Predictive visuo-motor communication through neural oscillations

Alessandro Benedetto[1], Paola Binda[1], Mauro Costagli[2,3], Michela Tosetti[2,3] & Maria Concetta Morrone[1,2*]

Author affiliation

[1]Department of Translational Research on New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy.

[2]IRCCS Stella Maris, Calambrone, Pisa, Italy.

[3]Imago 7 Research Foundation, Calambrone, Pisa, Italy.

[*]Correspondence: concetta.morrone@unipi.it

Summary

The mechanisms coordinating action and perception over time are poorly understood. The sensory cortex needs to prepare for upcoming changes contingent on displacement caused by action, and this requires temporally precise communication that takes into account the variable delays between sensory and motor processing. Here we demonstrate that synchronization of endogenous oscillatory neuronal activities can facilitate sensorimotor communication between distal brain areas. We show that a simple voluntary action synchronizes theta-range oscillations, which are revealed in rhythmic modulation of V1 BOLD response as well as in visual sensitivity. V1-M1 functional connectivity also oscillates at the same theta rhythm, and M1 BOLD signals during motor preparation predict V1 BOLD activity for several seconds during and after the action. Thus each voluntary action causes theta-range oscillations in V1 response, which last for hundreds of milliseconds, well after action completion. This mechanism has the potential to communicate the time of action set by the motor cortex to the sensory cortex, and to encode it with high precision over the extended period when relevant sensory signals may arise. By demonstrating predictive temporal coupling between primary motor and sensory cortices, our results strongly implicate synchronization of oscillatory activity patterns in mediating the precise sensorimotor coordination required for efficient interactions with the external world.

Introduction

Perception and action are usually conceived as separate systems, but in everyday life the two need to be continuously coordinated. What neural mechanisms mediate the effortless synchronization of perception and action in space and time? Von Helmholtz[1] introduced the idea that a representation of the motor action (later termed "corollary discharge"[2] or "efference copy"[3]) is copied to sensory areas to inform perceptual systems of the upcoming action. These and later models assume that perception is an active process, shaped by the continual flow of information between motor and sensory systems[4,5]. But how the brain solves the many difficult problems intrinsic to sensory-motor synchronization remains an open question. The synchronization problem is exacerbated by the fact that the flow of temporal information in the sensory-motor network has variable processing speeds and delays, which can vary considerably with the features of the sensory stimulation or the pattern of the programmed movement. In order to account for this variability, sensory-motor communication needs to rely on a temporally precise, yet flexible protocol.

Several theorists have proposed synchronization of endogenous oscillatory activity as the basis for an efficient and flexible communication protocol between distal brain areas[6], a concept known as "communication through coherence". This view assigns an important functional role to the ongoing periodic variations of activity or excitability commonly observed in most regions of the brain[7,8]. Endogenous oscillations span multiple frequency ranges, with low-frequency oscillations (theta and alpha ranges, 4 to 12 Hz) playing a critical role in setting the phase and amplitude of the higher frequency rhythms[9-14].

Recent data demonstrate that phase and amplitude of endogenous neural oscillations impact on perception[15,16], suggesting that they play an important role in coordinating behaviour. In order to reveal their impact on sensitivity and response times, the endogenous oscillations need to be synchronized to allow averaging of responses over many trials. One way to synchronize the oscillations is to present a salient sensory stimulus, such as a flash or a sound[17-21]. Another way is by asking participants to make a voluntary action[22-28]. Oscillations revealed by action start before action-onset and continue for more than a second after completion, implying that action planning changes perception of future sensory stimuli. This could provide a mechanism for predictive communication between cortical areas dedicated to motor-planning and perception, by periodic modulation of excitability (for review [29,30]). Assuming that the frequency of oscillation is known, sensory areas could decode the delay from action-onset at any given moment from the level of excitability, allowing for flexible and precise control over the

coordination of motor and sensory signals. Action planning could transmit a prediction of the upcoming change to the sensory cortex, which may act like a *prior* (in the Bayesian sense), to be combined with the sensory data to form a predictive perceptual representation[31-34].

Here we introduce a novel fMRI paradigm to probe fast oscillations, based on the concept of perturbative signals, which overcomes the low temporal resolution of BOLD signals. The assumption is that a synchronized endogenous rhythm will modulate cortical excitability rhythmically, which should be reflected in the BOLD responses to brief stimuli presented at different phases of the oscillation cycle. Taking advantage of the high signal-to-noise ratio of ultra-high field MR imaging, we measured the BOLD response to brief visual stimuli presented at small and variable delays from a voluntary motor act. This time-resolved fMRI method allowed us to demonstrate that motor planning affects the excitability of visual cortex and the predictive coupling between motor and visual cortex.

Results

Oscillation of behavioural sensitivity

We first measured accuracy in a spatial frequency discrimination task outside the scanner, to reveal behavioural oscillations in sensitivity after a voluntary action. Figure 1 plots percentage of correct responses as a function of time of stimulus presentation after participant keypress, which initiated the trial. Accuracy modulates rhythmically, and is well fit by a sinusoidal function of 5.4 Hz ($R^2 = 0.43$, $p_{corrected} = 0.005$: inset Figure 1B). Note that the first minimum in accuracy occurs 70 ms after action onset, which has been referred to as "motor-induced suppression"[35]. However, the suppression repeats periodically, generating an oscillatory pattern in the theta range. This result corroborates previous findings of behavioural oscillations in orientation and contrast discrimination tasks synchronized with the onset of action[22,24,27].

Oscillation of evoked BOLD responses in visual cortex

We repeated the psychophysical experiment in an ultra-high magnetic field (7T) scanner, measuring the BOLD response evoked by a keypress followed by the visual grating. Although the visual stimuli were very brief (33 ms), the BOLD response in V1 was strong and reliable, both when the stimuli were presented after the keypress action (*vision and action*), and when they were presented in a more typical condition without the preceding keypress (*vision only*). However, the time-courses of the two responses were qualitatively different (Figure 2A). For stimuli not preceded by the action, the BOLD

response followed the typical V1 haemodynamic (Figure 2A, black symbols); for stimuli preceded by the keypress action, the BOLD response reached a similar peak, but attenuated more rapidly (Figure 2A, blue-green symbols).

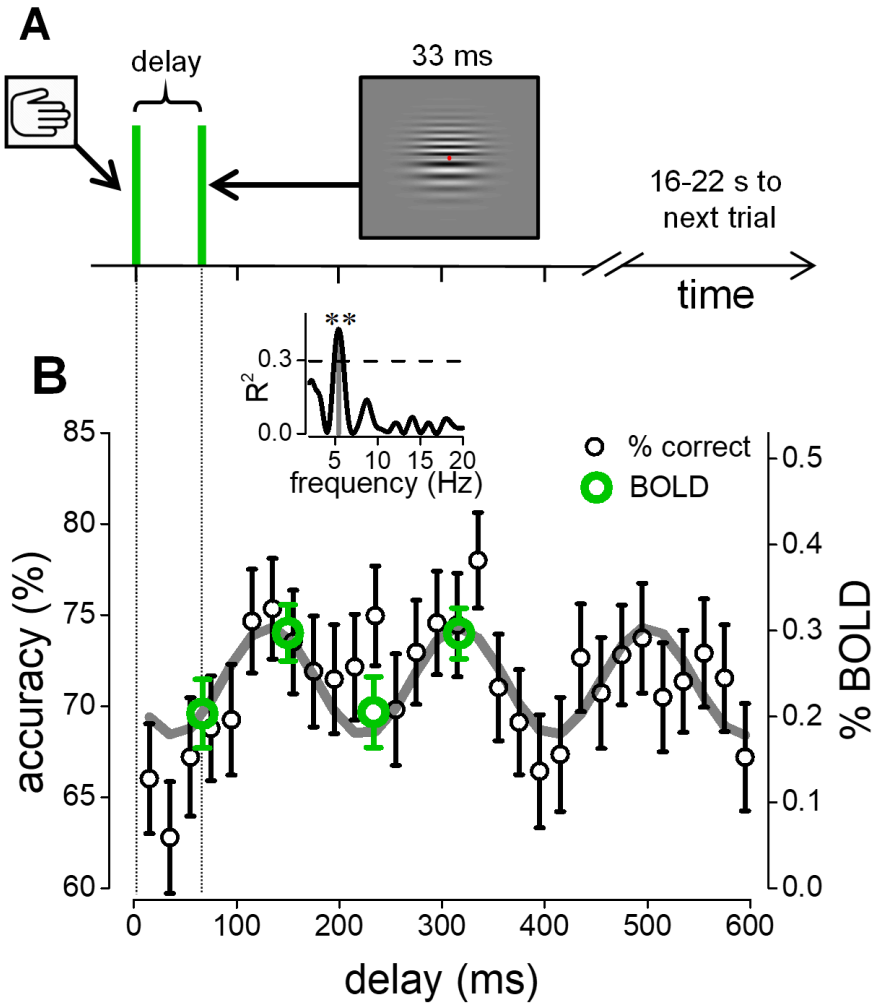

Figure 1. Experimental procedure and behavioural oscillations. A: Schematics of the time-resolved fMRI design measuring BOLD activity in response to visual stimuli elicited by a voluntary action. The delay between stimulus presentation and action onset varied randomly between 0 and 600 ms for the behavioural experiment and between 70, 150, 230 and 310 ms for the fMRI experiment. **B**: Black symbols show accuracy in the spatial frequency discrimination task (mean ± s.e.m.; aggregate observer, N = 7) as a function of visuo-motor delay with the best sinusoidal fit of the accuracy time-course (grey line; 5.4 Hz). Green circles show % BOLD signal change in V1 at the four visuo-motor delays tested (extracted from Figure 2: mean ± s.e.m.; N = 17). **Inset**: Goodness of the sinusoidal fit to the accuracy data in B; frequency was varied between 3 and 20 Hz, yielding a strong and significant peak at 5.4 Hz (R^2 = 0.43, p = 0.005, corrected for multiple comparisons) above the 95th percentile of the R^2 distribution obtained from permuted dataset.

Having assessed in the scanner that the stimuli were near discrimination threshold for all participants (see Methods), we studied how the V1 BOLD response depended on the delay between the voluntary action and stimulus presentation. Stimuli were briefly presented at various delays after keypress, corresponding to minima (70 and 230 ms) or maxima (150 and 310 ms) in accuracy (see Figure 1B). Figure 2B shows average haemodynamic responses to these four stimuli as a function of time. The responses to stimuli at delays with minimal psychophysical accuracy (blue curves) were clearly less than to those at delays with maximal psychophysical accuracy (green curves). To quantify the effect, we calculated for each participant the integral of the timecourse of BOLD responses for the 15 s interval after stimulus presentation. Figure 2C shows the response integral for the four delays. Repeated measures ANOVA showed a significant effect of stimulus delay ($F(3,48) = 4.17$; $p = 0.01$; $\eta^2 = 0.21$), with delays associated with minima in accuracy (70 and 230 ms) significantly lower than those associated with maxima (post-hoc t-tests reported in figure caption). Figure 2D shows individual data for the 17 participants, plotting responses for peak delays against troughs. Despite the large variability in response amplitude, 27 out of 34 points fall above the equality line, showing that the rhythmic modulation of V1 BOLD activity occurred in most participants.

Figure 1B shows the V1 BOLD response amplitudes (green circles) superimposed on the behavioural data collected outside the scanner. There is strong agreement between the results, with the minima and the maxima of BOLD responses coinciding with the minima and maxima of performance in the spatial frequency discrimination task, suggesting that both visual sensitivity and V1 BOLD responses are similarly modulated at about 5 Hz. However, the four data points of BOLD modulation are also consistent with oscillations at twice the frequency (10 Hz). To test and dismiss this possibility, we measured in a subset of participants (N=5) V1 responses at the additional timepoint of 180 ms, mid-way between 150 and 230 ms. If there is modulation at the second harmonic (10 Hz), the response at this delay should be a minimum, similar to the response at 70 ms and significantly lower than that at 150 ms. However, the V1 BOLD response for stimuli at 180 ms was statistically indistinguishable from those for 150 ms ($t(4) = 0.84$, $p = 0.22$, one-tailed paired t-test), and significantly higher than that at 70 ms delay ($t(4) = 4.19$, $p = 0.006$, one-tailed paired t-test). This result is not consistent with higher frequency oscillation in the alpha range.

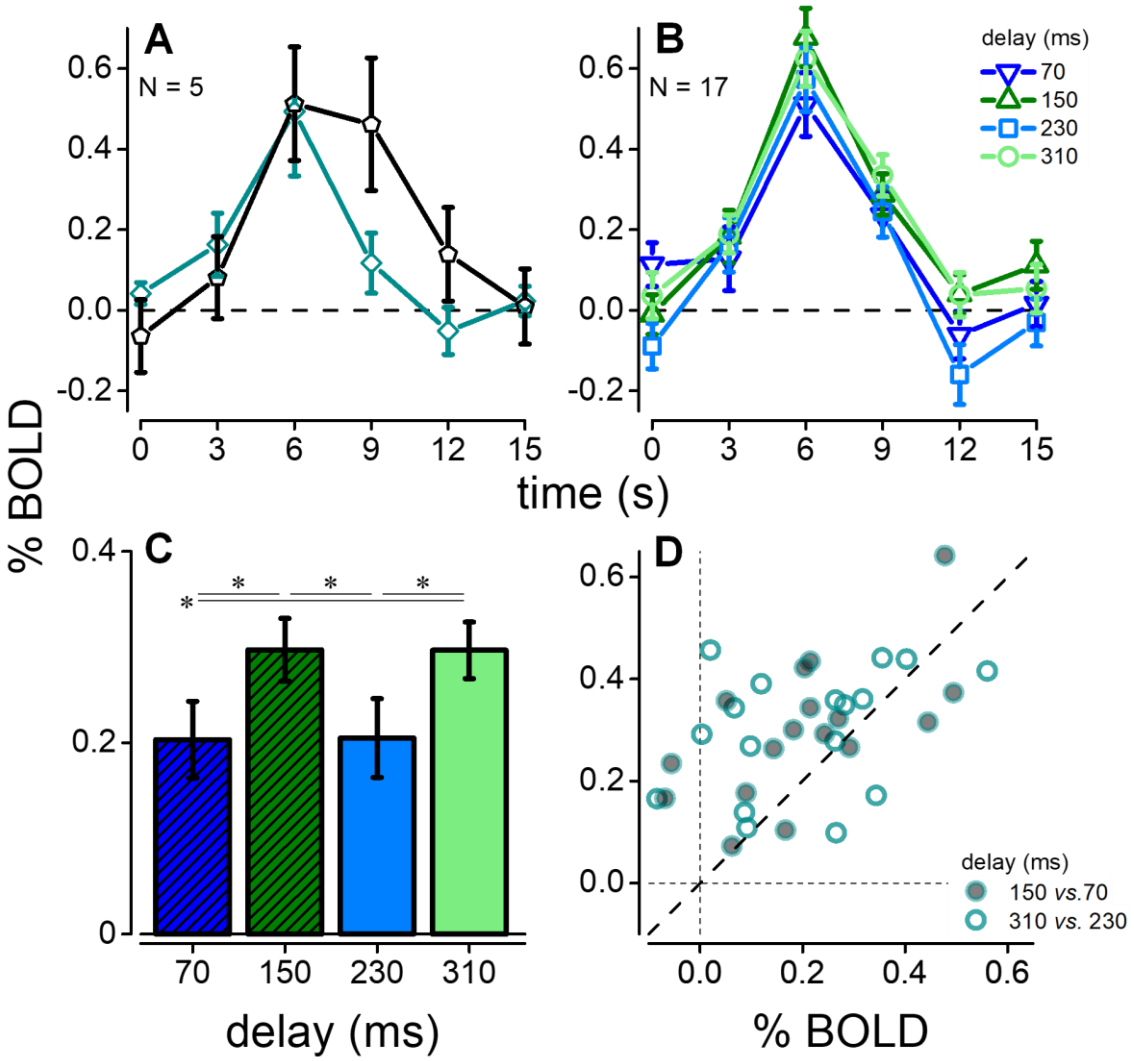

Figure 2. Temporal oscillation in V1 BOLD responses. A: GLM beta weights estimated for the V1 subregion representing the stimulus area; means with s.e.m. from a subset of N=5 participants in the *vision only* (black curve) and *vision and action* condition (blue-green curve). **B**: Timecourse of the GLM beta weights (mean and s.e.m. for all N=17 participants) representing the BOLD response to *vision and action* events in the stimulated V1, separately for the 70, 150, 230 and 310 ms visuo-motor delays. Figure S2 shows the GLM beta weights aligned to stimulus onset and not action onset, obtaining very similar results. **C**: average V1 response amplitude as function of visuo-motor delay. Asterisks mark statistical significance (0.05 > * > 0.01) of post-hoc one-tailed t-tests (Bonferroni-Holm corrected for multiple comparisons) comparing pairs of visuo-motor delays: 70 ms vs 150 ms: $t(16) = 2.88$, $p = 0.01$; 150 ms vs 230 ms: $t(16) = 2.54$, $p = 0.021$; 230 ms vs 310 ms: $t(16) = 2.20$, $p = 0.042$; 70 ms vs 310 ms: $t(16) = 2.33$, $p = 0.033$). **D:** V1 response amplitude in individual participants, plotting delay 150 ms against 70 ms (filled symbols) or delay 310 ms against 230 ms (empty symbols).

Figure 3A shows the BOLD responses at the four delays mapped onto the posterior cortical surface. Consistent with the transient nature of the stimuli, BOLD activity at all delays is particularly strong in the V1/V2 region, with additional foci of activity in V3, V4 and the intra-parietal and parieto-occipital sulci. Visual inspection of the maps suggests that V1 activity depends on visuo-motor delay, with BOLD responses at the 150 ms and 310 ms delays stronger and more diffuse than at 70 or 230. This is quantified in Figure 3B, which shows that the effect of delay is seen across all V1 eccentricities, even beyond the area covered by the stimulus (the bottom icon shows stimulus contrast across eccentricity). This implies that the effect of delay is independent of the stimulus characteristics, and probably driven by automatic and pervasive modulation of the entire early visual cortex, which oscillates rhythmically in time with action onset (see Model Section).

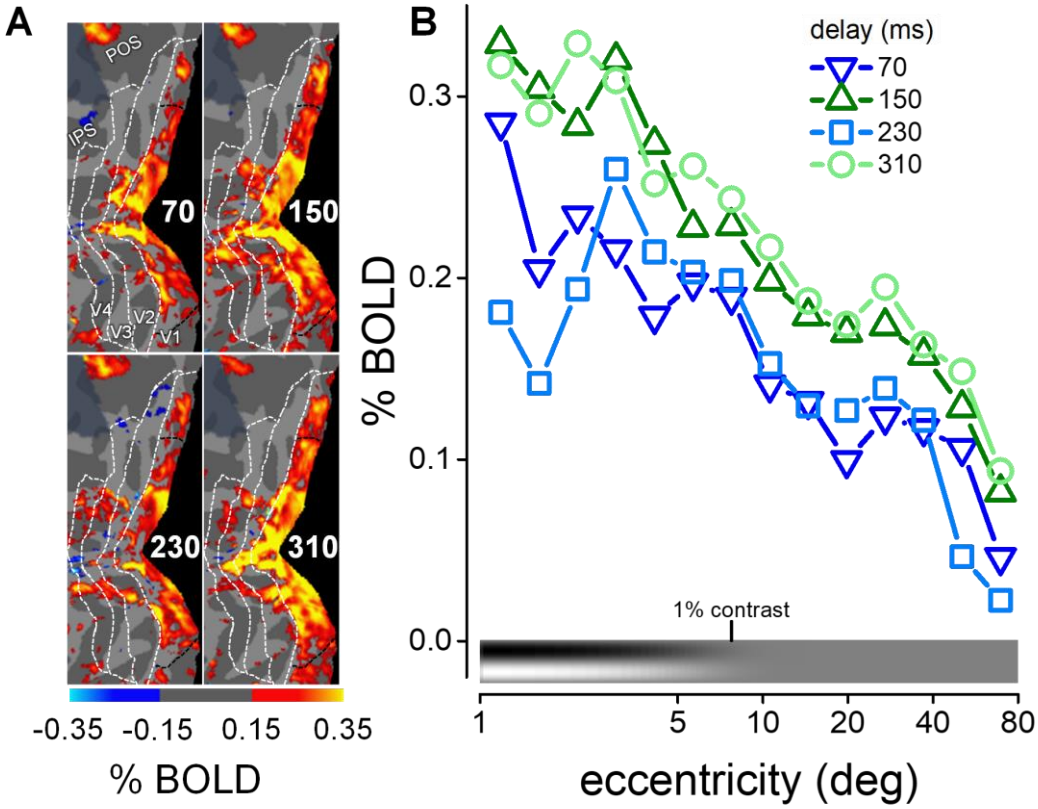

Figure 3. Effect of eccentricity on BOLD oscillations. A: BOLD response projected on the mirror-symmetric template of the posterior cortex averaged across hemispheres and participants. White dashed lines indicate the borders between V1, V2, V3 and V4; dark dashed lines delimit the part of V1 stimulated by the gratings. The left panels show activity for the 70 and 230 ms visuo-motor delays, and the right panels for the 150 and 310 ms delays. Figure S1A shows the complete brain surface map in response to all stimulus delays. **B**: V1 BOLD response as function of eccentricity in 14 non-overlapping logarithmic steps from 1 to 80 deg. The icon by the x-axis represents the variation of stimulus contrast over eccentricity.

BOLD correlations in the sensory-motor network

Having established that BOLD responses in V1 modulate rhythmically at around 5 Hz, we checked whether a similar oscillation also occurs in motor or somatosensory areas. For each participant, we defined ROIs in motor (M1) and somatosensory (S1) primary cortex using the BOLD response from independent trials where participants also performed the perceptual tasks. The M1 and S1 ROIs were located close to the Rolandic sulcus with distinct foci of activity (see example maps and ROI in Supplementary Figure 1S). Figure S2 shows the haemodynamics of M1 and S1 in response to *vision and action* events, which are different from those in V1 (compare panel A to panels C-E in Figure S2): the response occurs earlier (reflecting activity during action preparation[36]) and displays a more pronounced negative lobe. Both M1 and S1 responded strongly to the *vision and action* event, but the response was not modulated by the visuo-motor delay (supplementary Figure S2, panel D-F). This is expected given that the paradigm requires a visual response to act as a probe of the underlying cortical excitability.

To investigate whether synchronous oscillatory activity in V1 is shared with the motor area, we applied a correlation analysis. We factored out the V1 and M1 evoked responses and measured the shared trial-to-trial variability of BOLD signals between areas V1 and M1. Specifically, we extracted the residuals from the best fit GLM model and evaluated the cross-correlation at various lags between cortical areas, to estimate the amount of common noise. In essence, this is a measure of functional connectivity as a function of time lag. Figure 4A shows the correlation matrix between residuals in V1 and M1 for individual time lags from 3 s before action-onset to 12 s after, collapsing data for all visuo-motor delays. The matrix is clearly asymmetric around the diagonal. The correlation is strongest for time-lags when M1 leads (top-left half of the matrix) than for time-lags where M1 trails (bottom-right half), implying that M1 signals are predictive of V1, but not vice versa.

Figure 4B shows the average correlation coefficients along the diagonal (black outlined checks; 0 s lag), computed separately for each visuomotor delay; Figure 4C shows the same for the two upper off-diagonals where M1 leads V1 by 3 or 6 s. In both cases, the correlation shows an oscillatory pattern, with stronger correlations at delays of 150 and 310 ms. Importantly there was no modulation with delay for correlations where M1 trails (Figure 4D). These results show that stochastic variations in M1 BOLD signals predict V1 BOLD modulations with an anticipation of 3 to 6 s. They further show that the predictive coupling of M1 to V1, as well as simultaneous connectivity, is enhanced precisely when stimulus discriminability and V1 cortical excitability are highest, suggesting a multiplicative effect of the visual response. Note that the increased correlations are not a by-product of the enhanced V1 BOLD

responses, as these were factored out from the analysis. Furthermore, they are specific for M1 leading, confirming that they are not stimulus-driven and further implicating the directionality of the connectivity, from M1 to V1.

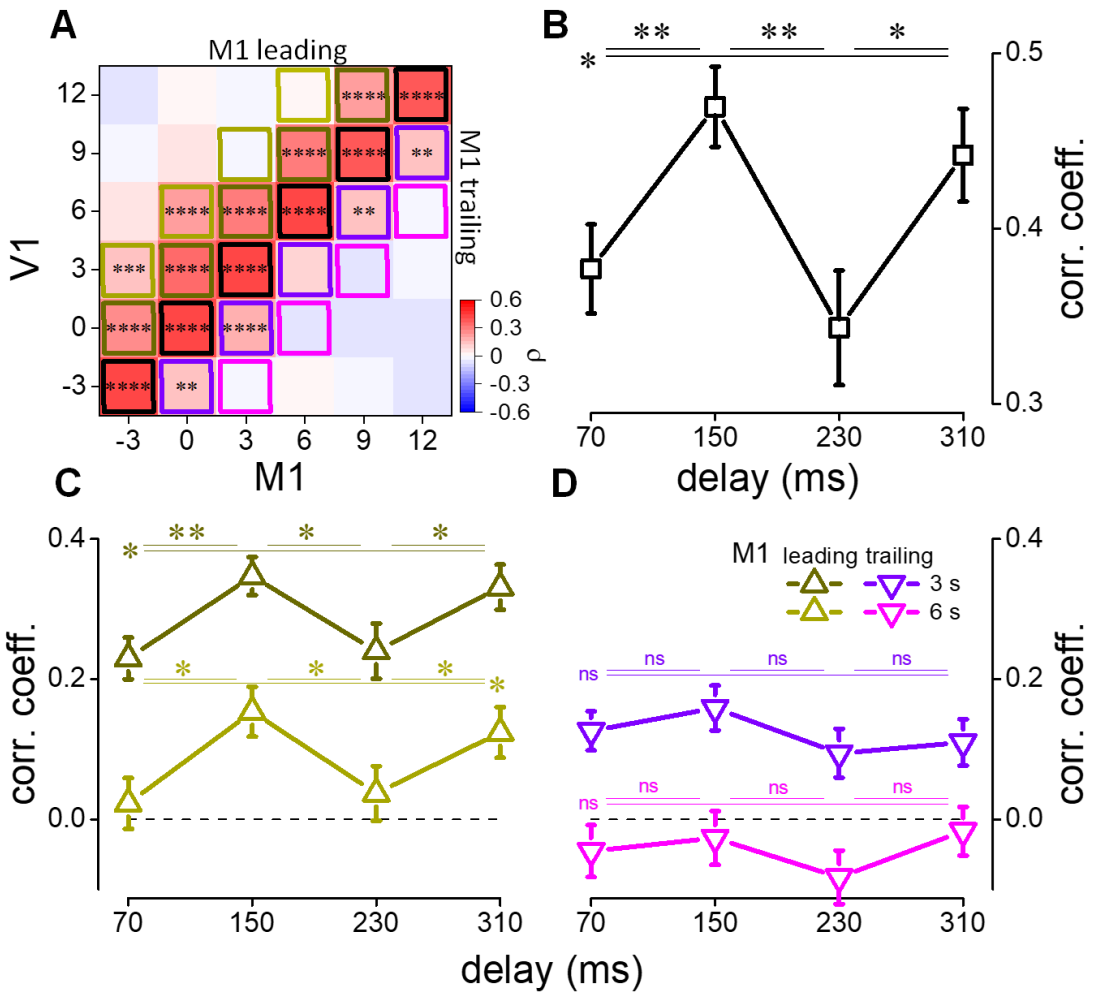

Figure 4. V1-M1 correlations. A: Correlation matrix of V1 and M1 residuals, computed between at times from −3 to +12 s from action execution, pooling across *vision and action* events with different visuo-motor delays; the map shows correlation coefficients (Pearson's r shown in the colour bar). Asterisks indicate statistical significance 0.05 > * > 0.01 > ** > 0.001 > *** > 0.0001 > ****. **B-D**: Correlation coefficient of V1-M1 residuals plotted as a function of visuo-motor delay and averaged across different lags of the matrix highlighted by colour-matching boxes in panel A; in B for the diagonal highlighted with black contour; in C for the off-diagonals of the matrix highlighter in yellow contour corresponding to M1 leading V1 by 3 s or 6 s; in D for the off-diagonals of the matrix highlighted in magenta contour corresponding to lags of M1 trailing V1 by 3 s or 6 s. Symbols and error bars are mean and s.e.m from the bootstrapped aggregate observer. Asterisks indicate statistical significance (Bonferroni-Holm corrected): ns > 0.05 > * > 0.01 > ** > 0.001 > ***.

Model of the BOLD response component of the visuo-motor system

The overall pattern of results strongly suggests that visual and motor cortices become synchronized during a simple visuo-motor task (keypress followed by a visual stimulus). The specific findings can be summarized with five major points:

1) A voluntary action modulates the BOLD responses in V1 (Figure 2A).

2) The effect on BOLD response is not only multiplicative but includes an additive negative component (Figure 2A).

3) The BOLD modulation oscillates rhythmically (Figure 2B) with similar timecourse to the oscillation in accuracy (Figure 1B), suggesting that both are driven by synchronization of endogenous oscillations.

4) The modulatory signal affects the entire V1 region, extending beyond the stimulated region (Figure 3). There is also a small but significant BOLD response in V1 following a voluntary action that is not followed by a visual stimulus (Figure 5, red square symbols).

5) The modulatory signal is already present before stimulus presentation, given that the predictive correlation of M1 and V1 BOLD starts at least 6 s before action onset.

To capture all these effects, we developed a toy model simulating some well-known physiological properties of the visuo-motor loop, including corollary signals during motor preparation, oscillatory rhythms and gain modulation of visual responses (see Figure 5). The key hypotheses are: 1) gain is rhythmically modulated in V1 at a theta frequency synchronously with motor preparation, hence with action onset; 2) V1 excitability is modulated in synchrony with motor signals, both before and after action.

The V1 response to *vision only* stimuli is well fit by the standard haemodynamic response function ($R^2 = 0.39$, black trace), assuming that the visual input is a delta function with amplitude chosen to maximize goodness of fit (black trace in the inset). To model the timecourse of the BOLD response to *vision and action* events, we assumed that the visual input is combined with a motor preparation signal (or with an equivalent synchronous signal).

We modelled this motor preparation signal with two well established components of the visuomotor loop. The first is the "readiness potential" associated with the keypress, which we represented as a ramp signal starting 1.5 s before the action onset[37,38]. The readiness potential precedes all voluntary

action independently of whether the action is followed by a visual stimulus, explaining the V1 response to keypress without visual stimuli (red curve in Figure 5). Using a standard haemodynamic function, we fitted the V1 BOLD response following a voluntary action (with no visual stimuli) to determine the ramp amplitude that best fitted the data (R^2 = 0.38). The second component is a negative signal, which we assume to be related to the synchronization of endogenous oscillations with the action onset. This is based on previous work showing that negative BOLD is associated with synchronized endogenous rhythms[39,40]. We simulate this signal as a boxcar starting at action-onset and lasting for several seconds (inset in Figure 5, blue-green trace), as it has been shown that behavioural oscillations persist over a large temporal window after the action[24,27]. The response to this negative signal, summed to the responses to the motor preparation and visual stimulus, predicted the time course of the response to *action and visual* condition for the 150 ms stimulus presentation (Figure 5, green curve).

Finally, we postulated a multiplicative gain modulation of visual responses as function of visuo-motor delay as the result of cortical excitability that follows the theta rhythm after action onset. This allowed us to model the response at 70 ms delay from that obtained at 150 ms delay, simply by scaling the amplitude of the delta function that represents the visual input (Figure 5, blue trace). The gain variation required to fit the data is small, about 33%, which would represent the excitation modulation produced by the synchronization of endogenous oscillations (R^2 = 0.38 and 0.40 for visuo-motor delays of 70 and 150 ms, respectively). We fitted the BOLD response for visuo-motor delays of 230 and 310 ms with the same model (not shown in the figure for simplicity). The fit for all delay conditions were similarly good (R^2 = 0.52 and 0.41 for 230 and 310 ms delays, respectively).

Using very few parameters, the model captured quantitatively the five different aspects of V1 modulations listed above.

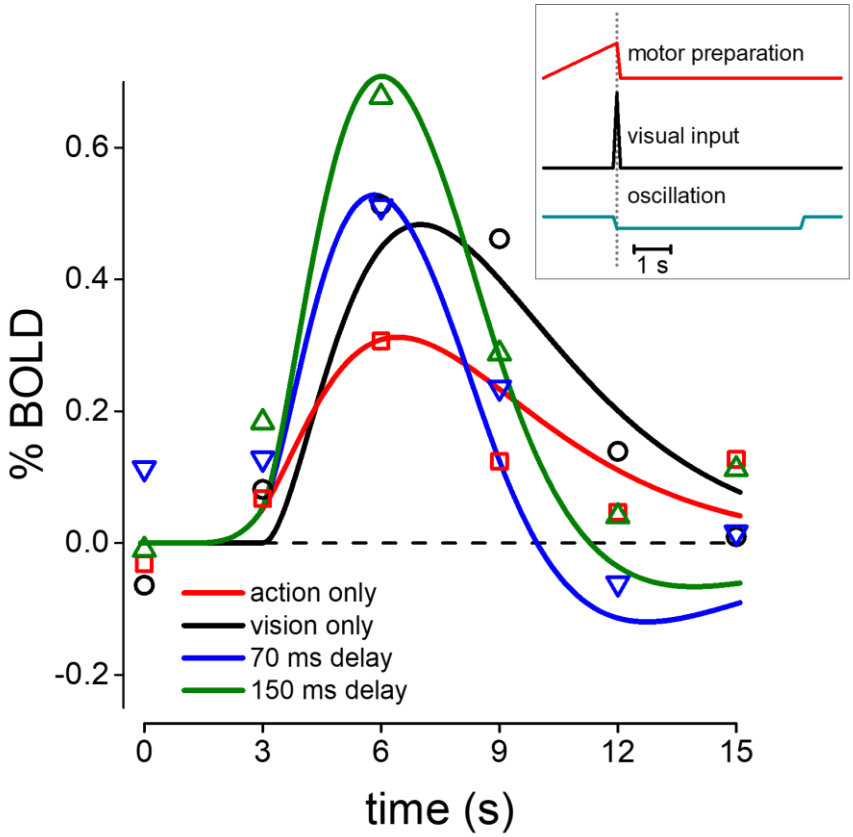

Figure 5. Simulation of V1 BOLD response. Simulated (continuous lines) and observed (symbols) V1 BOLD responses to visual events presented without keypress (*vision only*, black curve), to *vision and action* events at delays of 70 ms (blue curve) or 150 ms (green curve), and to keypress-only events (black). Simulated curves were obtained by convolving the appropriate inputs in the **inset** with a standard haemodynamic function[41] with decay time $\tau = 2$ s, order = 3 and delay $\delta = 3$ s.

Discussion

We applied a time-resolved perturbative-probe fMRI paradigm to study the neural mechanisms underlying the M1-V1 sensory-motor loop. We measured V1 BOLD responses to visual probes briefly presented at varying delays after a voluntary keypress action. BOLD response amplitude fluctuated rhythmically after keypress (Figure 2B), following closely the theta rhythm of the behavioural oscillations in sensitivity (Figure 1B). Similar oscillations modulated the functional coupling between M1 and V1, at time-lags when M1 leads V1. The state of the motor cortex before action-onset predicted the activity of V1 over several seconds, demonstrating predictive sensory-motor communication across distal cortical areas over a temporal interval of several seconds. These results indicate that perception is shaped by motor planning as reflected in M1 activity before action onset. They suggest that a predetermined common synchrony orchestrates the premotor signals, inducing the subsequent rhythmic modulation of V1, and ultimately the rhythmic modulation of visual discrimination. All these effects can be modelled by assuming that V1 and M1 activity oscillate synchronously at the time of a voluntary action, and that this oscillation provides a precise and long-lasting representation of the time-of-action within V1, which can be used to shape visual processing from its earliest stages.

Our results are consistent with prior evidence that preparation of voluntary actions, such as finger-tapping, button-press, reaching, grasping and even isometric muscle contraction can modulate the perception of visual stimuli (for review [42]). From hundreds of milliseconds before action onset (during motor preparation) to up to 1 second after action offset, visual sensitivity oscillates within the theta range, phase-reset by the action[22,24,27]. We replicated this behavioural oscillation with a discrimination task in an fMRI scanner, and observed V1 BOLD responses that oscillate rhythmically around 5 Hz, strongly implicating V1 as the origin of the perceptual oscillation. This result is also consistent with recordings in monkeys showing rhythmic theta-range modulation in V1[11,14,43], and with neurophysiological data from humans during an active task[25]. Interestingly, the resolution of the delay-dependent BOLD modulation was finer than that required for humans to judge simultaneity of action-onset and a visual stimulus (around 150 ms[44]), excluding cognitive effects.

Our results provide strong evidence that voluntary action modulates V1 activity: not only was the visual response after a voluntary action different from that recorded during passive viewing, but a voluntary action on its own, unaccompanied by any visual stimulus, also elicited a significant V1 BOLD response. The observed pattern of V1 responses is consistent with the models of active sensing or embodied perception, which propose that processing of sensory inputs is profoundly altered when input

is actively sought through eye-, hand-, or body-movements, compared with passively delivered inputs[45,46]. Another well-described phenomenon associated with active sensing is motor-induced suppression of sensitivity, typically occurring in the first 100 ms after action onset. The phenomenon is very obvious after saccadic eye movements[47], but also occurs after hand movements[48,49], and can cause a measurable change in V1 BOLD response[50]. The attenuation of the response to stimulus presentation that we observe at 70 ms delay might be interpreted as the consequence of this motor-inhibition. However, if this modulation of the haemodynamic response were simply transient motor-suppression[35], we would expect a monotonic increase with larger delays, rather than the rhythmic modulation we observe[24,26]. Rather than a simple transient suppression, our results point to a cyclic alternation starting before action-onset. On this view, the function of the modulation is not to eliminate unstable visual signals when sensors move, but rather is part of a cyclic mechanism of temporal binding. The observed oscillation of BOLD responses and functional connectivity can be explained by synchronization of activity in motor and visual cortex by a common oscillatory rhythm present during preparation phase. The BOLD connectivity between M1 and V1 oscillating with the stimulus delay strongly suggests that the noise in the synchronization of the motor cortex is reflected in the excitability of the V1. Importantly, the same effects are also seen in correlations between motor activity lagging visual activity, suggesting that the noise in synchronisation propagates over time.

We modelled all the complex features of our results, including the perceptual modulation, by assuming that synchronous oscillations cause a general subtractive effect on BOLD signal in V1[39,40], and that cortical excitability undergoes a multiplicative oscillatory gain change, which follows the phase of the theta rhythm[9,11-14]. Both our key assumptions are consistent with the BOLD literature (for review [51]). The model can be easily extended to predict the rhythmic modulation of functional connectivity with stimulus delay, and the predictive power of later V1 activity, by introducing noise in the ramp and oscillation signal. Overall, the model framework exploits a mechanism of "communication through coherence"[6] for transferring information between the sensory and the motor systems. This protocol is assumed to operate between distal cortical areas by synchronization of neural activity in local neuronal assemblies[6]. Synchronized activity enhances effective connectivity within a local assembly and selectively improves the communication between assemblies[6,8]. Although synchronization typically occurs in the gamma-range (30-90 Hz)[8], slower oscillations in the alpha and theta ranges can modulate the gamma rhythm and thereby affect communication[9,11-14].

Other examples of communication through coherence exist in sensory-motor systems. Corticospinal gamma-band coherence occurs during movement preparation, and the coherence correlates with reaction times[52]. Functional connectivity between cortical areas changes, depending on the local phase of the endogenous theta oscillation in humans[53]. Theta range synchronous modulation has been interpreted as a multiplexing mechanism to allocate processing resources to sensory and motor functions[54,55]. As with most multiplexing communication systems, this oscillatory communication may also be efficient and easy to time-lock. Our results corroborate these findings, and together suggest an efficient, predictive, and flexible communication system through synchronization of endogenous rhythms for sensorimotor coordination. Flexibility is a critical issue, given that sensorimotor synchronization needs to rapidly recalibrate to accommodate internal and external changes (for example, fatigue can affect motor delays, and illumination visual delays). This complex task may become more manageable if time is encoded in a cyclical function, represented within primary sensory areas in the form of periodic modulations of cortical excitability. Our observation that M1 signals in the motor preparation stage predict the oscillation of V1 response (Figure 4) strongly supports this notion.

Several studies propose that phase of theta oscillations drives selective and spatial attention[17,19,20,56-59], which is well known to modulate low-level areas such as V1 (for a review [60]). Attention may in principle contribute to our results, but this seems unlikely for several reasons. Firstly, the oscillation of V1 BOLD responses was not limited to the central visual field encompassing the stimulus representation, but extended to the far periphery of V1, which was neither stimulated nor task-relevant. Secondly, V1 representations of upper and lower halves of the stimulus oscillated in-phase, consistent with evidence that oscillations synchronized by voluntary actions have the same phase across multiple locations[22], whereas attentional fluctuations typically have opposite phases across locations[17,19]. Thirdly, the V1 modulation was not stronger in the lower visual field, despite the lower-hemifield bias of attentional processes[61].

Unlike V1, responses in M1 and S1 did not modulate rhythmically with presentation time of the visual stimulus. This is to be expected, given that M1 and S1 do not respond to visual stimuli, so a visual stimulus will not be an adequate probe of cortical excitability. Interestingly, using a similar time-resolved fMRI approach, and measuring reaction-times, Guo et al.[62] observed rhythmic modulations of M1, consistently with previous report[63]. We interpret their oscillation in M1 to be related to the programming of speeded responses, whose latency oscillates with a theta rhythm[20,56,62], like behavioural sensitivity. Alternatively, the oscillation may relate to the memory trace of the stimulus. Theta rhythms are

characteristic of the reactivation of memory traces[64,65], and of the binding of sensory signals for memory recall[66]. Both interpretations are consistent with the lack of M1 oscillation observed here, given that participants were not required to make motor actions to report perceptual decisions.

Active sensing and the impact of action on sensory processing have been the subject of research since the intuition of the corollary discharge by von Helmholtz[1], later taken up by Sperry[2] and von Holst and Mittelstaedt[3]. However, the details of the underlying neuronal mechanisms remain largely unknown. For visual perception during eye movements, it is well established that pre-motor signals can dramatically change visual processing and perception (for review [4,5]). These complex effects may generalize to many visuo-motor tasks. The execution of an action is consistently recorded in the activity of human primary visual cortex[50,67-69], and similarly strong modulations are apparent during locomotion in the primary cortex of mice[70]. Here we find that visuo-motor interactions have high temporal specificity, since perceptually indistinguishable changes in the delay between motor and visual events are sufficient to produce dramatic changes in the pattern of V1 activity, as well as of M1-V1 functional connectivity. Moreover, premotor signals (readiness potential) actively participate in the temporal recalibration of cortical activity necessary after adaptation to altered visuo-motor temporal delays for efficient processing[71]. All these examples may be interpreted as evidence for predictive coding of the future consequence of the action[32,34,48,72,73] elaborated during motor preparation and represented within sensory cortex. These predictions may be instrumental in stabilizing perception across movements, as well as in endowering a sense of agency. Our finding that M1 activity during action planning predicts V1 signals demonstrates the existence of a predictive signal, exchanged between these distal areas through the synchronization of oscillatory activity patterns. Thus, communication through coherence could provide an efficient protocol for predictive coding, signalling the onset of a voluntary action across the brain, and thereby shaping sensory processing.

Acknowledgments

This project was supported by the European Research Council (ERC) under the European Union's Horizon 2020 program (Grant Agreement No 832813-GENPERCEPT and No 801715-PUPILTRAITS) and by MIUR - PRIN 2017 - grant 2017SBCPZY_02. We thank Eleanor Reynolds and David Burr for proofreading the manuscript. The authors would like to thank Dr. Brian Burns (GE Healthcare) for his implementation of the 3D-MP2RAGE sequence on the MR950 system.

Author contributions

Conceptualization, A.B, P.B., and M.C.M.; Methodology, all authors. and M.C.M.; Investigation, A.B, M.C. and P.B.; Writing, all authors.

Declaration of interests

The authors declare no competing interests.

Keywords

Endogenous oscillations; visual discrimination; vision and action; 7T BOLD; visual cortex; sensorimotor cortex; communication through coherence; phase-reset.

STAR Methods

RESOURCE AVAILABILITY

Lead Contact and Materials Availability

Further information and requests for resources should be directed to the Lead Contact, Maria Concetta Morrone (concetta.morrone@unipi.it).

Data and Code Availability

Analyses were performed with Freesurfer v6.0.0, SPM12, BrainVoyager 20.6, and FSL (version 5.0.10) packages And MATLAB (version R2020b). Experimental data will be made available before publication on a repository in Zenodo.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human subjects

Experimental procedures are in line with the declaration of Helsinki and were approved by the regional ethics committee (*Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer*—Firenze (FI)) and by the Italian Ministry of Health, under the protocol '*Plasticità e multimodalità delle prime aree visive: studio in risonanza magnetica a campo ultra-alto (7T)*' version #1 dated 11/11/2015. Written informed consent was obtained from each participant, which included consent to process and preserve the data and publish them in anonymous form.

We recruited seven participants for the behavioural experiment (26-36 years old, mean age = 30, 3 females and 4 males), with normal or corrected to normal vision. Twenty-four healthy participants with normal or corrected-to-normal vision were recruited for the fMRI experiment. Three of them did not complete the scanning session and were excluded from the analyses. Of the remaining 21 participants (24-63 years old, mean age = 30, 9 females and 12 males), 17 took part in the main experiment that included 60±17 visual presentations during 5±2 fMRI acquisitions (mean ± standard deviation.

METHOD DETAILS

Behavioural experiment: setup and procedure

Behavioural measurements were made in a quiet room in dim light condition. The stimuli were generated in MATLAB (MATLAB r2010a; MathWorks), displayed with a graphics board ViSaGe (Cambridge Research System) on a linearized CRT monitor (Barco Calibrator 40×30 cm, resolution of 800×600 pxl, refresh rate 100 Hz, mean luminance 25 cd/m^2) at 57 cm from the observer. Keypress responses were acquired by CB6 response box (Cambridge Research System).

The visual stimulus comprised two horizontal gratings (10% contrast) presented for 30 ms in the upper and lower visual field, both vignetted by a Gaussian window (σ = 5 deg) at screen centre (see inset Figure 1). The two gratings were separated by a Gaussian horizontal blank with space constant of 0.5 deg (σ). The spatial frequencies of the upper and lower gratings varied randomly between 1 cpd and 1.1 cpd from trial to trial. The phase of each grating also varied randomly across trials. Participants maintained fixation on a small red dot (0.15 deg) at the centre of the display and pressed a key to start the trial (with the fourth finger of the right hand). The visual stimulus was displayed at 60 possible stimulus delays from button press, chosen randomly on each trial from a uniform distribution between 0 and 600 ms in steps of 10 ms. The participant's task was to indicate which grating (upper or lower) had the higher spatial frequency (i.e. a 2AFC spatial frequency discrimination task) by pressing a button with the index or the middle finger of the right hand.

Participants were instructed to wait at least 1 s between two successive keypresses (this interval was 1.6 s on average, with 1 s standard deviation). Trials with an ITI of less than 1 s (1.3% of trials) were excluded from further analysis. We collected 1008±187 trials for each participant (mean and standard deviation).

Behavioural oscillation analysis

To test whether visual accuracy fluctuated rhythmically over time, we pooled all trials from all participants, sorted them in terms of visuo-motor delays and grouped them into 20 ms nonoverlapping bins. For each bin, we computed the proportion of correct responses and fitted the resulting time series with the following sine function:

$$y(x) = \beta_0 + \beta_1 \sin(2\pi f x) + \beta_2 \cos(2\pi f x) \qquad \text{[eq. 1]}$$

where f represents the frequency of the oscillation and can vary between 2 and 20 Hz in steps of 0.1 Hz; β_0, β_1 and β_2 are free parameters to be estimated. Statistical significance was evaluated by

comparing the goodness of fit (R^2) for the original dataset, with the distribution of goodness of fits obtained on permuted datasets created by shuffling the original responses (10000 permutations) with frequency f as a free parameter. This automatically corrects for multiple frequencies comparisons.

fMRI experiment: setup and procedures

Visual stimuli were presented by using a magnetic resonance-compatible goggle set (VisuaStimDigital, Resonance Technologies Inc., Northridge, California) with visual field of approximately 32×24 deg, a resolution of 800×600 pxl, refresh rate 60 Hz, mean luminance 25 cd/m^2. Keypresses were recorded through a magnetic resonance-compatible response-box (Evoke Response Pad System, Resonance Technologies Inc., Northridge, California), positioned on the right hand of volunteers.

The same visual stimulus and procedure of the behavioural experiment were used in the fMRI scans, with the following differences. To accommodate for the different refresh rate of the monitor, the duration of the stimulus was 33 ms. Participants were instructed to wait at least 15 s between two successive keypresses. Only in a subset of trials (not used for the main analysis) participants were required to report perceptual judgments. As for the behavioural experiment, to start the trial, participant pressed a button on the response-box below the fourth finger.

In a first set of runs (not included in the main analyses), all participants (N=21) performed the discrimination task, reporting their decision with a second keypress, delayed by several seconds (15 s on average, with 4 s standard deviation) to minimize interference with the visual presentation. These trials were used to assess that stimulus discriminability in the scanner was similar as in the psychophysical experiment (62±3% on average; the paucity of trials did not allow for tracking oscillations in accuracy as function of visuo-motor asynchrony). We used the BOLD modulation associated with the decision keypress to locate, for each participant, cortical areas holding a representation of the keypress action both in M1 and S1 (see supplementary Figure S1). Primary visual cortex V1 was localized by separate retinotopic mapping data (two 45 deg wedges centred around the horizontal or vertical meridian, presented alternately for 5 TRs each, with no blanks, and six repetitions in total).

Data for the main experiment (Figures 1-5) were collected in 17 of the 21 subjects. Participants were not required to report their visual discrimination judgments. Visual stimuli were presented every 19 s on average (3 s standard deviation). We tested four possible asynchronies between the keypress and

the visual stimulus (delay = 70, 150, 230, or 310 ms), selected to match the maxima and minima of performance as assessed in the behavioural experiment. The average number of trials (± 1 standard deviation) per participant were: 17±6 for delay 70 ms; 17±6 for delay 150 ms; 13±5 for delay 230 ms; 13±4 for delays 310 ms. As a control, in a subset of 5 participants we additionally tested delays of 180 ms vs 150 ms (51±12 trials each). Given that participants started the trial at their own pace, neither keypress nor visual presentation could be time-locked to the TR. In order to define the design matrix used for the GLM analyses, we assigned each *vision and action* event to the TR preceding action onset. This meant that in about 3% of trials, the visual presentation actually occurred 1TR later than indicated in the design matrix (this happened in about 5% of trials for the 310 ms delay, and 4%, 3%, and 2% for the 230, 150 and 70 ms delays respectively). We verified that aligning the design matrix to the visual stimulus rather than the action onset did not change the results (Figure S2, panel A). We also verified that the temporal distribution of events was homogeneous in the TR, using circular statistics (Rayleigh test for 70 ms delays: $p = 0.34$; 150 ms delays: $p = 0.55$; 230 ms: $p = 0.89$; 310 ms delays: $p = 0.45$). The results ensured that there were no temporal biases between conditions and therefore that event timing could not explain response amplitude differences.

We also ran a feasibility experiment testing a subset of N = 5 participants with *vision-only*, where the visual stimulus was delivered alone, with no accompanying keypress.

MRI scanning

Scanning was performed on a MR950 7T whole body MRI system (GE Healthcare, Milwaukee, WI, USA) equipped with a 2-channel transmit coil driven in quadrature mode, a 32-channel receive coil (Nova Medical, Wilmington, MA, USA) and a high-performance gradient system (50 mT/m maximum amplitude and 200 mT/m/ms slew rate).

For twelve participants, the anatomical images were acquired at 1 mm isotropic resolution using a T1-weighted magnetization-prepared Fast Spoiled Gradient Echo (FSPGR) sequence with the following parameters: TR = 6 ms, TE = 2.2 ms, flip angle = 12 deg, receiver Bandwidth = 50 kHz, TI = 450 ms, ASSET = 2. For the remaining 9 participants, high-resolution anatomical images with 0.8 mm isotropic resolution were acquired with a modified magnetization-prepared rapid gradient echo sequence (MP2RAGE) with the following parameters: TR = 6.6 ms, TE = 2.5 ms, flip angle = 5 deg, receiver Bandwidth = 62.5 kHz, TI = 1000 ms and 3200 ms, ARC factor = 2.

Functional images were acquired with a gradient-echo EPI sequence with Field of View = 192 mm, matrix size = 128x128, resulting in a spatial resolution of 1.5x1.5 mm^2 in-plane, and slice thickness 1.4 mm with slice spacing = 0.1mm, TR = 3 s, TE = 23 ms, flip angle = 60 deg, receiver Bandwidth = 250 kHz, ASSET acceleration factor = 2, phase encoding direction: Anterior-to-Posterior. No resampling was performed during the reconstruction. For each EPI sequence, we acquired an additional volume with the reversed phase encoding direction (Posterior-to-Anterior), used for distortion correction (see pre-processing).

Pre-processing of imaging data

Anatomical images were processed by a standard procedure for segmentation implemented in Freesurfer (recon-all). In addition, hemispheres were aligned to a template of the cortical surface (fsaverage) as well as to the left/right symmetric version of the same template (fsaverage_sym).

Functional images were corrected for subject movements, slice-time acquisition. Geometrical distortions were compensated by using EPI images with reversed phase encoding direction (Brain Voyager COPE plug-in). The pre-processed images were aligned to each participant's anatomical image using a boundary-based registration algorithm (Freesurfer "*bbergister"* algorithm) and projected to the cortical surface of each hemisphere. All analyses were conducted on data in the individual subject space, resampled to a resolution of 3x3x3 mm. For each run, we removed the first 6 TRs (to allow the MR signal to reach steady-state), then estimated the temporal linear trend and subtracted it from the fMRI time-course. Time-courses were averaged within each ROI (see below) and converted into percentage signal change by normalizing them to the mean BOLD.

QUANTIFICATION AND STATISTICAL ANALYSIS

Definition of Regions of Interest (ROI)

For each participant, we defined three ROIs (bilateral V1, left M1 and left S1) in the 3D native space of each participant. V1 was defined using the responses to the vertical and horizontal meridians from the retinotopic mapping scans and included the cortical representation of the stimulus area, in both hemispheres of about 7.3 cm^3. The M1 and S1 ROIs were defined based on the pattern of BOLD responses to keypress associated with the perceptual decision that generated two separate foci located

rostrally and caudally of the left Rolandic sulcus, consistent with M1 and Brodmann Area 3a. The M1 and S1 ROIs comprised on average 5.8 cm^3 of cortex.

Evaluation of fMRI Activity

We used the Finite Impulse Response deconvolution approach for estimating the Haemodynamic Response Function (HRF) over a window of 7 TRs following each event (an "event" being a visual stimulus presentation, a voluntary keypress or both). This analysis approach does not assume a particular shape of the haemodynamic response but, rather, estimates it directly by modelling each of the 7 TRs after the events as a separate GLM predictor and thereby assigning each with a beta-weight[74]. Given that visual stimuli and the keypress that started the trial were separated by less than 350 ms, we aligned the predictor with the TR that included the keypress (Figure S2A shows V1 BOLD responses computed after aligning to the visual stimulus instead: compare Figure S2A with Figure 2 in the main text). In each ROI, seven beta values per each event type were estimated with GLM (e.g. for the main analyses, we considered four types of events corresponding to the 4 visuo-motor delays, making 28 predictors to which we added a constant term); the average goodness of fit was 20±2% (average number of timepoints: 629).

Statistical comparisons were performed after extracting GLM beta-weights in individual participants and in each of our *a priori* defined ROIs.

We assessed pairwise correlations between M1 and V1 in the aggregate observer data, after regressing out the event-related modulations, i.e. using the residuals of the GLM model. Specifically, we performed 4 steps: 1) compute the predicted time-course given the estimated GLM weights for each run; 2) subtract it from the observed time-course to obtain the residuals; 3) concatenate residuals from all runs and all participants; 4) compute the Pearson's correlation between residuals of the two ROIs for the corresponding time lags. Statistical significance of difference of correlation across delays was assessed by bootstrap (10000 repetitions with replacement, one-tailed test), which means that step #4 was repeated 10000 times, every time selecting a different set of events.

We also generated surface maps of the GLM beta weights at the TRs of interest (the peak of haemodynamic response at TR2 for Figure S1, or the average beta weights within TR1 and TR4 for Figure 3) by projecting the pre-processed BOLD timecourses from each participant on a common mirror-symmetric template of the cortical surface, concatenating them and then applying the GLM analysis. For Figure S1, data from the two hemispheres were analyzed separately; for Figure 3, the concatenated

BOLD timecourses were averaged across hemispheres before the GLM analysis. Activity maps were masked for statistical significance ($p_{FDR} < 0.05$). Eccentricities in primary visual cortex were defined according to the Benson template for fsaverage[75].

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and algorithms		
MATLAB r2010a, r2020b	The MathWork	http://www.mathworks.com/products/matlab/; RRID: SCR_001622
Freesurfer v6.0.0	[76]	http://surfer.nmr.mgh.harvard.edu/; RRID: SCR_001847
BrainVoyager 20.6	[77]	http://www.brainvoyager.com/products/brainvoyagerqx.html; RRID: SCR_013057
FSL v5.0.10	[78]	http://www.fmrib.ox.ac.uk/fsl/; RRID: SCR_002823
SPM12	[79]	http://www.fil.ion.ucl.ac.uk/spm/; RRID: SCR_007037
Psychophysics Toolbox v3	[80]	http://psychtoolbox.org/; RRID: SCR_002881

References

1. Helmholtz, H.v. (1866). Handbuch der physiologischen Optik, (Leipzig Leopold Voss).
2. Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. J Comp Physiol Psychol *43*, 482-489.
3. von Holst, E., and Mittelstaedt, H. (1950). Das Reafferenzprinzip. Naturwissenschaften *37*, 464-476.
4. Binda, P., and Morrone, M.C. (2018). Vision During Saccadic Eye Movements. Annu Rev Vis Sci *4*, 193-213.
5. Wurtz, R.H. (2018). Corollary Discharge Contributions to Perceptual Continuity Across Saccades. Annu Rev Vis Sci *4*, 215-237.
6. Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. Neuron *88*, 220-235.
7. Buzsáki, G. (2006). Rhythms of the brain, (Oxford ; New York: Oxford University Press).
8. Womelsdorf, T., Schoffelen, J.M., Oostenveld, R., Singer, W., Desimone, R., Engel, A.K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. Science *316*, 1609-1612.
9. Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger, M.S., Barbaro, N.M., and Knight, R.T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. Science *313*, 1626-1628.
10. Haegens, S., Nacher, V., Luna, R., Romo, R., and Jensen, O. (2011). alpha-Oscillations in the monkey sensorimotor network influence discrimination performance by rhythmical inhibition of neuronal spiking. Proceedings of the National Academy of Sciences *108*, 19377-19382.
11. Bosman, C.A., Schoffelen, J.M., Brunet, N., Oostenveld, R., Bastos, A.M., Womelsdorf, T., Rubehn, B., Stieglitz, T., De Weerd, P., and Fries, P. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. Neuron *75*, 875-888.
12. Landau, A.N., Schreyer, H.M., van Pelt, S., and Fries, P. (2015). Distributed Attention Is Implemented through Theta-Rhythmic Gamma Modulation. Curr Biol *25*, 2332-2337.
13. Helfrich, R.F., Fiebelkorn, I.C., Szczepanski, S.M., Lin, J.J., Parvizi, J., Knight, R.T., and Kastner, S. (2018). Neural Mechanisms of Sustained Attention Are Rhythmic. Neuron *99*, 854-865 e855.
14. Spyropoulos, G., Bosman, C.A., and Fries, P. (2018). A theta rhythm in macaque visual cortex and its attentional modulation. Proc Natl Acad Sci U S A *115*, E5614-E5623.
15. Busch, N.A., Dubois, J., and VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts Visual Perception. Journal of Neuroscience *29*, 7869-7876.
16. VanRullen, R. (2016). Perceptual Cycles. Trends Cogn Sci *20*, 723-735.
17. Landau, A.N., and Fries, P. (2012). Attention samples stimuli rhythmically. Curr Biol *22*, 1000-1004.
18. Romei, V., Gross, J., and Thut, G. (2012). Sounds reset rhythms of visual cortex and corresponding human visual perception. Curr Biol *22*, 807-813.
19. Fiebelkorn, I.C., Saalmann, Y.B., and Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. Curr Biol *23*, 2553-2558.
20. Song, K., Meng, M., Chen, L., Zhou, K., and Luo, H. (2014). Behavioral oscillations in attention: rhythmic alpha pulses mediated through theta band. J Neurosci *34*, 4837-4844.

21. Ho, H.T., Leung, J., Burr, D.C., Alais, D., and Morrone, M.C. (2017). Auditory Sensitivity and Decision Criteria Oscillate at Different Frequencies Separately for the Two Ears. Curr Biol *27*, 3643-3649 e3643.
22. Tomassini, A., Spinelli, D., Jacono, M., Sandini, G., and Morrone, M.C. (2015). Rhythmic oscillations of visual contrast sensitivity synchronized with action. J Neurosci *35*, 7019-7029.
23. Wutz, A., Muschter, E., van Koningsbruggen, M.G., Weisz, N., and Melcher, D. (2016). Temporal Integration Windows in Neural Processing and Perception Aligned to Saccadic Eye Movements. Curr Biol *26*, 1659-1668.
24. Benedetto, A., Spinelli, D., and Morrone, M.C. (2016). Rhythmic modulation of visual contrast discrimination triggered by action. Proc Biol Sci *283*, 20160692.
25. Tomassini, A., Ambrogioni, L., Medendorp, W.P., and Maris, E. (2017). Theta oscillations locked to intended actions rhythmically modulate perception. Elife *6*.
26. Benedetto, A., and Morrone, M.C. (2017). Saccadic Suppression Is Embedded Within Extended Oscillatory Modulation of Sensitivity. J Neurosci *37*, 3661-3670.
27. Zhang, H., Morrone, M.C., and Alais, D. (2019). Behavioural oscillations in visual orientation discrimination reveal distinct modulation rates for both sensitivity and response bias. Sci Rep *9*, 1115.
28. Assaneo, M.F., Rimmele, J.M., Sanz Perl, Y., and Poeppel, D. (2021). Speaking rhythmically can shape hearing. Nat Hum Behav *5*, 71-82.
29. Melloni, L., Schwiedrzik, C.M., Rodriguez, E., and Singer, W. (2009). (Micro)Saccades, corollary activity and cortical oscillations. Trends Cogn Sci *13*, 239-245.
30. Schroeder, C.E., Wilson, D.A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of Active Sensing and perceptual selection. Curr Opin Neurobiol *20*, 172-176.
31. Engel, A.K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. Nat Rev Neurosci *2*, 704-716.
32. Friston, K. (2005). A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci *360*, 815-836.
33. Arnal, L.H., and Giraud, A.L. (2012). Cortical oscillations and sensory predictions. Trends Cogn Sci *16*, 390-398.
34. Friston, K.J., Bastos, A.M., Pinotsis, D., and Litvak, V. (2015). LFP and oscillations-what do they tell us? Curr Opin Neurobiol *31*, 1-6.
35. Aliu, S.O., Houde, J.F., and Nagarajan, S.S. (2009). Motor-induced suppression of the auditory cortex. J Cogn Neurosci *21*, 791-802.
36. Cunnington, R., Windischberger, C., Deecke, L., and Moser, E. (2003). The preparation and readiness for voluntary movement: a high-field event-related fMRI study of the Bereitschafts-BOLD response. Neuroimage *20*, 404-412.
37. Libet, B., Gleason, C.A., Wright, E.W., and Pearl, D.K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. Brain *106 (Pt 3)*, 623-642.
38. Shibasaki, H., and Hallett, M. (2006). What is the Bereitschaftspotential? Clin Neurophysiol *117*, 2341-2356.
39. Becker, R., Reinacher, M., Freyer, F., Villringer, A., and Ritter, P. (2011). How ongoing neuronal oscillations account for evoked fMRI variability. J Neurosci *31*, 11016-11027.
40. Scheeringa, R., Fries, P., Petersson, K.M., Oostenveld, R., Grothe, I., Norris, D.G., Hagoort, P., and Bastiaansen, M.C. (2011). Neuronal dynamics underlying high- and low-frequency EEG oscillations contribute independently to the human BOLD signal. Neuron *69*, 572-583.

41. Boynton, G.M., Engel, S.A., Glover, G.H., and Heeger, D.J. (1996). Linear Systems Analysis of Functional Magnetic Resonance Imaging in Human V1. The Journal of Neuroscience *16*, 4207-4221.
42. Benedetto, A., Morrone, M.C., and Tomassini, A. (2020). The Common Rhythm of Action and Perception. J Cogn Neurosci *32*, 187-200.
43. Kienitz, R., Cox, M.A., Dougherty, K., Saunders, R.C., Schmiedt, J.T., Leopold, D.A., Maier, A., and Schmid, M.C. (2021). Theta, but Not Gamma Oscillations in Area V4 Depend on Input from Primary Visual Cortex. Curr Biol *31*, 635-642 e633.
44. Stetson, C., Cui, X., Montague, P.R., and Eagleman, D.M. (2006). Motor-sensory recalibration leads to an illusory reversal of action and sensation. Neuron *51*, 651-659.
45. Ballard, D.H. (1991). Animate Vision. Artif Intell *48*, 57-86.
46. Harman, K.L., Humphrey, G.K., and Goodale, M.A. (1999). Active manual control of object views facilitates visual recognition. Curr Biol *9*, 1315-1318.
47. Burr, D.C., Morrone, M.C., and Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. Nature *371*, 511-513.
48. Wolpert, D.M., Ghahramani, Z., and Jordan, M.I. (1995). An internal model for sensorimotor integration. Science *269*, 1880-1882.
49. Blakemore, S.J., Wolpert, D.M., and Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. Nat Neurosci *1*, 635-640.
50. Straube, B., van Kemenade, B.M., Arikan, B.E., Fiehler, K., Leube, D.T., Harris, L.R., and Kircher, T. (2017). Predicting the Multisensory Consequences of One's Own Action: BOLD Suppression in Auditory and Visual Cortices. PLoS One *12*, e0169131.
51. Scheeringa, R., and Fries, P. (2019). Cortical layers, rhythms and BOLD signals. Neuroimage *197*, 689-698.
52. Schoffelen, J.M., Oostenveld, R., and Fries, P. (2005). Neuronal coherence as a mechanism of effective corticospinal interaction. Science *308*, 111-113.
53. Hanslmayr, S., Volberg, G., Wimber, M., Dalal, S.S., and Greenlee, M.W. (2013). Prestimulus oscillatory phase at 7 Hz gates cortical information flow and visual perception. Curr Biol *23*, 2273-2278.
54. Fiebelkorn, I.C., Pinsk, M.A., and Kastner, S. (2018). A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention. Neuron *99*, 842-853 e848.
55. Fiebelkorn, I.C., and Kastner, S. (2019). A Rhythmic Theory of Attention. Trends Cogn Sci *23*, 87-101.
56. Huang, Y., Chen, L., and Luo, H. (2015). Behavioral oscillation in priming: competing perceptual predictions conveyed in alternating theta-band rhythms. J Neurosci *35*, 2830-2837.
57. Jia, J., Liu, L., Fang, F., and Luo, H. (2017). Sequential sampling of visual objects during sustained attention. PLoS Biol *15*, e2001903.
58. Mo, C., Lu, J., Wu, B., Jia, J., Luo, H., and Fang, F. (2019). Competing rhythmic neural representations of orientations during concurrent attention to multiple orientation features. Nat Commun *10*, 5264.
59. Re, D., Inbar, M., Richter, C.G., and Landau, A.N. (2019). Feature-Based Attention Samples Stimuli Rhythmically. Curr Biol *29*, 693-699 e694.
60. Carrasco, M. (2011). Visual attention: the past 25 years. Vision Res *51*, 1484-1525.
61. He, S., Cavanagh, P., and Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. Nature *383*, 334-337.
62. Guo, B., Lu, Z., Goold, J.E., Luo, H., and Meng, M. (2020). Fluctuations of fMRI activation patterns in visual object priming. Human Behaviour & Brain *1*.

63. Wood, D.K., Gu, C., Corneil, B.D., Gribble, P.L., and Goodale, M.A. (2015). Transient visual responses reset the phase of low-frequency oscillations in the skeletomotor periphery. Eur J Neurosci *42*, 1919-1932.
64. Duzel, E., Penny, W.D., and Burgess, N. (2010). Brain oscillations and memory. Curr Opin Neurobiol *20*, 143-149.
65. Fuentemilla, L., Penny, W.D., Cashdollar, N., Bunzeck, N., and Duzel, E. (2010). Theta-coupled periodic replay in working memory. Curr Biol *20*, 606-612.
66. Lee, H., Simpson, G.V., Logothetis, N.K., and Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. Neuron *45*, 147-156.
67. Gutteling, T.P., Petridou, N., Dumoulin, S.O., Harvey, B.M., Aarnoutse, E.J., Kenemans, J.L., and Neggers, S.F. (2015). Action preparation shapes processing in early visual cortex. J Neurosci *35*, 6472-6480.
68. Gallivan, J.P., Chapman, C.S., Gale, D.J., Flanagan, J.R., and Culham, J.C. (2019). Selective Modulation of Early Visual Cortical Activity by Movement Intention. Cereb Cortex *29*, 4662-4678.
69. Monaco, S., Malfatti, G., Culham, J.C., Cattaneo, L., and Turella, L. (2020). Decoding motor imagery and action planning in the early visual cortex: Overlapping but distinct neural mechanisms. Neuroimage *218*, 116981.
70. Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron *65*, 472-479.
71. Cai, C., Ogawa, K., Kochiyama, T., Tanaka, H., and Imamizu, H. (2018). Temporal recalibration of motor and visual potentials in lag adaptation in voluntary movement. Neuroimage *172*, 654-662.
72. Rao, R.P., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci *2*, 79-87.
73. Brown, H., Adams, R.A., Parees, I., Edwards, M., and Friston, K. (2013). Active inference, sensory attenuation and illusions. Cogn Process *14*, 411-427.
74. Costagli, M., Ueno, K., Sun, P., Gardner, J.L., Wan, X., Ricciardi, E., Pietrini, P., Tanaka, K., and Cheng, K. (2014). Functional Signalers of Changes in Visual Stimuli: Cortical Responses to Increments and Decrements in Motion Coherence. Cerebral Cortex *24*, 110-118.
75. Benson, N.C., Butt, O.H., Datta, R., Radoeva, P.D., Brainard, D.H., and Aguirre, G.K. (2012). The retinotopic organization of striate cortex is well predicted by surface topology. Curr Biol *22*, 2081-2085.
76. Fischl, B. (2012). FreeSurfer. Neuroimage *62*, 774-781.
77. Goebel, R. (2012). BrainVoyager--past, present, future. Neuroimage *62*, 748-756.
78. Jenkinson, M., Beckmann, C.F., Behrens, T.E., Woolrich, M.W., and Smith, S.M. (2012). Fsl. Neuroimage *62*, 782-790.
79. Ashburner, J. (2012). SPM: a history. Neuroimage *62*, 791-800.
80. Brainard, D.H. (1997). The Psychophysics Toolbox. Spat Vis *10*, 433-436.

Supplementary figures

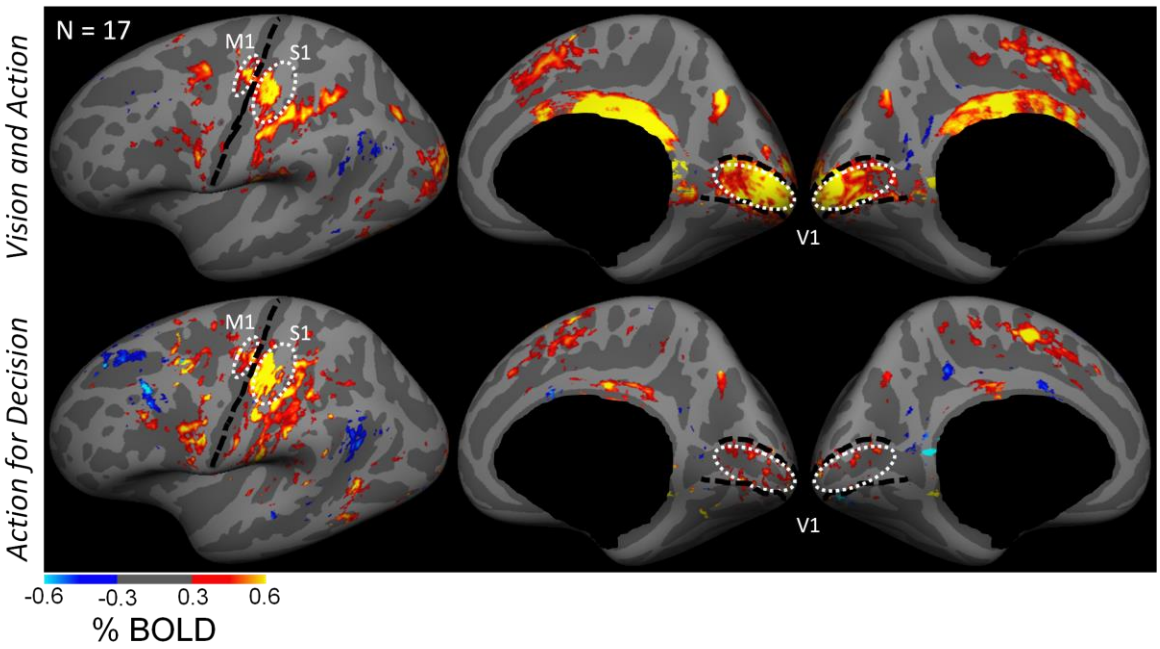

Figure S1. Cortical activations during the task. A: BOLD response (estimated at the peak of the HDR, at 6s) to the *vision and action* event. Black lines report the central sulcus and V1 borders on the surface template (fsaverage_sym). White ovals point to the position of the ROIs when projected onto the surface. BOLD activities are computed on the aggregate observer (N = 17) and masked for statistical significance (FDR correction). **B**: BOLD response to the *action for decision* event corresponding to the keypress performed to report the perceptual decision.

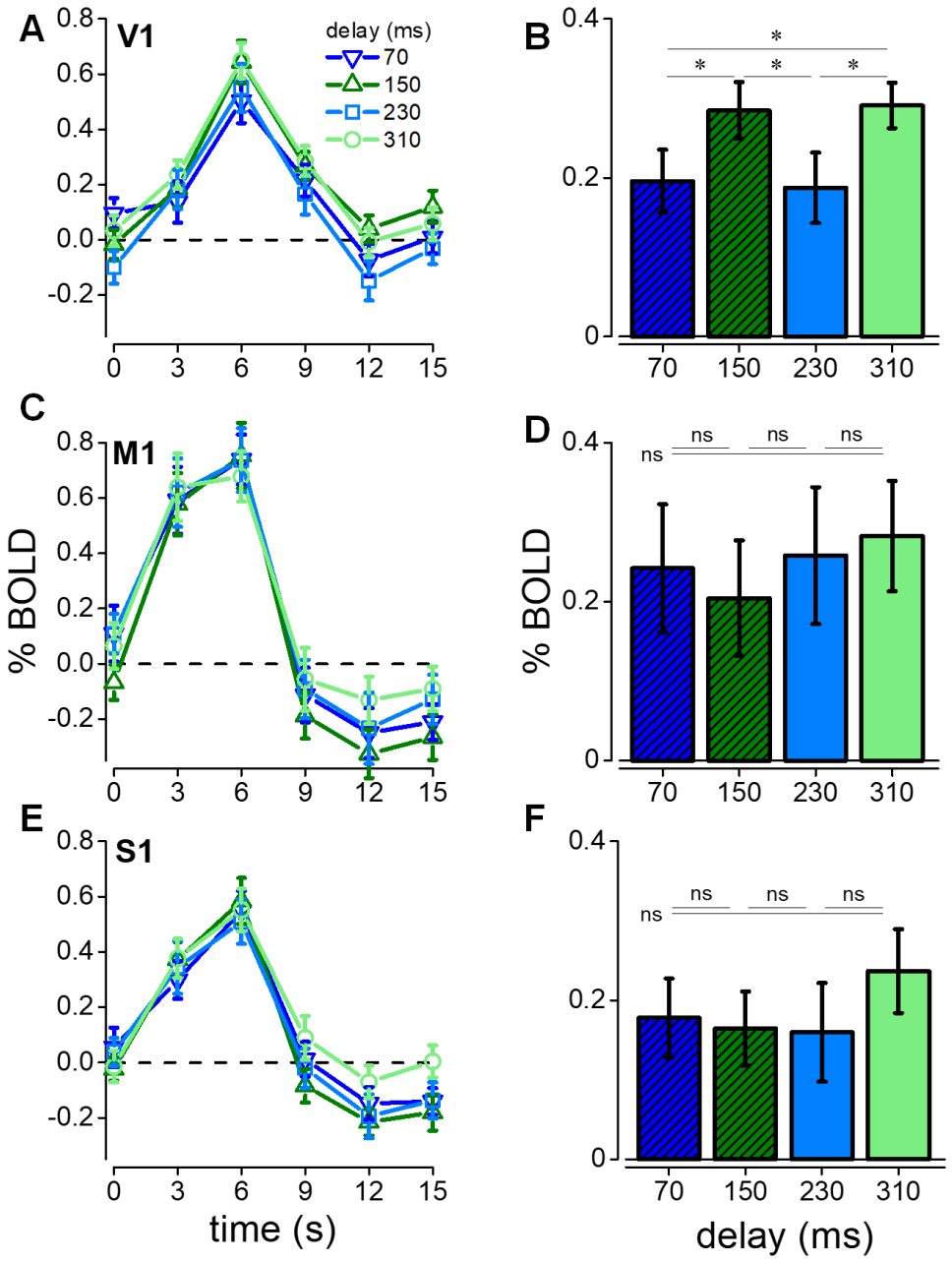

Figure S2. V1-M1-S1 BOLD responses. A-C-E: Timecourse of the GLM beta weights (mean and s.e.m.) representing the BOLD response to *vision and action* events in V1 (panel A), M1 (panel C) and S1 (panel E) for different delays from button press (down triangle: 70 ms, up-triangle: 150 ms; square: 230 ms; circle: 310 ms) averaged across subjects. M1 and S1 are synchronized with the button-press, as in the main analyses reported in the paper. Here V1 events are synchronized with stimulus onset (compare with Figure 2B-C, showing almost identical results). **B-D-F**: GLM beta weights for the different delays for V1 (panel B), M1 (panel D) and S1 (panel F). Asterisks indicate statistical significance (Bonferroni-Holm corrected): ns > 0.05 > * > 0.01.