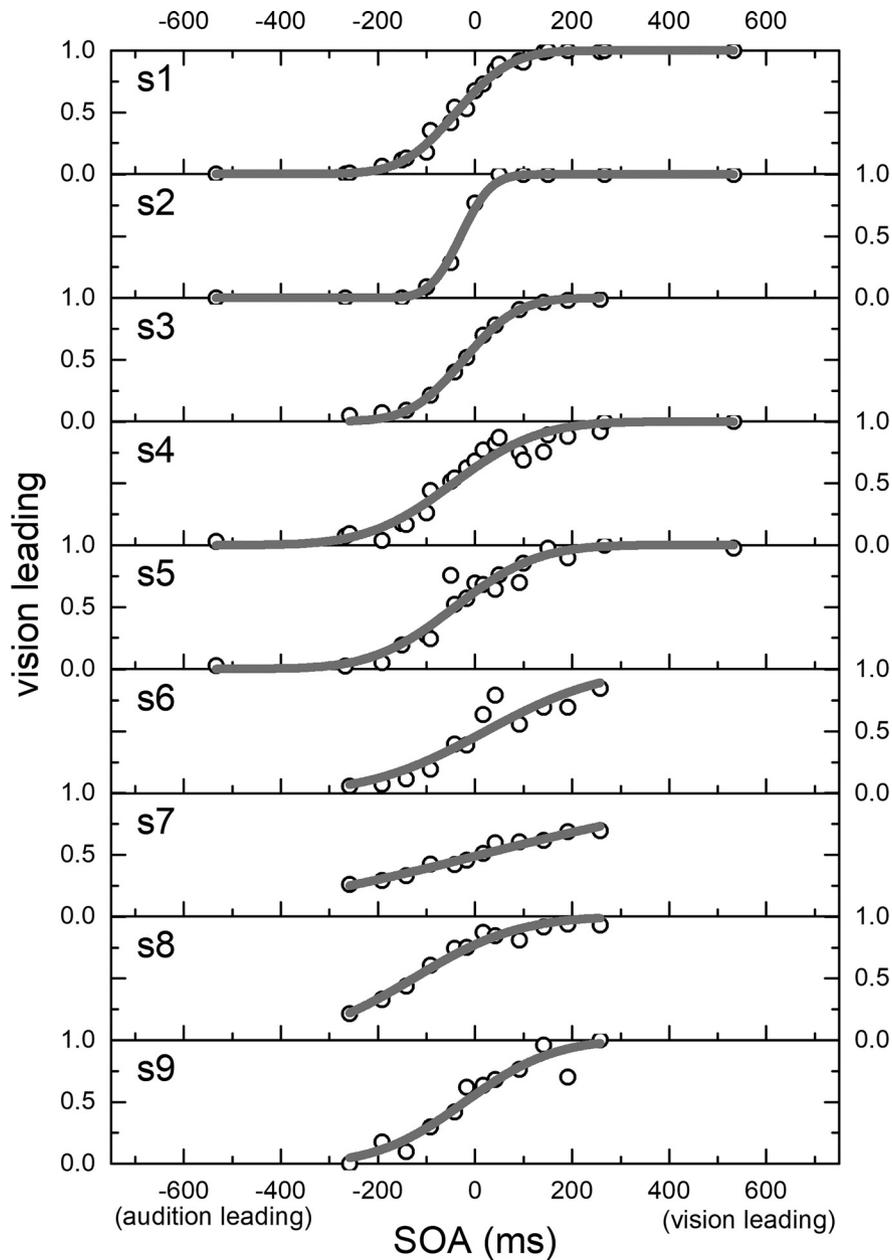


Figure 1. Data and fitted psychometric functions for all participants ($N = 9$) pooling all trials at all delays (1197 ± 257 trials per participant, mean \pm SEM). Proportion of visual leading responses is plotted as a function of stimulus onset asynchrony (SOA), defined as positive when vision leads. Gray curves show the fitted psychometric function.

plitude and phases of the significant components) by means of the bivariate Hotelling’s T -squared statistic (as in Tomassini et al., 2017):

$$T^2 = k(\bar{\beta}_1, \bar{\beta}_2)'S^{-1}(\bar{\beta}_1, \bar{\beta}_2),$$

where k denotes the number of subjects, $(\bar{\beta}_1, \bar{\beta}_2)$ is the sample mean (across subjects) of the vector of linear regression coefficients (β_1, β_2) , and S^{-1} is the inverse of the sample covariance matrix of these vector-valued regression coefficients. This Hotelling’s T -square test is an extension of the Student’s t test to the multivariate domain, and it provides significant results only if the regres-

sion coefficients are large and have the same sign across subjects (Tomassini et al., 2017).

We also investigated group-mean fluctuations in bias with glm analysis. We applied the analysis to each individual subject, and then averaged—for each tested frequency—the norm of the resulting sine and cosine vector (norm of vector average). It is noteworthy that this approach, similarly to the aggregate observer analysis, is aimed at revealing the phase-locked oscillations across participants, the modulation of the bias that is shared by all subjects. The error of the spectrum was computed via a jackknife resampling. Additionally, to further assure us that the significant oscillation was restricted to those

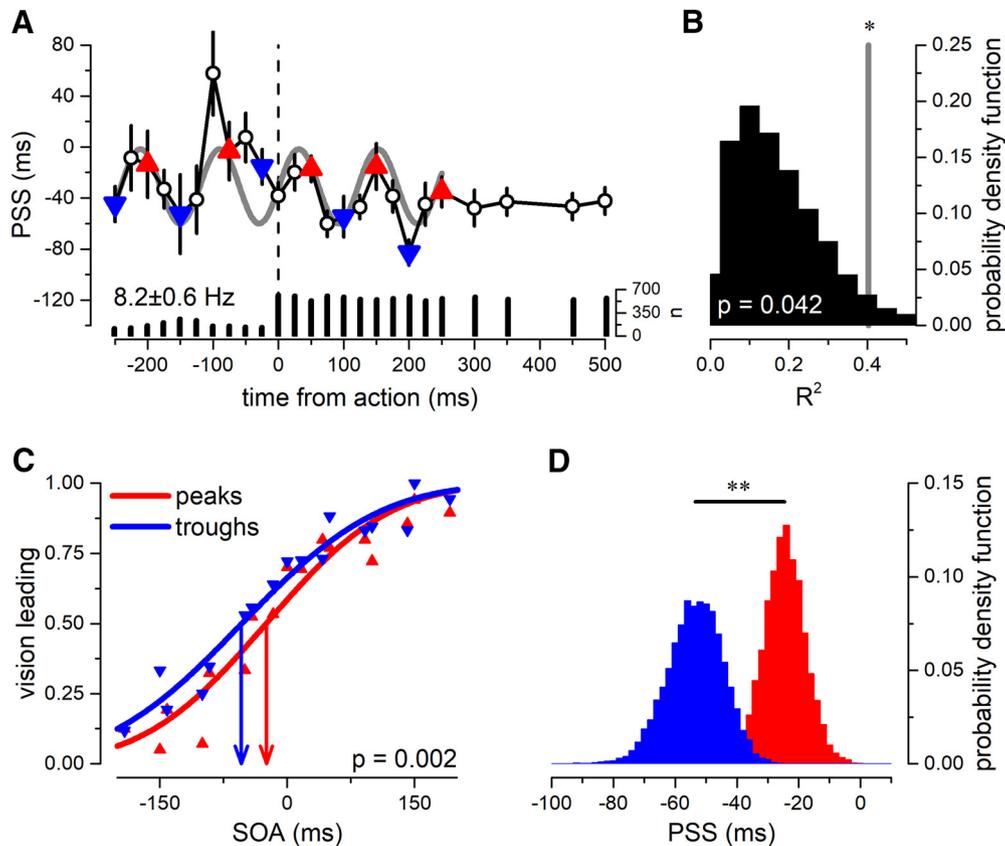


Figure 2. Dynamics of PSS for the aggregate observer ($N = 9$). **A**, $PSS \pm 1$ SEM as a function of time from action execution. The gray line reports the best sinusoidal fit, with a frequency of 8.2 Hz. Red upright triangles, points nearest to the predicted peaks of the oscillation; blue inverted triangles, points nearest the troughs of the oscillation. **B**, Probability density function of the goodness of fit obtained by fitting the random shuffled data with its best sinusoidal fit. The vertical gray line reports the goodness of fit of the PSS oscillatory model being higher than what expected from chance ($p = 0.042$). **C**, Proportion of vision-leading responses as a function of SOA, pooling together all trials falling within the peak (red triangles) or trough (blue triangles) of the best-fit oscillation in **A**. Thick curves show the psychometric function computed on the datasets; vertical lines report the PSS for the peak and trough trials, in red and blue, respectively. **D**, Bootstrap distribution (10,000 simulations, with replacement) of PSS for peak (red) and trough (blue) trials. The bootstrap t test revealed a significant difference between the two distributions ($p = 0.002$). Asterisks mark the statistical difference between the two distributions (p -value: $0.01 > ** > 0.001$).

frequencies with higher across-subject coherence, the Hotelling T -squared test was extended to all the tested frequencies (3–15 Hz).

Results

Our experiment investigated the bias and sensitivity dynamics in a multisensory TOJ task. Fig. 1 shows the psychometric functions for the TOJ task (proportion of trials where vision was seen as leading, as a function of SOA), for all participants ($N = 9$), computed pooling together all trials at all delays from button press (go signal). The slopes of psychometric functions were shallow (JND mean and SEM, 105 ± 24 ms; interquartile range, 58 ms), suggesting the task was demanding for participants and sensitivity poor as previously observed (Arrighi et al., 2006; Binda et al., 2009). Consistent with previous results, participants showed a weak but consistent bias toward “vision first,” with average negative points of subjective synchrony (PSS mean and SEM, -33 ± 14 ms; interquartile range, 29 ms).

Fig. 2A shows the dynamics of PSS as a function of time from action. The PSS was not constant but fluctuated rhythmically as a function of time-from-action, at a frequency of 8.2 Hz, with amplitude of ± 29 ms. A time difference of ~ 50 ms of subjective simultaneity is a large effect for audiovisual signals (Fujisaki et al., 2004). Clear oscillations emerged for 500 ms around action-execution, then attenuating in amplitude. To evaluate the significance of the sinusoidal fit, we compared the R -squared values of the best fit with the distribution of R -squared obtained by fitting the best sinusoidal wave form to surrogate data. Fig. 2B shows that the goodness of fit of the PSS model was statistically higher than that expected from a noise distribution (R -squared = 0.4, $p = 0.042$).

As a further test of the significance of the modulation, we selected all the trials in the bins at the predicted peaks or troughs of the best-fitting oscillation (Fig. 2A). As shown in Fig. 2C, the two independent psychometric functions, one for each selected dataset, had a difference in PSS of ~ 30 ms, but the same slope (i.e., sensitivity).

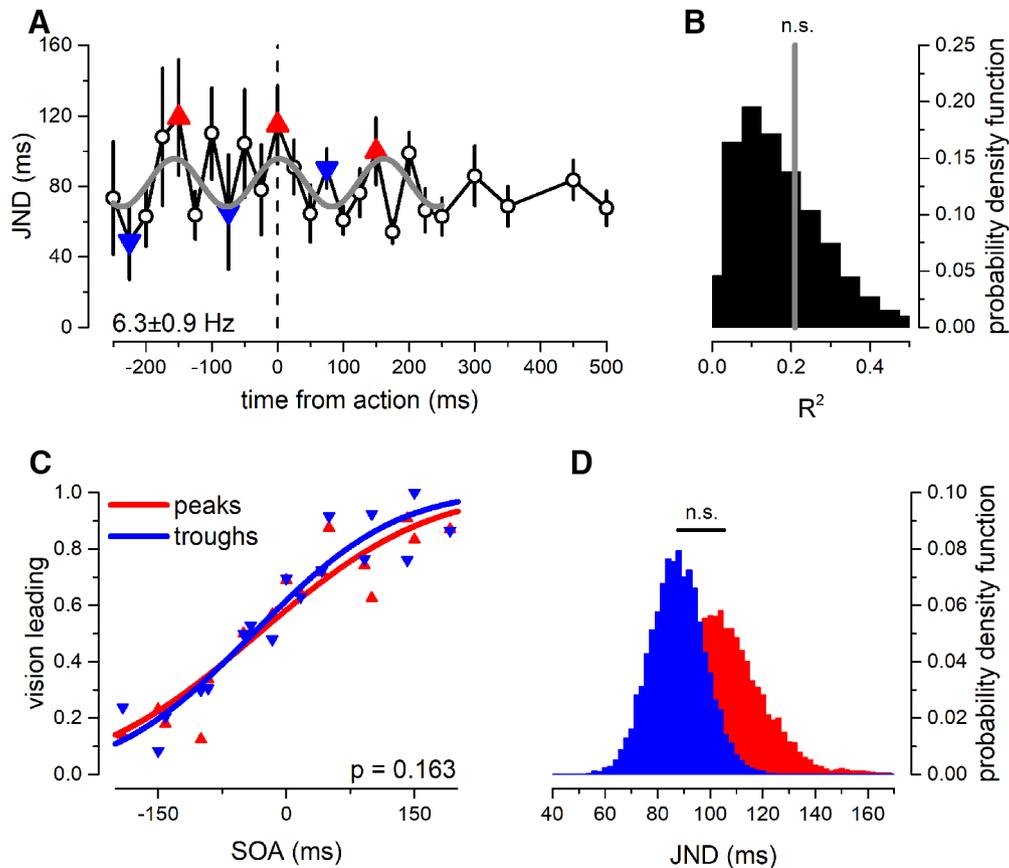


Figure 3. Dynamics of JND for aggregate observer ($N = 9$). **A**, $JND \pm 1$ SEM as a function of time from action-execution. The gray line reports the best sinusoidal fit, of frequency 6.3 Hz. Red upright triangles, points nearest to the peaks of the oscillation; blue inverted triangles, points nearest to the troughs of the oscillation. **B**, Goodness-of-fit probability distribution obtained by fitting the random shuffled data with its best sinusoidal fit. The gray line reports the goodness-of-fit of the JND oscillatory model that does not exceed the chance-level threshold ($p = 0.3$). **C**, Proportion of vision-leading response as a function of SOA, pooling together all trials falling within the peak (red triangles) or trough (blue triangles) of the best-fit oscillation in **A**. Thick curves show the psychometric function computed on the datasets. **D**, Bootstrap distribution (10,000 simulations, with replacement) of JND for peak (red) and trough (blue) trials. The bootstrap t test revealed that difference between the two distributions was not significant ($p = 0.163$).

The difference in PSS was significant ($p = 0.002$, bootstrap sign-test with 10,000 reiterations; Fig. 2D).

Fig. 3A reports the dynamics of the JND of the psychometric function, again as a function of time from action-execution. No significant sinusoidal fit was obtained: the best fit was at 6.3 Hz, but this was not significant (amplitude = 13 ms; R-squared = 0.21, $p = 0.03$; Fig. 3B). Even after pooling all trials of the bins nearest to the predicted peak or trough of the best-fit oscillation, there was no significant difference in the slope of the psychometric functions ($p = 0.163$; Fig. 3D).

We further investigated the frequency spectrum of the bias dynamics with a single-trial multivariate glm analysis, which allowed us to eliminate the binning procedure. As an indirect index of bias, we computed the proportion of “vision leading” responses as a function of time from action (Fig. 4D), for all SOAs between visual and auditory stimuli less than ± 100 ms and within a temporal range of -450 to 250 ms from action execution (see Methods). Fig. 4A reports the result of this analysis for the aggregate observer, run on the probability of reporting “vision leading,” with time aligned with the timing of the first stimulus

of the audiovisual pair (black thick curve), the middle of the pair (red dotted curve), or the last (blue dashed curve). When aligning the action with the presentation of the first stimulus, PSS showed strong and statistically significant theta oscillations at 7.6 Hz (amplitude and phase: 0.034 and 216° ; $p = 0.012$, corrected for multiple comparison: see Methods). When the same analysis was run with responses aligned to the onset of the last stimulus, or with the stimulus midline, no statistical significant oscillations emerged in the PSS ($p > 0.05$, corrected for multiple comparisons). This suggests that the oscillations in PSS modulate the time of the first sensory event.

The results obtained from the multivariate glm analysis on single trials are consistent with those analyzing the best fits of the binned time-series. The two methods predicted similar sinusoidal oscillations in frequency, amplitude, and phase as shown by the gray and green curves of Fig. 4D, showing the separate estimates of predicted oscillations: gray line, best-fit analysis, adapted from Fig. 2A; green line, single-trial frequency analysis, from Fig. 4B). Both functions fit well the oscillation of the proportion of trials

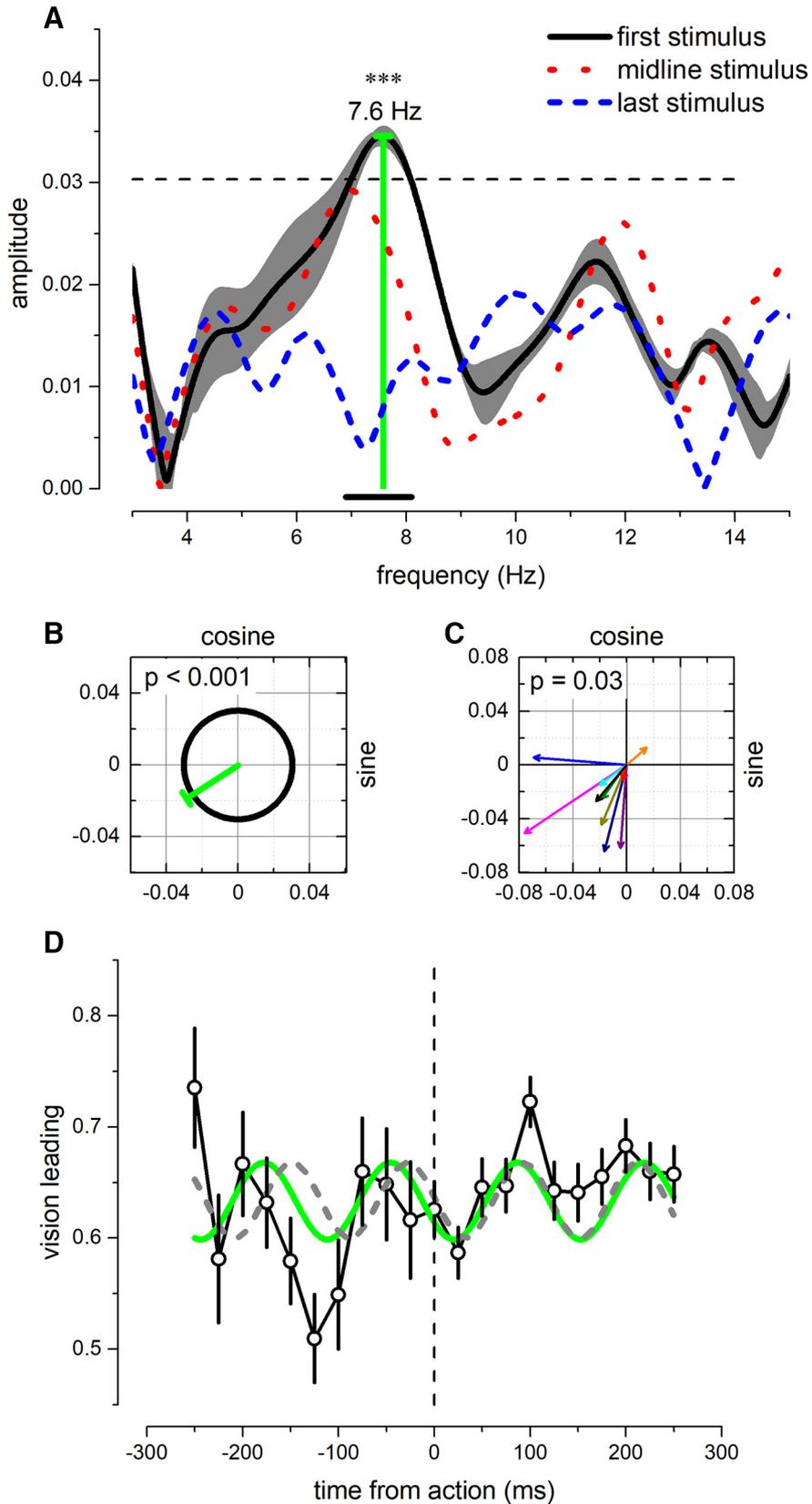


Figure 4. A, Multivariate glm analysis for the aggregate observer (for the interval range within -450 and 250 ms from action execution). The thick black curve shows the amplitude modulation as a function of frequency, for trials aligned to the first stimulus; the gray shadow reports the interquartile range of the amplitude error estimated with a jackknife resampling; dotted red and dashed blue curve represent the amplitude for trials aligned to midline or last stimulus, respectively. Only frequencies ~7.6 Hz for trials aligned to the

continued

first stimulus were significant (vertical green line; uncorrected $p < 0.001$, corrected $p = 0.013$). Horizontal dashed line marks the 95th percentile of noise distribution computed via the maximum statistics for permuted data. **B**, Phase distribution of the strongest frequency above noise level, at 7.6 Hz. Green and black lines report the norm of the betas computed for 7.6 Hz and the 95th percentile of noise distribution calculated for the best sinusoidal fitted frequency for each permutation (see Methods). **C**, Results from the group-mean analysis computed for the aggregate observer's significant frequency of 7.6 Hz. Colored arrows report single subjects' beta distribution (sine and cosine); black arrow shows the group-mean betas. The vectors scatter around the third quadrant, suggesting strong phase-coherence across participants. The Hotelling T^2 revealed that the beta distributions were statistically significant ($p = 0.03$). **D**, Proportion (± 1 SEM) of vision leading response as a function of time from action ($N = 9$). Gray curve shows the reverse best-fit sine model estimated on the PSS, from the dataset of Fig. 2A; green curve reports the function obtained by the multivariate glm methods shown in **A**.

reported as "vision leading," calculated from the independent bins of 25 ms.

Finally, we performed an additional control analysis on single subjects to test whether the aggregate-observer results could have been driven by a few participants with strong phase-locked oscillations in bias. We selected the frequency of 7.6 Hz, suggested by the aggregate observer analysis (Fig. 4A). For each subject, we estimated the sine and cosine components via glm analysis and computed the Hotelling t -squared statistic to assess the statistical significance of the beta distribution (see Methods). Fig. 4C shows the results of this analysis. Confirming and validating the aggregate-observer results, we found a consistent phase-locking within subjects at ~ 7.6 Hz (Hotelling $t_8^2 = 12.6$, $p = 0.036$). The average amplitude and phase of the individual subject's vectors at 7.6 Hz were 0.035 and 228° , respectively, very similar to that of the aggregate observer. Globally, these results indicate the presence of a strong theta oscillation in bias for audiovisual TOJ, synchronized with the action onset and encompassing it, showing very similar temporal dynamics across participants.

To better assess the interindividual differences, we also performed a glm analysis on single subjects. The spectrum of the group-mean analysis is reported in Fig. 5A, which shows, for each tested frequency, the length of the mean vector across participants. Confirming the aggregate-observer result, the spectrum shows a peak at ~ 7.6 Hz, indicating a general increase in phase coherence for that frequency. To further confirm the presence of a single oscillatory spot of modulation in the bias, the Hotelling t -squared statistic was extended to all possible frequencies (Fig. 5C). The t -squared distribution confirms the aggregate-observer results, indicating a group-mean phase locking at ~ 7.6 Hz only. Fig. 5B additionally shows single subject modulations of the z-scored proportion of visual leading responses for 6 representative subjects. These plots confirm the well-known subject-by-subject variability previously reported for behavioral oscillations (Fiebelkorn et al., 2013; Tomassini et al., 2015; Ho et al., 2017), but nevertheless show similar oscillatory modulation for each participant. The best sinusoidal fit revealed a mean frequency and interquartile range of 6.6 ± 2.5 Hz, with R -squared of 0.269 ± 0.13 . Importantly, it is fundamental to point out that despite the overall interindividual differences, our analysis clearly shows a consistent phase-locked oscillation in bias, for all 9 participants.

Discussion

We measured sensitivity and bias dynamics in temporal order judgment of transient audiovisual stimuli presented around the time of executing a voluntary action (i.e., button-press). We found that audiovisual temporal bias oscillates in synchrony with action-execution, in the high-theta range (at ~ 7 –8 Hz), while no significant oscillations were observed for sensitivity of temporal order. The oscillations in bias were phase-locked with the first stimulus of the temporal sequence, not with the second or the barycenter of the two. The temporal rhythmicities commence ~ 500 ms before action-execution and last for up to 250 ms after button-press, with strong phase coherence across subjects. Interestingly, the phase of the prominent theta oscillation in bias reveals a continuous dynamic that encompasses the action onset, with no phase-reset or frequency modulation at the actual action onset, suggesting a close link between the programming signals that precedes action-execution and perceptual processing.

Despite the strong phenomenological impression of a unitary sense of time, it is not uncommon for the brain to deal with multiple processing speeds in its computations: time has been shown to vary across sensory modalities and features of the sensory stimulation (Johnston et al., 2006; Kanai et al., 2006; Burr et al., 2011; Harrington et al., 2011; Tomassini et al., 2011). Considering all these modality- and input-specific clocks or oscillators in the brain, how do we achieve a stable and unitary sense of time? It has been shown that time perception and motor timing rely on similar cerebral structures (Schubotz et al., 2000; Nobre and O'Reilly, 2004), suggesting that motor systems play a key role in shaping a unitary sense of time, possibly by synchronizing the dynamics of local processing. According to this hypothesis, brain oscillations strongly correlate with time perception (Varela et al., 1981; Large and Jones, 1999; Barnes and Jones, 2000; Herrmann et al., 2013; Samaha and Postle, 2015; Benedetto et al., 2017; Ronconi et al., 2017).

Recent electrophysiological and behavioral experiments have shown that the motor system and perceptual rhythms are synchronized over time (Gaarder et al., 1966; Tomassini et al., 2015, 2017; Benedetto et al., 2016; Benedetto and Morrone, 2017; Hogendoorn, 2016; Tomassini and D'Ausilio, 2018; Wutz et al., 2016). Neural oscillations have been proposed to reflect ongoing modulation of cortical excitability (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Jensen et al., 2012, 2014). In agreement with this hypothesis, several studies have shown

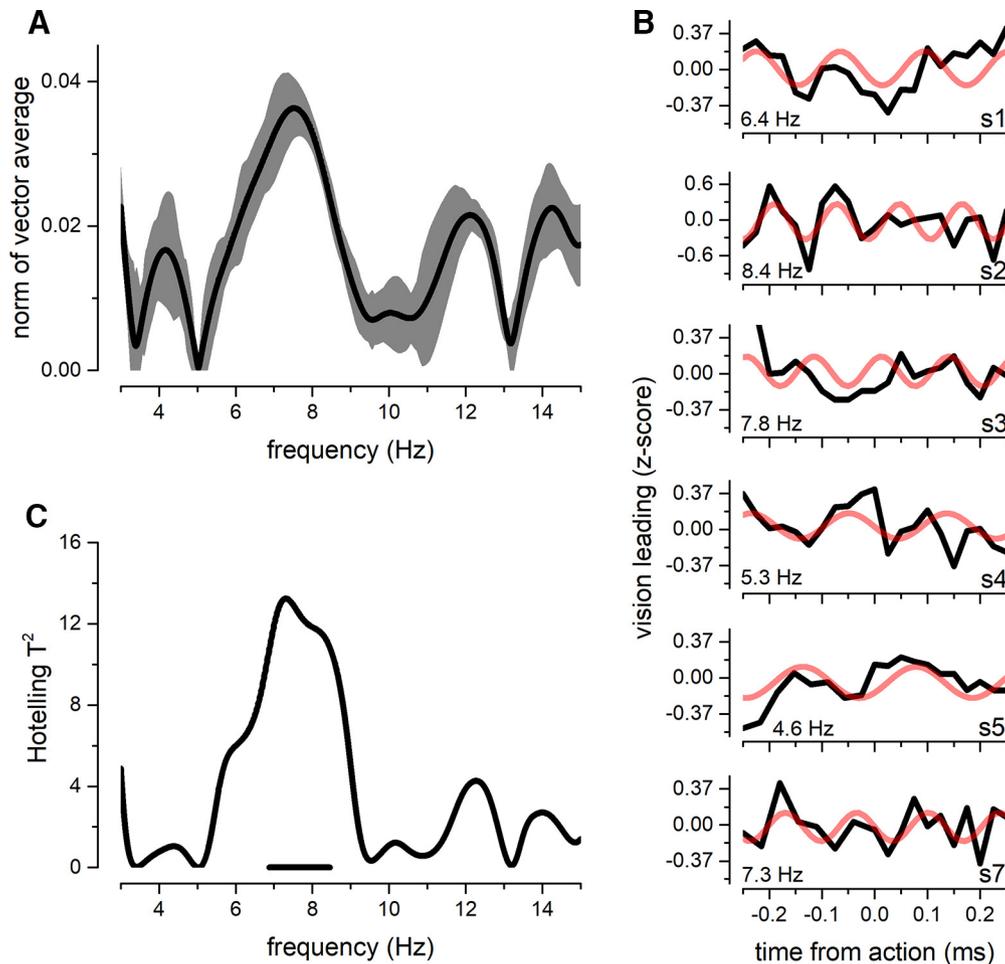


Figure 5. **A**, Norm of the group-mean average vector obtained for the proportion of vision leading time courses. Shaded gray area reports the interquartile range obtained from jackknife resampling procedure. **B**, Single-subject fluctuations in z-scored vision leading response for six representative subjects. The red line shows the best sinusoidal fit for each individual participant. **C**, Hotelling T -squared distribution for all the tested frequencies. The only significant spot of coherence is ~ 7.6 Hz. The horizontal thick line reports the significant frequencies ($p < 0.05$).

that ongoing oscillations in the theta-alpha range are related to fluctuations in perceptual sensitivity across different modalities (Ng et al., 2012; Busch et al., 2009; Fiebelkorn et al., 2013; Milton and Pleydell-Pearce, 2016; Craddock et al., 2017). It has been demonstrated that cross-modal interactions are also achieved by synchronization of cortical neural oscillations in the delta, theta, and gamma frequencies (Lakatos et al., 2007), and that rhythmic synchronization between different brain areas might serve to bind cross-modal information (Senkowski et al., 2008).

In our experiment, we investigated audiovisual temporal order judgments, reporting oscillations in perceptual bias—in the theta range—phase-locked with the time of action-execution (button-press). Theta oscillations have been proposed to mediate human anticipatory mechanisms (predictive priors), by modulating faster neuronal synchronization and facilitating neuronal communication among distant brain areas (Lakatos et al., 2008; Stefanics et al., 2010). A recent experiment has shown that the phase of the prestimulus occipital alpha oscillations can predict the subjects' perceptual decision in a detection

task (Sherman et al., 2016): the oscillations in temporal bias reported here may indeed reflect rhythmic, top-down influences of perceptual predictions, modulating our decisions according to the timing of action execution. It is noteworthy that (in our paradigm) we are able to record behavioral oscillations only if these rhythmicities are phase-locked with the motor execution. The results show not only that perceptual decisions oscillate in synchrony with the button-press, but also that these oscillations precede the action execution (by about half a second), indicating a synchronization between our perceptual decisions and the intention to move. Furthermore, the strong phase-coherence reported across participants suggests not only that this oscillatory dynamic is very precise, but also that it is likely to reflect a general sensorimotor rhythm in the brain, pointing to the presence of a single shared clock driving both perception and action.

The dynamics of the modulation in criterion shows a local maximum of decision bias just before and after the action execution (Figs. 2A and 4D, as indicated by maximal distance between the oscillatory model and the ideal no-bias point), with reduced perceptual bias around the

time of action-execution (0–100 ms from button-press). We speculate that this precise sensorimotor coordination has the goal of minimizing prior interference around the time of action-execution, when the subject is actively interacting with the environment and likely to be producing perceptual events. On the contrary, during action planning and motor-induced suppression periods (around –50 and 100 ms from button press, respectively), the weight of priors seems to be maximum, resulting in a stronger bias. Furthermore, these oscillations were phase-locked with the first stimulus of the sequence; no oscillations were observable when the action was aligned with the last or the midline stimulus. This fact serves as an important sanity check for the robustness of our spectral analysis, as well as speaking to the dynamics of the oscillatory process under description. The results are consistent with the idea that oscillations in bias mostly affect the first stimulus, which may in turn induce a reorganization of local-global brain oscillations (Romei et al., 2012).

We observed oscillations in temporal bias, but no oscillations in temporal sensitivity. The scientific literature on perceptual cycles has mainly focused on the effect of oscillations on sensitivity and reaction time (VanRullen, 2016). However, a recent experiment has shown that sensitivity is not the only perceptual feature that exhibits rhythmic dynamics, but that decision criteria also oscillate (Ho et al., 2017). Crucially, the two dynamics seem to be largely independent, oscillating at different frequencies and phases (Ho et al., 2017), suggesting the presence of two distinct driving mechanisms shaping the rhythms of perception. This idea is supported by several electrophysiological studies showing that alpha and theta oscillations are instrumental in transmitting feedback signals in the brain (Pastoll et al., 2013; van Kerkoerle et al., 2014; Bastos et al., 2015; Jensen et al., 2015; Andreou et al., 2017). It has also been shown that alpha phase and power can specifically predict decision (Sherman et al., 2016; Benwell et al., 2017; Craddock et al., 2017). Our results are consistent with the idea of two separated oscillatory mechanisms for sensitivity and bias. Sensitivity oscillation may be linked with local modulations of cortical excitability (Klimesch et al., 2007; Jensen and Mazaheri, 2010; Jensen et al., 2012, 2014), while decision criteria oscillations may be driven by top-down predictive mechanisms (Sherman et al., 2016; Ho et al., 2017). We further show here that oscillations in decision can be synchronized with voluntary action, emerging around half a second before executing a movement. However, there is still no conclusive evidence about the causal direction of the synchronization between action and perception: either the intention-to-move signal arrives in perceptual areas and synchronizes oscillations there; or both perceptual and motor actions are gated by an endogenous neural rhythm. Both mechanisms are equally possible.

To conclude, we demonstrate that audiovisual temporal bias, but not sensitivity, oscillates in the theta range, synchronized with action execution. These oscillations are instantiated hundreds of milliseconds before the movement onset, indicating that the organization of a voluntary action is directly linked to the way we process sensory

information about temporal order. These rhythmic influxes are very precise in time and similar across subjects, suggesting a fundamental functional role in shaping our criterion over time and pointing to the presence of a single shared clock between perception and action.

References

- Andreou C, Frielinghaus H, Rauh J, Mußmann M, Vauth S, Braun P, Leicht G, Muler C (2017) Theta and high-beta networks for feedback processing: a simultaneous EEG-fMRI study in healthy male subjects. *Transl Psychiatry* 7:e1016. [CrossRef Medline](#)
- Arrighi R, Alais D, Burr D (2006) Perceptual synchrony of audiovisual streams for natural and artificial motion sequences. *J Vis* 6:6. [CrossRef](#)
- Barnes R, Jones MR (2000) Expectancy, attention, and time. *Cogn Psychol* 41:254–311. [CrossRef Medline](#)
- Bastos AM, Vezoli J, Bosman CA, Schoffelen J-M, Oostenveld R, Dowdall JR, De Weerd P, Kennedy H, Fries P (2015) Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* 85:390–401. [CrossRef Medline](#)
- Benedetto A, Lozano-Soldevilla D, Vanrullen R (2017) Ambient luminance changes modulate oscillatory properties of the visual system. *J Vis* 17:724. [CrossRef](#)
- Benedetto A, Morrone MC (2017) Saccadic suppression is embedded within extended oscillatory modulation of sensitivity. *J Neurosci* 37:3661–3670. [CrossRef Medline](#)
- Benedetto A, Spinelli D, Morrone MC (2016) Rhythmic modulation of visual contrast discrimination triggered by action. *Proc Biol Sci* 283:3536–3544. [CrossRef](#)
- Benwell CSY, Keitel C, Harvey M, Gross J, Thut G (2017) Trial-by-trial co-variation of pre-stimulus EEG alpha power and visuospatial bias reflects a mixture of stochastic and deterministic effects. *Eur J Neurosci* doi: 10.1111/ejn.13688.
- Binda P, Cicchini GM, Burr DC, Morrone MC (2009) Spatiotemporal distortions of visual perception at the time of saccades. *J Neurosci* 29:13147–13157. [CrossRef Medline](#)
- Burr DC, Cicchini GM, Arrighi R, Morrone MC (2011) Spatiotopic selectivity of adaptation-based compression of event duration. *J Vis* 11:21–21. [CrossRef](#)
- Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29:7869–7876. [CrossRef Medline](#)
- Cecere R, Rees G, Romei V (2015) Individual differences in alpha frequency drive crossmodal illusory perception. *Curr Biol* 25:231–235. [CrossRef Medline](#)
- Craddock M, Poliakoff E, El-derey W, Klepousniotou E, Lloyd DM (2017) Pre-stimulus alpha oscillations over somatosensory cortex predict tactile misperceptions. *Neuropsychologia* 96:9–18. [Cross-Ref Medline](#)
- de Lange FP, Rahnev DA, Donner TH, Lau H (2013) Prestimulus oscillatory activity over motor cortex reflects perceptual expectations. *J Neurosci* 33:1400–1410. [CrossRef Medline](#)
- Drewes J, VanRullen R (2011) This is the rhythm of your eyes: the phase of ongoing electroencephalogram oscillations modulates saccadic reaction time. *J Neurosci* 31:4698–4708. [CrossRef Medline](#)
- Engel AK, Singer W (2001) Temporal binding and the neural correlates of sensory awareness. *Trends Cogn Sci* 5:16–25. [Medline](#)
- Fiebelkorn IC, Foxe JJ, Butler JS, Mercier MR, Snyder AC, Molholm S (2011) Ready, set, reset: stimulus-locked periodicity in behavioral performance demonstrates the consequences of cross-sensory phase reset. *J Neurosci* 31:9971–9981. [CrossRef Medline](#)
- Fiebelkorn IC, Saalman YB, Kastner S (2013) Rhythmic sampling within and between objects despite sustained attention at a cued location. *Curr Biol* 23:2553–2558. [CrossRef](#)
- Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci* 9:474–480. [CrossRef Medline](#)

- Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778. [CrossRef Medline](#)
- Gaarder K, Koresko R, Kropfl W (1966) The phasic relation of a component of alpha rhythm to fixation saccadic eye movements. *Electroencephalogr Clin Neurophysiol* 21:544–551. [Medline](#)
- Green D, Swets JA (1966) *Signal Detection Theory and Psychophysics*. New York: Peninsula Pub.
- Haggard P, Clark S, Kalogeras J (2002) Voluntary action and conscious awareness. *Nat Neurosci* 5:382–385. [CrossRef Medline](#)
- Hagura N, Kanai R, Orgs G, Haggard P (2012) Ready steady slow: action preparation slows the subjective passage of time. *Proc Biol Sci* 279:4399–4406. [CrossRef](#)
- Hanslmayr S, Volberg G, Wimber M, Dalal SS, Greenlee MW (2013) Prestimulus oscillatory phase at 7 Hz gates cortical information flow and visual perception. *Curr Biol* 23:2273–2278. [CrossRef Medline](#)
- Harrington DL, Castillo GN, Fong CH, Reed JD (2011) Neural underpinnings of distortions in the experience of time across senses. *Front Integr Neurosci* 5:32. [CrossRef Medline](#)
- Herrmann B, Henry MJ, Grigutsch M, Obleser J (2013) Oscillatory phase dynamics in neural entrainment underpin illusory percepts of time. *J Neurosci* 33:15799–15809. [CrossRef Medline](#)
- Heusser AC, Poeppel D, Ezzyat Y, Davachi L (2016) Episodic sequence memory is supported by a theta-gamma phase code. *Nat Neurosci* 19:1374–1380.
- Ho HT, Leung J, Burr DC, Alais D, Morrone MC (2017) Auditory sensitivity and decision criteria oscillate at different frequencies separately for the two ears. *Curr Biol* 27:3643–3649. [CrossRef Medline](#)
- Hogendoorn H (2016) Voluntary saccadic eye movements ride the attentional rhythm. *J Cogn Neurosci* 28:1625–1635.
- Jensen O, Bonnefond M, Marshall TR, Tiesinga P (2015) Oscillatory mechanisms of feedforward and feedback visual processing. *Trends Neurosci* 38:192–194. [CrossRef Medline](#)
- Jensen O, Bonnefond M, VanRullen R (2012) An oscillatory mechanism for prioritizing salient unattended stimuli. *Trends Cogn Sci* 16:200–206. [CrossRef Medline](#)
- Jensen O, Gips B, Bergmann TO, Bonnefond M (2014) Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci* 37:357–369. [CrossRef Medline](#)
- Jensen O, Mazaheri A (2010) Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front Hum Neurosci* 4:186. [CrossRef Medline](#)
- Johnston A, Arnold DH, Nishida S (2006) Spatially localized distortions of event time. *Curr Biol* 16:472–479. [CrossRef Medline](#)
- Kanai R, Paffen CLE, Hogendoorn H, Verstraten FAJ (2006) Time dilation in dynamic visual display. *J Vis* 6:8. [CrossRef](#)
- Klimesch W, Sauseng P, Hanslmayr S (2007) EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Res Rev* 53:63–88. [Cross-Ref Medline](#)
- Lakatos P, Chen C-M, O’Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53:279–292. [CrossRef](#)
- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–113. [CrossRef Medline](#)
- Landau AN, Fries P (2012) Attention samples stimuli rhythmically. *Curr Biol* 22:1000–1004. [CrossRef Medline](#)
- Lansing RW (1957) Relation of brain and tremor rhythms to visual reaction time. *Electroencephalogr Clin Neurophysiol* 9:497–504. [Medline](#)
- Large EW, Jones MR (1999) The dynamics of attending: how people track time-varying events. *Psychol Rev* 106:119–159. [CrossRef](#)
- Linkenkaer-Hansen K, Nikulin VV, Palva S, Ilmoniemi RJ, Palva JM (2004) Prestimulus oscillations enhance psychophysical performance in humans. *J Neurosci* 24:10186–10190. [CrossRef Medline](#)
- Milton A, Pleydell-Pearce CW (2016) The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *Neuroimage* 133:53–61. [CrossRef Medline](#)
- Morrone MC, Ross J, Burr DC (2005) Saccadic eye movements cause compression of time as well as space. *Nat Neurosci* 8:950–954. [CrossRef Medline](#)
- Ng BSW, Schroeder T, Kayser C (2012) A precluding but not ensuring role of entrained low-frequency oscillations for auditory perception. *J Neurosci* 32:12268–12276. [CrossRef](#)
- Nobre AC, O’Reilly J (2004) Time is of the essence. *Trends Cogn Sci* 8:387–389. [CrossRef Medline](#)
- Park J, Schlag-Rey M, Schlag J (2003) Voluntary action expands perceived duration of its sensory consequence. *Exp Brain Res* 149:527–529. [CrossRef Medline](#)
- Pastoll H, Solanka L, van Rossum MCW, Nolan MF (2013) Feedback inhibition enables theta-nested gamma oscillations and grid firing fields. *Neuron* 77:141–154. [CrossRef](#)
- Romei V, Gross J, Thut G (2012) Sounds reset rhythms of visual cortex and corresponding human visual perception. *Curr Biol* 22:807–813.
- Ronconi L, Oosterhof NN, Bonmassar C, Melcher D (2017) Multiple oscillatory rhythms determine the temporal organization of perception. *Proc Natl Acad Sci U S A* 114:13435–13440. [CrossRef Medline](#)
- Samaha J, Postle BR (2015) The speed of alpha-band oscillations predicts the temporal resolution of visual perception. *Curr Biol* 25:2985–2990.
- Schubotz RI, Friederici AD, von Cramon DY (2000) Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage* 11:1–12. [CrossRef Medline](#)
- Senkowski D, Schneider TR, Foxe JJ, Engel AK (2008) Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci* 31:401–409. [CrossRef Medline](#)
- Sherman MT, Kanai R, Seth AK, VanRullen R (2016) Rhythmic influence of top–down perceptual priors in the phase of prestimulus occipital alpha oscillations. *J Cogn Neurosci* 28:1318–1330. [Cross-Ref Medline](#)
- Stefanics G, Hangya B, Hernádi I, Winkler I, Lakatos P, Ulbert I (2010) Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J Neurosci* 30:13578–13585. [CrossRef](#)
- Surwillo WW (1961) Frequency of the “alpha” rhythm, reaction time and age. *Nature* 191:823–824. [CrossRef](#)
- Tomassini A, Ambrogioni L, Medendorp WP, Maris E (2017) Theta oscillations locked to intended actions rhythmically modulate perception. *eLife* 6:e25618.
- Tomassini A, D’Ausilio A (2018) Passive sensorimotor stimulation triggers long lasting alpha-band fluctuations in visual perception. *J Neurophysiol* 119:380–388. [CrossRef](#)
- Tomassini A, Gori M, Baud-Bovy G, Sandini G, Morrone MC (2014) Motor commands induce time compression for tactile stimuli. *J Neurosci* 34:9164–9172. [CrossRef Medline](#)
- Tomassini A, Gori M, Burr D, Sandini G, Morrone MC (2012) Active movement restores veridical event-timing after tactile adaptation. *J Neurophysiol* 108:2092–2100. [CrossRef Medline](#)
- Tomassini A, Gori M, Burr D, Sandini G, Morrone MC (2011) Perceived duration of visual and tactile stimuli depends on perceived speed. *Front Integr Neurosci* 5:51. [CrossRef Medline](#)
- Tomassini A, Morrone MC (2016) Perceived visual time depends on motor preparation and direction of hand movements. *Sci Rep* 6:27947. [CrossRef Medline](#)
- Tomassini A, Spinelli D, Jacono M, Sandini G, Morrone MC (2015) Rhythmic oscillations of visual contrast sensitivity synchronized with action. *J Neurosci* 35:7019–7029. [CrossRef Medline](#)
- van Kerkoerle T, Self MW, Dagnino B, Gariel-Mathis M-A, Poort J, van der Togt C, Roelfsema PR (2014) Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc Natl Acad Sci U S A* 111:14332–14341. [CrossRef Medline](#)
- VanRullen R (2016) Perceptual cycles. *Trends Cogn Sci* 20:723–735. [CrossRef Medline](#)
- VanRullen R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213. [CrossRef](#)

- Varela FJ, Toro A, John ER, Schwartz EL (1981) Perceptual framing and cortical alpha rhythm. *Neuropsychologia* 19:675–686. [Medline](#)
- Wood DK, Gu C, Corneil BD, Gribble PL, Goodale MA (2015) Transient visual responses reset the phase of low-frequency oscillations in the skeletomotor periphery. *Eur J Neurosci* 42:1919–1932. [CrossRef](#) [Medline](#)
- Wutz A, Muschter E, van Koningsbruggen MG, Weisz N, Melcher D (2016) Temporal integration windows in neural processing and perception aligned to saccadic eye movements. *Curr Biol* 26:1659–1668. [CrossRef](#) [Medline](#)
- Yarrow K, Rothwell JC (2003) Manual chronostasis: tactile perception precedes physical contact. *Curr Biol* 13:1134–1139. [Medline](#)