bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Prefrontal state fluctuations control access to consciousness

Abhilash Dwarakanath[1, *], Vishal Kapoor[1, *], Joachim Werner[1], Shervin Safavi[1,2], Leonid A. Fedorov[1], Nikos K. Logothetis[1, 3§], Theofanis I. Panagiotaropoulos [1,4§]

[1] Department of Physiology of Cognitive Processes, Max Planck Institute for Biological Cybernetics, Tübingen 72076, Germany

[2] International Max Planck Research School, Tübingen 72076, Germany

[3] Division of Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, UK

[4] Cognitive Neuroimaging Unit, CEA, DSV/I2BM, INSERM, Universite Paris-Sud, Universite Paris-Saclay, Neurospin Center, 91191 Gif/Yvette, France

[*] *These authors contributed equally to this work*

§ Co-senior authors

Co-correspondence:

abhilash.dwarakanath@tuebingen.mpg.de

theofanis.panagiotaropoulos@tuebingen.mpg.de

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Abstract

In perceptual multistability, the content of consciousness alternates spontaneously between different interpretations of unchanged sensory input. The source of these internally driven transitions in conscious perception is unknown. Here we show that transient, low frequency (1-9 Hz) perisynaptic bursts in the macaque lateral prefrontal cortex precede spontaneous perceptual transitions in a no-report binocular motion rivalry task. These low-frequency transients suppress 20-40 Hz oscillatory bursts that selectively synchronise the discharge activity of neuronal ensembles signalling conscious content. Similar ongoing state changes, with dynamics resembling the temporal structure of spontaneous perceptual alternations during rivalry, dominate the prefrontal cortex during resting-state, thus pointing to their default, endogenous nature. Our results suggest that prefrontal state fluctuations control access to consciousness through a reorganisation in the activity of feature-specific neuronal ensembles.

One sentence summary

Prefrontal state transitions precede spontaneous transitions in the content of consciousness.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

When the visual system is confronted with ambiguous sensory information, conscious awareness spontaneously fluctuates between different possible perceptual interpretations (*1*, *2*). In an unpredictable manner, one of the competing sensory representations temporarily gains access to consciousness while the others become perceptually suppressed, therefore dissociating sensory input from conscious perception. The mechanism that is responsible for this internally driven passage of sensory input from non-conscious processing to conscious access and vice versa is currently unknown.

Binocular rivalry (BR) is a well-known and commonly studied paradigm of multistable perception in the effort to identify the mechanism that gives rise to conscious awareness. In BR, the content of consciousness spontaneously alternates between two disparate stimuli that are continuously presented in each eye (*3–5*). Several mechanisms localised in different cortical areas, including competition between monocular neurons in the primary visual cortex (V1) and activation of a widespread cortical network driven from a non-linear ignition event in the prefrontal cortex (PFC) have been hypothesised to drive these spontaneous perceptual transitions and conscious access (*6*, *7*). However, monocular V1 neuron competition seems insufficient to explain BR (*8*) which is an instance of perceptual multistability involving competition between higher order perceptual representations, that are not bound to eye-specific input (*9*). Furthermore, although the PFC has been suggested to mediate visual consciousness (*10–16*), whether prefrontal signals contribute to or drive the emergence of conscious awareness, or simply reflect the cognitive consequences of reporting the content of perception like monitoring and introspection is still unclear (*17–22*). The confusion largely stems from probing the PFC using different signals (e.g. hemodynamic response using functional magnetic resonance imaging (fMRI) vs. electrophysiology) during voluntary perceptual report or no-report paradigms of exogenous or intrinsically driven changes in conscious perception (*15*, *19*, *20*, *23–26*).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Here we studied the mechanisms of conscious awareness in the macaque lateral PFC using a no-report BR paradigm. This allowed us to detect intrinsically driven transitions in conscious perception of opposing directions-of-motion. We combined this task with multielectrode recordings of local field potentials (LFPs) and simultaneously sampled direction-of-motion selective, competing ensembles. By using the optokinetic nystagmus (OKN) reflex as an objective criterion of perceptual state transitions, we removed any effects of voluntary motor reports on neural activity, identifying signals directly related to spontaneous transitions in the content of consciousness.

Results

We used a no-report paradigm of binocular motion rivalry coupled with multielectrode extracellular recordings of LFPs and direction-of-motion selective neuronal ensembles in the inferior convexity of the macaque PFC (Fig. 1A). Two types of trials were employed: a) physical alternation (PA) of monocular alternating, opposing directions of motion and b) binocular rivalry (BR) where the initial direction of motion was not removed but was followed by a flashed, opposing direction of motion in the contralateral eye (Fig. 1B, upper panel). Initially, this manipulation results in an externally induced period of perceptual suppression of variable duration for the first stimulus (binocular flash suppression - BFS) which is followed by spontaneous perceptual transitions since the two competing representations start to rival for access to consciousness. In order to exclude the effect of voluntary perceptual reports on neural activity, the macaques were not trained to report their percept. Instead, the polarity of their motion-induced optokinetic nystagmus reflex (OKN) elicited during passive observation of the stimuli (in both conditions, i.e. BR and PA), and previously shown to provide an accurate perceptual state read-out in both humans and macaques (*25*, *27*), was used to infer perceptual dominance periods (Fig. 1B, lower panel). These dominance durations followed a gamma

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

distribution, a hallmark of multistable perception, with a median dominance duration of 1.54±1.28s (median±SD) for spontaneous transitions in BR and 2.25±2.21s for transitions involving exogenous perceptual suppression in BFS (Fig. 1C, SI Fig. 1).

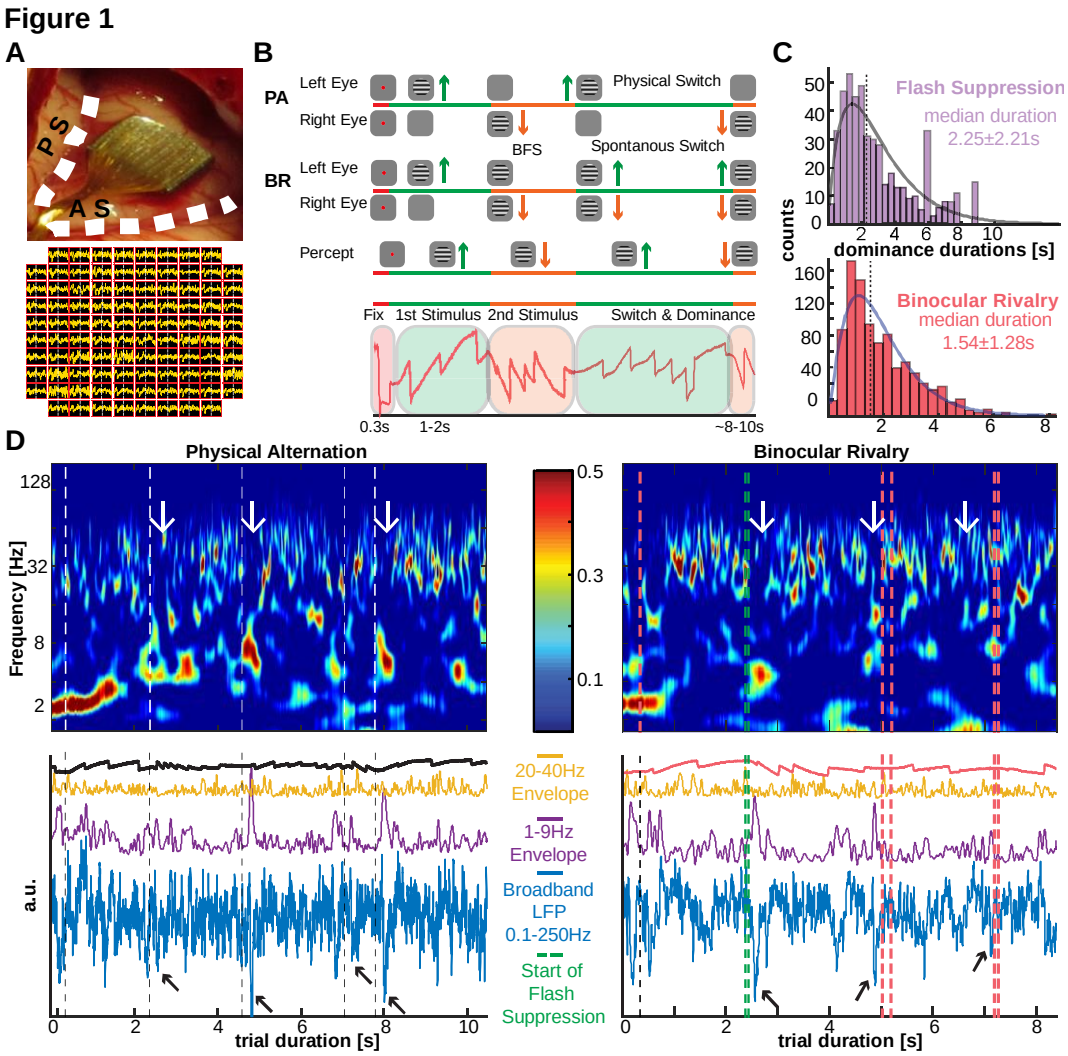


Figure 1 | Experiment and typical perisynaptic signals.

A. Multielectrode array in the inferior convexity of the PFC (top). AS: arcuate sulcus; PS: principal sulcus. Below, example spatial map of 0.1-250 Hz LFP signals around a spontaneous (white line) perceptual transition.

B. Task structure: For both PA and BR trials, following binocular fusion and fixation of a red dot (0.2°) for 300ms a grating (or 200 random dots at 100% coherence in one session) moving upward (90°, green) or downward (270°, orange) was initially presented in one eye (8°, 12-13°/sec, 100% contrast). After 1000-2000 ms the first stimulus was removed in PA trials and followed by the presentation of the opposing motion direction and subsequently multiple monocular switches up to 8-10 seconds. In BR trials, after the presentation of the opposing motion direction (resulting in BFS) the two stimuli were let to compete for conscious access resulting in spontaneous

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

perceptual switches. Different OKN patterns (red trace, highlighted in green and orange) elicited from two directions of motion allowed decoding of the conscious percept.

C. Histogram and gamma distribution function fit of perceptual dominance times during BFS (2.25±2.21s, median±SD) and spontaneous perceptual transitions during BR (1.54±1.28s, median±SD).

D. Channel-averaged normalized (z-score) time-frequency spectrograms for a single trial/observation period of PA (left) and BR (right) is shown at the top. White lines in PA reflect the manually marked change in the OKN polarity after the onset of the exogenous, monocular stimulus changes. Green lines in BR represent the start of the flash suppression phase (at 2.3sec) whereas the red lines represent the subsequent spontaneous perceptual transitions. 1-9 Hz bursts suppressing 20-40 Hz activity occurred following a switch in PA but before a transition in BR (white arrows). Bottom panels: Broadband LFP, instantaneous 1-9 Hz and 20-40 Hz signal amplitude and OKN traces for the same observation period. Black arrows point to negative deflections in the 0.1-250 Hz LFP trace (for display, all traces were normalised and plotted with an arbitrary shift to clearly delineate the different regimes).

Prefrontal state fluctuations precede conscious awareness

To identify mesoscopic prefrontal signals preceding spontaneous perceptual transitions, we first analysed the perisynaptic activity dynamics reflected in the LFPs (Figure 1D, SI Fig. 2). Transient negative deflections of the channel-averaged raw (0.1-250 Hz) LFPs (blue traces in Fig. 1D), suggestive of brief depolarising states, were observed to disrupt a default state of oscillatory bursts in the beta range (20-40 Hz) throughout the observation periods in both PA and BR trials. However, the strongest negative deflections appeared to occur just after the change in the OKN polarity induced by external stimulus changes in PA (Fig. 1D, left) but just before spontaneous perceptual transitions in BR (Fig. 1D, right). Pooling all transitions in PA (n=802) revealed that the power of these transient negative deflections was concentrated immediately after exogenous stimulus changes in a low-frequency (1-9 Hz) range that resulted in a temporally transient suppression of the ongoing beta activity (20-40 Hz) (Fig. 2A, left).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

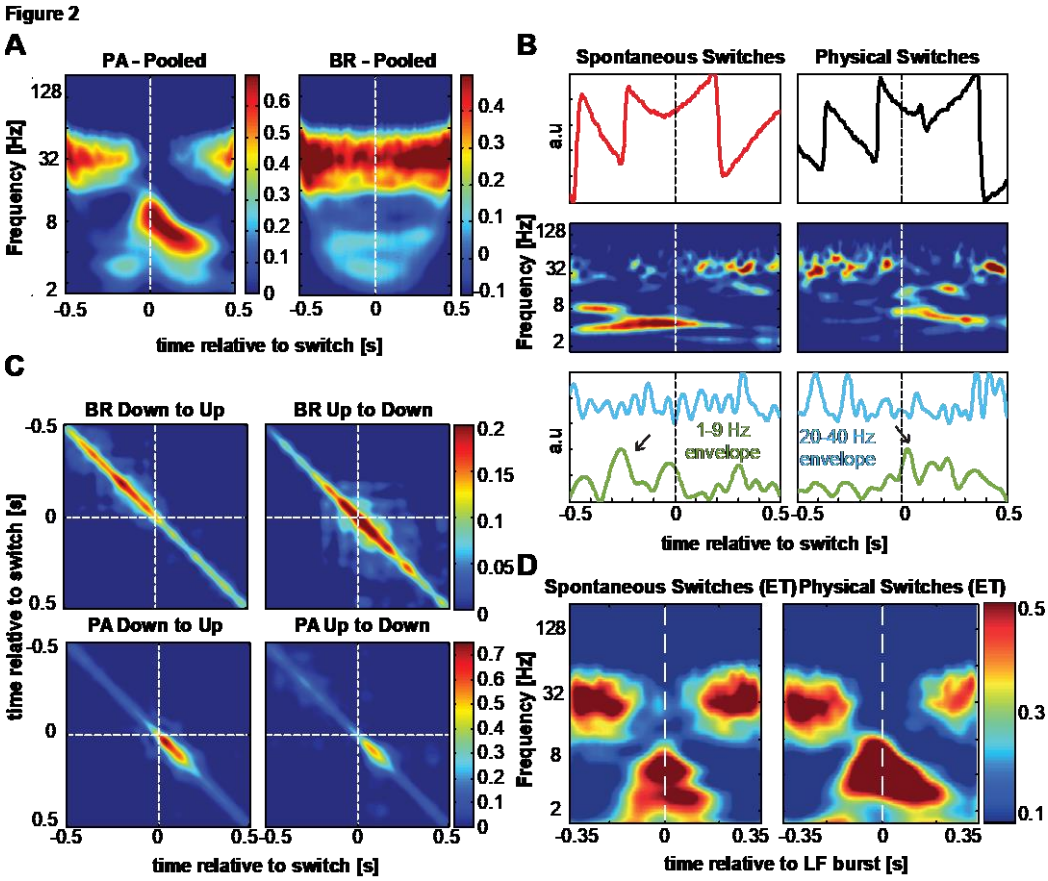


Figure 2| Time-frequency LFP analysis.

A. Grand average time-frequency analysis of all physical (left) and spontaneous (right) perceptual transitions. Spectrograms are aligned (t=0) to manually marked OKN changes in both PA and BR for periods of stable perceptual dominance before and after the switch.

B. Upper panel: OKN traces around a single physical (black) and spontaneous (red) transition. Middle panel: channel-averaged normalised spectrograms aligned to the OKN slope change for the two conditions. Lower panel: Normalised instantaneous amplitudes of the two modulated frequency bands (i.e. 1-9 Hz, green and 20-40 Hz, cyan) identified from the spectrograms. Low-frequency bursts occur after the physical switch but before the spontaneous transition (black arrows).

C. Differences in the onset of the low-frequency activity across physical and spontaneous transitions in direction of motion are reflected in the temporal auto-covariance of the low frequency envelopes across the array recorded simultaneously for every transition type (down to up, left and up to down, right). Most of the similarity is observed before a transition in BR (upper panel) but after a transition in PA (bottom panel).

D. Grand average low-frequency peak aligned spectrograms after a physical and before a spontaneous transition.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Low frequency activation results in beta power suppression for around 300ms both PA and BR.

Low frequency induced beta suppression was also observed for intrinsically generated perceptual transitions in BR (n=573); however, it started well before (~ 400ms) the spontaneous OKN change (Fig. 2B, right). In BR, the absence of a consistent feedforward response, locked to an external change of the sensory input, resulted in a temporal jitter of the low-frequency transients across different transitions and neuronal sites (SI Fig. 3). In individual transitions, the low-frequency-triggered beta-suppression started clearly before the spontaneous OKN transition (Fig. 2B). Indeed, the temporal-covariance of the channel-averaged low-frequency signal across transitions was concentrated well before the perceptual change in BR, whereas it was heavily concentrated in the post-transition period during PA (Fig. 2C). Computing spectrograms aligned to the low-frequency event peaks detected before and after the transition in BR and PA respectively showed the similarity in the coupling of low-frequency transients and beta-burst suppression between the two conditions (Fig. 2D). Therefore, 1-9 Hz suppression of 20-40 Hz activity follows exogenous stimulus changes in PA, but precedes spontaneous OKN- inferred perceptual transitions in BR.

Since both low frequency and beta activity were not sustained oscillations but appeared to occur in bursts, we quantified the burst-rate of low-frequency and beta activity before and after the time of exogenous (PA) and endogenous (BR) perceptual transitions using a burst-rate metric (Methods). Low frequency burst rate (bursts/transition/channel) was significantly higher after the OKN change in PA (0.35±0.0045, n = 26812, post-transition, vs. 0.09±0.0016, pre-transition, n = 6612; $p<10^{-185}$ mean ± SEM), but before the OKN change in BR (0.17 ± 0.002, n = 9667, pre-transition vs 0.14 ± 0.002, n = 7730, post-transition, $p<10^{-43}$ mean ± SEM) (Fig. 3A and SI Figure 4). Furthermore, low frequency bursts were significantly more before a spontaneous perceptual transition than before a physical transition (0.17 ± 0.002, n = 9667, pre-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transition BR vs. 0.09 ± 0.0016, n = 6612, pre-transition PA; $p<10^{-147}$ mean ± SEM). Importantly, low-frequency bursts, including the last-detected burst before a transition, occurred on average even before the end of the last dominance period preceding a transition in BR (end of dominance: -97.4 ± 140ms, low-frequency bursts: -198 ± 133ms, median ± SD, $p<10^{-67}$, Fig. 3B, and Fig. 2B for an example transition. The distributions of all low-frequency bursts and the last event were the same). As expected, low frequency bursts occurred predominantly and significantly after the OKN change in PA (64ms ± 147ms, median ± SD).

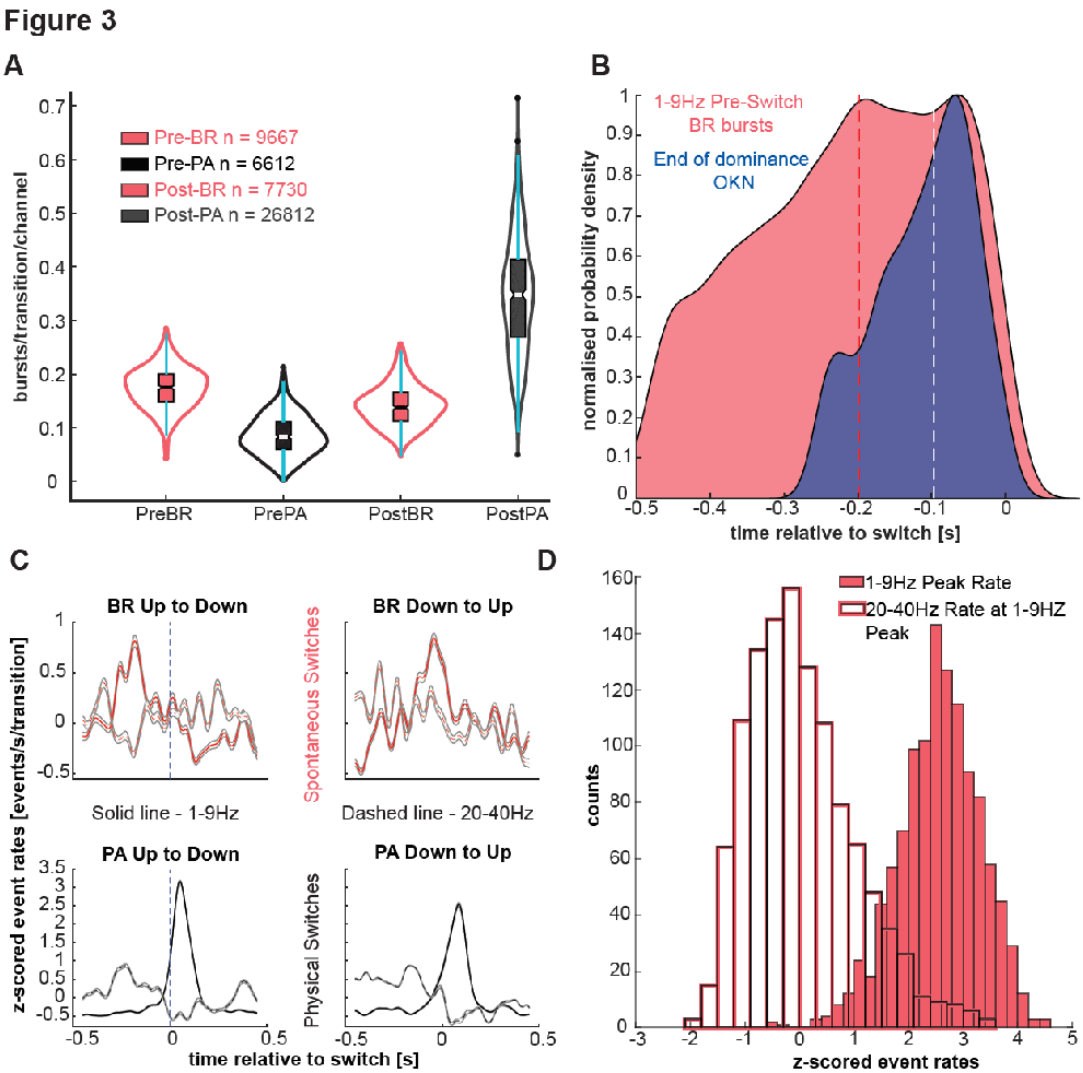


Figure 3| LFP burst rate analysis

A. Burst rate per transition per channel for periods before and after transitions in PA (red) and BR (black). In order from left to right: Pre-BR (dark pink), Pre-PA (black), Post-BR (light pink), Post-PA (grey). The whiskers

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

of the box-plots show the dispersion of the data. White lines depict median values. Matching colour dots represent outliers. More low frequency bursts occur before spontaneous, but after physical switches. Burst rate before a physical switch is very low, suggesting noise levels. This baseline burst-rate needs to be ramped up for a switch to occur. PA: (0.35 ± 0.0045, n = 26812, post-transition, vs. 0.09 ± 0.0016, pre-transition, n = 6612; $p<10^{-185}$ mean ± SEM) BR: (0.17 ± 0.002, n = 9667, pre-transition vs 0.14 ± 0.002, n = 7730, post-transition, $p<10^{-43}$ mean ± SEM).

B. Distribution of low frequency burst-times and OKN times marking the end of the previous dominance period before a spontaneous transition. We fit the probability distribution using a kernel density estimate with a variable width. These functions were then normalised for direct comparison. Low-frequency bursts on average occurred at -198±133ms before a switch, whereas the end of dominance occurred at -97.4±140ms before a transition to the competing representation ($p<10^{-67}$ median ± SD).

C. Normalised (z-score) burst rate in time (events/s/transition) during BR (red lines) and PA trials (white lines) for low-frequency (solid lines) and beta activity (dashed lines).

D. Distribution of beta-band rate at the peak of low-frequency rate before spontaneous perceptual transitions in BR. (r = -0.08, p=0.0071; pooled across both transition types)

The low-frequency activation before a spontaneous perceptual reversal in BR is better observed in the evolution of bursting activity in time (quasi-PSTH, i.e. the detected bursts are converted into binary event trains, smoothed and then averaged). In BR, the peak-rate of 1-9 Hz bursts occurred at -160 ± 237ms and -28 ± 199ms (median ± SD) before the spontaneous perceptual transitions for the two transition types respectively (Figure 3C, top row), while in PA at 52 ± 28ms (median ± SD) and 82 ± 64.5ms (median ± SD) following marked OKN change (Fig. 3C, bottom row). These differences were further enhanced when the bursts towards the end of the post-switch window (bursts occurring 150ms after the transition) were discarded (SI Fig. 5). Confirming the time-frequency analysis pattern in Fig. 2A and suggesting a frequency-specific competitive process (i.e. cortical state fluctuations) in the PFC, low-frequency and beta burst-rates were significantly anti-correlated in BR (r = -0.08, p = 0.0071;

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

pooled across both transition types; Fig. 3D).

Spatiotemporal build-up of prefrontal activity

Are the perisynaptic transients preceding a spontaneous change in the content of consciousness random large excursions from baseline (noise) activity or do they reflect a gradual spatiotemporal build-up process that is critical for inducing a spontaneous transition? Indeed, we noticed that in many instances before a spontaneous transition, the last transient low-frequency (1-9 Hz) burst was frequently preceded by similar but of lower amplitude bursts (Fig. 1D and SI Fig. 1, 6). When the instantaneous amplitude of the low-frequency activity triggered at every switch was averaged first across populations and then across transitions, we observed a gradual increase approaching spontaneous but not physical transitions (Fig. 4A). Fitting a linear model to the relationship between the transition-averaged low frequency burst amplitudes at every time point before a transition revealed a linearly increasing relationship between the two variables before a spontaneous (adjusted $R^2 = 0.34$) but not before a physical switch (adjusted $R^2 = - 0.003$) (Fig. 4B). While this low frequency burst amplitude exhibited a gradual linear increase, the number of activated prefrontal sites suddenly increased just before a spontaneous reversal, suggesting a non-linear increase in the spatial spread of prefrontal activation just before a spontaneous change in the content of consciousness (Fig. 4C and SI Fig. 7). Taken together, these results point to a mesoscopic, spatio-temporal spread of low frequency prefrontal bursts before spontaneous perceptual reversals. Both linear and non-linear effects in the amplitude and spatial spread of activation respectively suggest the operation of both in a prefrontal ignition (*6*) process during BR.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

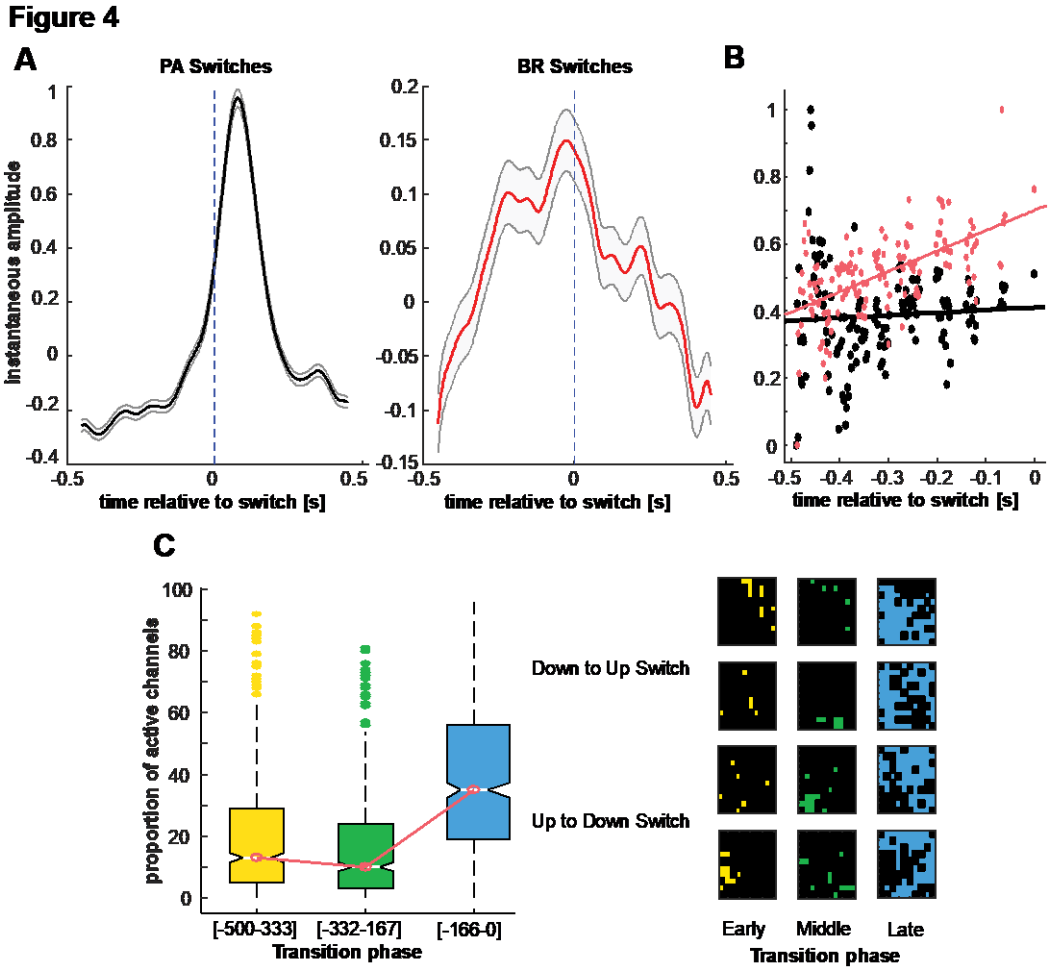


Figure 4| Spatiotemporal build-up preceding spontaneous transitions

A. Low-frequency instantaneous amplitude shows a slow climbing activity before a perceptual transition (right) but not before a physical transition (left). Curves reflect an average ± SEM across transitions of the channel-averaged activity for each collected transition.

B. Average build-up of low-frequency activity in time by fitting a linear model to the pooled and averaged amplitude at every time-point (red – BR, black – PA) across all measured neuronal sites. While before a spontaneous transition, the recorded PFC area ramps up its low-frequency activity in time, before a physical transition, it remains flat.

C. Spatio-temporal spread of low frequency activity shown in the proportion of channels for a given transition that displayed low-frequency burst peaks in early ([-500 to -333ms]), middle ([-332 to -167ms]) and late temporal windows ([-166 to 0ms]). Data are pooled across the two macaques. A significantly larger proportion ($p<0.01$) of neuronal sites peak closer to the switch, showing a sharp, non-continuous increase in the number of recruited channels (left). Activation of sites on the array shown for four examples of different types of spontaneous

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transitions from 2 macaques (right).

We further hypothesised that if an increase in low-frequency bursting is critical for inducing spontaneous perceptual reversals, then low-frequency amplitude should be significantly weaker when perceptual transitions were not complete, but resulted in piecemeal (PM) periods in which perception did not unambiguously favour either of the two competing directions of motion (Fig. 5A). Subtracting the time-frequency decomposition of transitions to a PM percept from that of clean BR perceptual transitions revealed a preponderance of low-frequency activity before a transition, suggesting that the amplitude of low frequency bursts is critical for completing a perceptual transition to another period of clear dominance. Moreover, we could recover the pattern observed during spontaneous transitions, when we subtracted from the spectrogram of these periods the averaged spectrogram of randomly triggered periods during sustained perceptual dominance (Fig. 5B). This suggests that low-frequency activity should be significantly up-modulated from noise level, potentially crossing a threshold to cause a perceptual transition (SI Fig. 8).

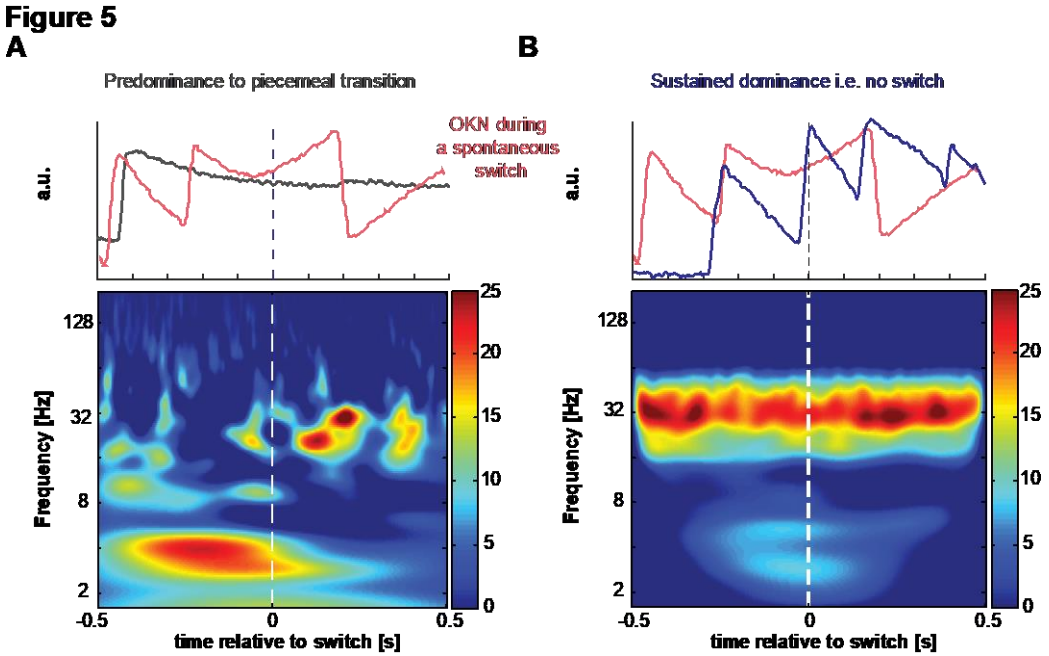


Figure 5| Low frequency amplitude is critical for inducing clear switches.

A. Top panel shows two typical OKN patterns elicited during a spontaneous transition (red) and during a transition

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

to piecemeal (grey). The subtracted spectrogram (BR minus PM) below shows a large difference in low-frequency activity before a spontaneous switch suggesting very weak low frequency activation during a transition to piecemeal perception.

B. Top panel shows two typical OKN patterns elicited during a spontaneous transition (red) and during sustained dominance where no switches occurred (blue). The subtracted spectrogram (BR minus no switches) recovers the time-frequency pattern observed during BR suggesting that the LFP activity during sustained perceptual dominance periods is at noise level.

Conscious content specific ensemble activity transitions and LFPs

To understand the temporal relationship between perisynaptic state fluctuations and neuronal populations reflecting the content of consciousness, we compared the convergence times (see Methods) of the normalised discharge activity of simultaneously recorded ensembles selective for the rivalling directions-of-motion (Fig. 6A, SI Fig. 9), to the low-frequency burst and peak-rate distributions. We found that discharge activity converged significantly later compared to both the low-frequency peak-rates (-60 ± 222ms for LFP event/s/transition vs. 77 ± 76ms median ± SD, convergence of spiking; $p<10^{-88)}$ and burst-times (-112 ± 190ms median burst time (truncated at 250ms post-switch) vs. 77 ± 76ms, median ± SD convergence of spiking, $p<10^{-74}$) in spontaneous perceptual transitions. In a large number of spontaneous transitions (86.2% compared to peak-rates and 89% compared to burst times), spiking activity crossovers occurred after the median truncated low-frequency peak-rates and burst-times (Fig. 6B). Therefore, low-frequency transients could reflect a pre-conscious process preceding an intrinsically generated perceptual transition, i.e. a change in the content of consciousness by feature-selective neural ensembles.

Are these low-frequency LFP modulations non-specific or do they show a relationship to the patterns seen in the spiking activity of selective ensembles that reflect the current content of consciousness? To answer this question, we contrasted the selective-ensemble-summed low-

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

frequency instantaneous amplitude to their respective summed spiking PSTHs. For the same transitions, the respective ensemble spiking activity shows a clear crossover around the OKN transition (both in BR and PA, Fig. 6C), pointing to a change in encoding of the dominant percept, while no such divergence is observed in the low frequency activity that is identical for the two populations. Low frequency activity is therefore distinguished from the spiking of selective neurons, and most likely reflects a pre-conscious process.

Finally, to understand the mechanism that prefrontal state fluctuations could result in spiking-network and therefore perceptual reorganisation, we computed the spike-field coherence (SFC) (*28*, *29*) of the simultaneously recorded, feature-selective ensemble activity and the global broadband LFP across all transitions. After a spontaneous perceptual transition, when the negative LFP deflections and therefore the low frequency (1-9 Hz) transients were less prevalent, the perceptually dominant ensemble was more coherent in the beta range (~25-40 Hz) compared to the suppressed ensemble ($p < 0.03$; Fig. 6D). However, in the period approaching a spontaneous transition when low frequency transients suppressed beta bursts, there were no differences between suppressed and dominant populations. These results suggest that prefrontal ensembles signalling the current content of consciousness are synchronised in the beta-band of the LFP. Low-frequency transients dissolve these beta-coherent ensembles, potentially increasing the likelihood for spiking in the suppressed population and therefore the likelihood for perceptual reorganisation.

Intrinsic nature of prefrontal state fluctuations

If the competition between low-frequency transients and beta-bursts found to control access to visual awareness is intrinsically generated, reflecting waking state fluctuations, traces of this process should also be observed during resting-state: i.e., in the absence of structured visual input. Indeed, in resting-state, low-frequency bursts suppressed beta bursts (Fig. 6E).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Periods of uninterrupted beta activity exhibited a gamma distribution with a duration of 1.2 ± 1.44s (median ± SD), which is close to the psychophysical distribution of stable perceptual dominance durations (1.54 ± 1.28s) that are also characterised by uninterrupted beta activity (Fig. 6F). Therefore, prefrontal state fluctuations reflect an intrinsically generated process in the macaque PFC.

Figure 6

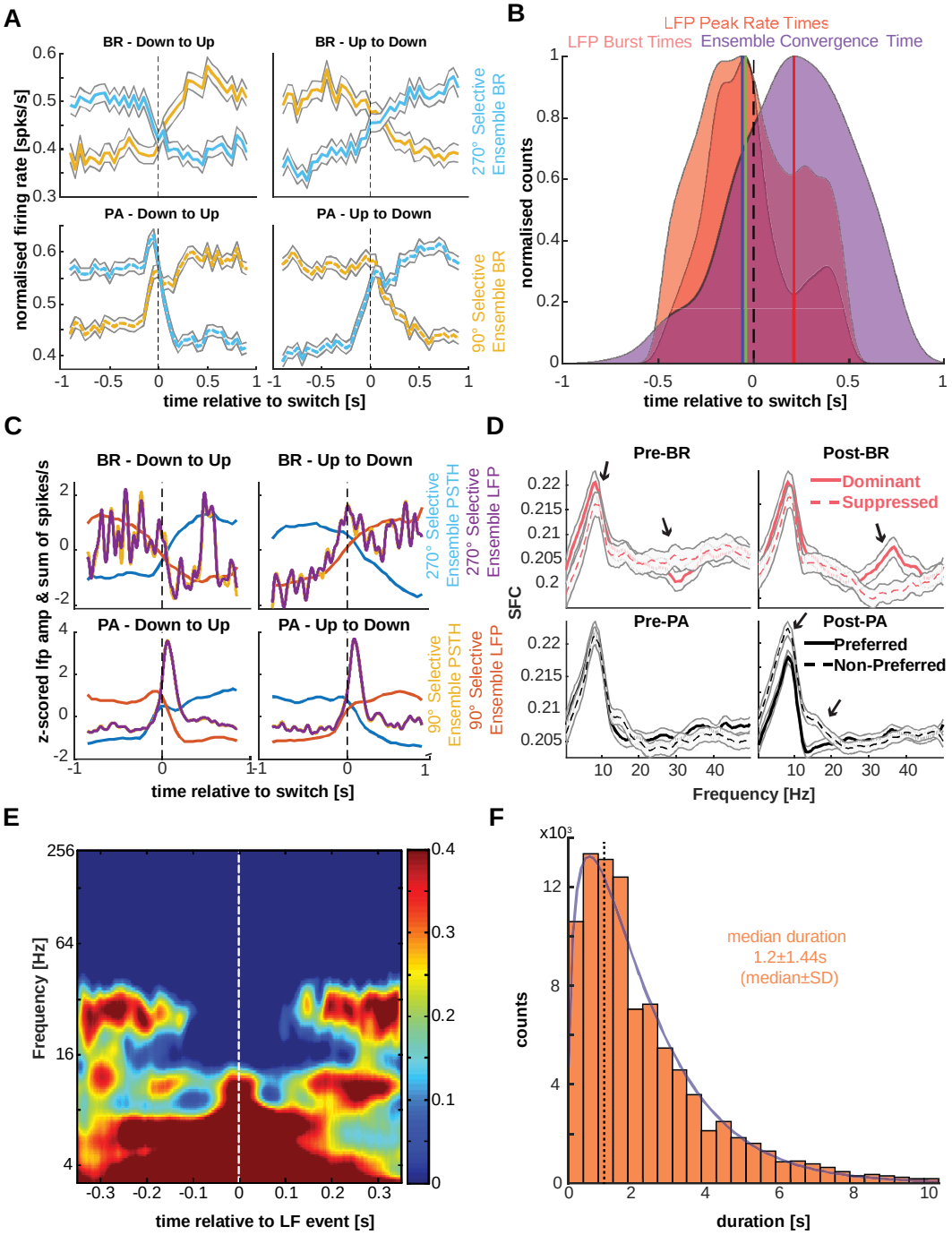


Figure 6| Ensemble spiking activity-LFP relationship and resting state analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

A. Top row: Average (across all collected transitions) population PSTHs of simultaneously recorded feature specific neuronal ensembles during BR. For each transition, we summed spiking activity across all selective single neurons belonging to each competing ensemble (upward and downward selective, blue and yellow, respectively). Bottom row: same for PA.

B. Distribution of the truncated (-500 to 250ms around the OKN transition) low-frequency burst times (pink), event-rate peak times (red) and the time of convergence of the PSTHs of each competing ensemble for each transition estimated by LOWESS smoothing and interpolation. Low-frequency activation preceded a change in the content of consciousness from neuronal populations (-60 ± 222ms for LFP event/s/transition and -112 ± 190ms vs. 77±76ms median ± SD, convergence of spiking; $p<10^{-74}$ for both comparisons).

C. Although the PSTHs of the two competing ensembles for both BR and PA show a clear divergence in their activity after the transitions, the low-frequency instantaneous amplitudes of the same ensembles show no significant differences in their pattern, pointing to a non-specific and global role of these LFPs in driving perceptual transitions.

D. After a BR transition, spiking of the dominant population was significantly more coherent with the beta LFP band compared to suppression (black arrows point to frequency bins with statistically significant differences between dominant and suppressed $p<0.05$). SFC during pre-switch BR periods when low-frequency transient bursts are more prevalent did not exhibit similar differences in the beta band. These effects in beta band were absent in physical transitions where SFC for a dominant preferred stimulus was significantly reduced in a lower frequency range.

E. Grand-average (n=480), low frequency burst-triggered spectrogram during resting state showing cortical state fluctuations between low frequency and beta activity in the absence of structured sensory input.

F. Periods of sustained beta activation (1.2 ± 1.44s, median ± SD) follow a gamma distribution (BIC (Bayesian information criterion (*30*)) = 2.3×10^5 for a gamma distribution vs BIC = 2.9×10^5 for an exponential) during resting state (activity periods longer than 10s were discarded to maintain equivalence with the experimental trial durations). The median duration is remarkably similar to the psychophysical gamma distribution (1.54 ± 1.28s, median ± SD).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Discussion

Contrary to the notion that PFC activity reflects the consequences of conscious perception (*18*, *19*, *21*, *26*, *31–33*), our results reveal that perisynaptic prefrontal state fluctuations precede spontaneous changes in the content of consciousness, with the latter reflected in feature specific prefrontal ensembles, in the absence of any voluntary report requirements. Therefore, the PFC appears to have a centrally important role in the emergence of conscious perception. Previous electrophysiological studies in the PFC revealed representations of the content of consciousness using exogenous perceptual manipulations (*23*, *34*) and preparatory activity before spontaneous perceptual changes, that could not however be dissociated from signals related to voluntary motor reports (*24*, *35*). Furthermore, studies using BOLD fMRI that indirectly reflects neural activity disagree on the impact of voluntary reports on PFC activity during spontaneous perceptual transitions (*12*, *19*). Here we used a no-report spontaneous perceptual transition paradigm coupled with multielectrode recordings of local field activity and feature-specific neuronal ensembles that was instrumental for revealing the mechanistic details of conscious awareness emergence.

Our findings reveal a gating-like mechanism where intrinsically generated cortical state fluctuations between low-frequency transients and periods of sustained beta-bursts control the access of competing perceptual representations to consciousness. The spatiotemporal build-up of 1-9 Hz bursts before spontaneous perceptual transitions is reminiscent of the Readiness Potential or *Bereitschaftspotential*; i.e., a steady accumulation of activity in the Anterior Cingulate Cortex (ACC) and Supplementary Motor Area (SMA) preceding the awareness of the volition to initiate a report (*36–39*). This motor-related process is reflected in a spiking activity build-up before the voluntary motor report of a perceptual transition in BR (*24*). Our results show that both linear and non-linear increases in the instantaneous amplitude and spatial-spread of low-frequency prefrontal bursts respectively precede spontaneous transitions

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

in the content of consciousness without voluntary behavioural reports.

Periods of beta burst suppression are thought to reflect a decrease in endogenous cortical processing, since beta is suppressed during cognitive processes like attention, decision-making and movement-planning (*40–43*). Beta activity could therefore reflect an intrinsic mode of cortical operation that shields ongoing behavioural and processing states ("status-quo") from interference and distractors (*40*, *41*). As a result, transient decreases in cortical beta activity could increase sensory information relay (*44*, *45*, *46*, *47*) providing a mechanism for controlling bottom-up sensory processing through top-down knowledge (*48*). Similar desynchronised brain states could mediate low-level awareness (*49*). Indeed, we found that a dissolution of beta-coherent dominant ensembles from low-frequency transient bursts precedes intrinsically generated changes between the two perceptual states. This is reminiscent of rule-selective prefrontal ensembles that are coherent in the beta band with low-frequency activity inhibiting a rule that is about to be deselected (*50*). This suggests that the underlying prefrontal mechanism for the emergence of conscious awareness and cognitive control might be the same. We propose that the intrinsically-generated prefrontal beta activity could reflect the neural correlate of the prefrontal threshold mechanism that has been long hypothesised to control access to consciousness (*34*, *51–54*). Disruption of intrinsic prefrontal beta from transient depolarisation events could provide temporal windows for a reorganisation in the discharge activity of neuronal ensembles that encode the competing representations and consequently increase the likelihood for a perceptual transition to happen. Indeed, beta-suppression occurred earlier than the flip in the population discharge activity encoding the content of consciousness (SI Fig. 10). This top-down mechanism of perceptual reorganisation is fundamentally different from bottom-up mechanisms proposing that competition between monocular neurons in primary visual cortex (V1) resolves BR (*7*, *8*). Neuronal activity in V1 is indeed only weakly modulated during BR in both monocular and binocular neurons (*8*) while

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

BOLD modulation of V1 is detected in superficial layers suggesting feedback from higher cortical areas (*55*). Furthermore, optical imaging signals in V1 during BR can also be observed during anaesthesia (*56*), suggesting that V1 activity is not sufficient for conscious visual perception.

Finally, spontaneous cortical activity can attain various states during wakefulness, and mimic sensory-driven activity (*57–59*). We observed suppression of beta bursts by low-frequency transients also during periods of resting state. This suggests that the source of spontaneous transitions in the content of consciousness is the operation of waking state fluctuations in the PFC. Taken together, our results reveal a pivotal role of prefrontal state fluctuations in the emergence of conscious awareness.

Methods

Electrophysiological recordings.

We performed extracellular electrophysiological recordings in the inferior convexity of the lateral PFC of 2 awake adult, male rhesus macaques (*Macaca mulatta*) using chronically implanted Utah microelectrode arrays (*60*) (Blackrock Microsystems, Salt Lake City, Utah USA). We implanted the arrays 1 - 2 millimetres anterior to the bank of the arcuate sulcus and below the ventral bank of the principal sulcus, thus covering a large part of the inferior convexity in the ventrolateral PFC, where neurons selective for direction of motion have been previously found (*61*, *62*). The arrays were 4x4mm wide, with a 10 by 10 electrode configuration and inter-electrode distance of 400µm. Electrodes were 1mm long therefore recording from the middle cortical layers. The monkeys were implanted with form-specific titanium head posts on the cranium after modelling the skull based on an anatomical MRI scan acquired in a vertical 7T scanner with a 60cm diameter bore (Biospec 47/40c; Bruker Medical,

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Ettlingen, Germany). All experiments were approved by the local authorities (Regierungspräsidium, protocol KY6/12 granted to TIP as the principal investigator) and were in full compliance with the guidelines of the European Community (EUVD 86/609/EEC) for the care and use of laboratory animals.

Data acquisition, spike sorting and local field potentials.

Broadband neural signals (0.0001–30 kHz) were recorded using Neural Signal Processors (NSPs) (Blackrock Microsystems). Signals from the Utah array were digitised, amplified, and then routed to the NSPs for acquisition. For the offline detection of action potentials, broadband data were filtered between 0.6 and 3 kHz using a second-order Butterworth filter (the filter was chosen such that it allowed a flat response in the passband while contributing the least phase distortion due to its low order, yet having an acceptable attenuation in the stop band, i.e. a roll-off starting at -20dB). The amplitude for spike detection was set to five times the median absolute deviation (MAD) (*63*). Spikes were rejected if they occurred within 0.5 ms of each other or if they were larger than 50 times the MAD. All of the collected spikes were aligned to the minimum. Automatic clustering to detect putative single neurons was performed by a Split and Merge Expectation-Maximisation (SMEM) algorithm that fits a mixture of Gaussians to the spike feature data which consisted of the first three principal components (*64*) (Klustakwik). The clusters were finalised manually using a cut-and-merge software (*65*) (Klusters). For the analysis of perisynaptic LFP activity, the broadband signal was decimated to 500 Hz sampling rate using a Type I Chebyshev Filter, preserving frequency components up to 200 Hz.

Visual stimulation and experimental paradigm.

Visual stimuli were generated by in-house software written in C/Tcl and used OpenGL implementation. Stimuli were displayed using a dedicated graphics workstation (TDZ 2000;

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Intergraph Systems, Huntsville, AL, USA) with a resolution of 1,280 × 1,024 and a 60 Hz refresh rate. An industrial PC with one Pentium CPU (Advantech) running the QNX real-time operating system (QNX Software Systems) controlled the timing of stimulus presentation, and the digital pulses to the electrophysiological data acquisition system. Eye movements were captured using an IR camera at 1kHz sampling rate using the software iView (SensoriMotoric Instruments GmBH, Germany). They were monitored online and stored for offline analysis using both the QNX-based acquisition system and the Blackrock data acquisition system. We were able to capture reliably the eye movements of the animals by positioning the IR camera in front of a cold mirror stereoscope.

Initially, the two monkeys (A11 and H07) were trained to fixate on a red square of size 0.2° of visual angle about 45cm away from the monitors that could be viewed through the stereoscope. This dot was first presented in one eye (the location of the red fixation square was adjusted to the single eye vergence of each individual monkey) and the eye-position was centred using a self-constructed linear offset amplifier. While the monkey was fixating the dot was removed and immediately presented in the other eye. Over multiple presentations, the offset between the two eyes was averaged to provide a horizontal correction factor to allow the two dots to be perfectly fused within the resolution limitations of the recording device (1/100th of a degree). The monkeys were trained to maintain fixation within a window of 2° of visual angle during initiation. After 300ms of fixation, a moving grating of size 8°, moving in the vertical direction (90° or 270°) at a speed of 12° (monkey H) and 13° (monkey A) per second, with a spatial frequency of 0.5 cycles/degree of visual angle and at 100% contrast was presented for 1000-2000ms. This marked the first monocular stimulus epoch in both conditions, viz. Binocular Rivalry (BR) and Physical Alternation (PA). At the end of 1000-2000ms, the second stimulus with the same properties as above but moving in the opposite direction was presented to the other eye. In the BR trials, this marked the “Flash Suppression” phase. These two

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

competing stimuli were allowed to rival with each other for a period of 6000-8000ms. In the PA trials, switches in the percept were mimicked by alternatively removing one stimulus based on the mean dominance time computed from the Gamma Distributions (tailored to each monkey's performance and statistics) acquired during multiple training sessions, and adjusted to be closer to a mean of 2000ms. Free viewing within the 8° window elicited the Optokinetic Nystagmus (OKN) reflex concomitant to the perceived direction of motion which served in lieu of a voluntary report, fulfilling the criterion of a "no-report paradigm". The monkeys were given a liquid reward (either water or juice) at the end of the trial, if their OKN successfully remained within the specified viewing window during the entire duration of the trial. Every successful trial was followed by a 2000-2500ms inter-trial period.

Detection of spontaneous transitions

The recorded eye-movement signal in the Y-coordinate was first low-pass filtered using a 3rd order Butterworth Filter below 20 Hz to remove involuntary jitter-induced high-frequency noise. A custom GUI written in MATLAB allowed us to manually identify the end of a dominance period and the beginning of the subsequent one. Manual marking (performed by two authors, AD and VK) was necessitated due to the large variability in the shapes that comprised the OKN complex. These events were based on the change in the slope of the slow-phase of the OKN. Such spontaneous switches were identified by the difference in the end of a dominance and the beginning of the next one; specifically, if this difference was less than 250ms (a fast switch). A "clean" transition was designated if the previous dominance and the subsequent one lasted for at least 500ms without being broken. Analogous to subjective reports, we aligned the LFP and the spiking activity on the beginning of the subsequent dominance period. This was performed in the same way for both BR and PA trials

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Treatment of the LFP data

Firstly, the decimated LFP signal (0.1-500 Hz) around the OKN transitions was decomposed into a time-frequency representation using a Continuous Wavelet Transform (CWT, MATLAB 2016b) with a Morse wavelet of 7 cycles. This allowed us to resolve 169 frequencies from 0.5 to 256 Hz (500 Hz sampling rate) while preserving the full temporal resolution. The CWT for each channel in each transition (BR, PA, PM and RT) was first z-scored in the frequency domain to visualise the relative changes in power and then pooled across all channels and averaged. To visualise the differences between spontaneous transitions, piecemeals and randomly-triggered periods, the latter two spectrograms were subtracted from the former, respectively.

To understand the evolution of the LFP activity, we first filtered the broadband LFP trace into two constituent bands that were identified to be modulated during the task from the time-frequency analysis, i.e., the low-frequency (1-9 Hz) and the beta band (20-40 Hz). We used a 4th and 8th order Chebyshev Type I filter respectively, with a maximum passband ripple of 0.001dB. To get the instantaneous amplitude in time, we transformed the signal into the Hilbert space and then computed the absolute value. Bursting events were detected at each transition in each channel using a threshold which was 4 times the standard deviation of the noise modelled as a Gaussian distribution. The minimum duration of each event to be detected was set as one full cycle of the highest frequency in that band, i.e. 111ms for the 1-9 Hz band and 25ms for the 20-40 Hz band (*66*). The event-rate in time was computed as a quasi-PSTH by turning the detected bursts into a binary spike-train and smoothed with a Gaussian kernel of width 25ms, and then averaged across all channels. The burst rate was computed as the sum of low-frequency bursts normalised by the number of transitions and channels (bursts/transition/channel). To compute the build-up in the low-frequency activity, the amplitude at each detected time-point was averaged first across all channels for a given

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

transition, and then averaged across all transitions. A line was then fit to this mean scatter-plot using the CurveFit Toolbox in MATLAB.

Construction of direction of motion specific neural ensembles

Single neuron selectivity was assessed during perceptual transition periods of binocular rivalry (perceptual switches) and physical alternation (stimulus switches). During binocular rivalry trials, these periods were selected according to the following criteria: 1. Perceptual dominance (judged from the OKN signal) must be maintained for at least 1000 milliseconds post a perceptual switch 2. A preceding perceptual dominance for the competing stimulus must be maintained for 1000 milliseconds, and finally 3. The delay between the end and the beginning of the two dominance phases was not more than 250 milliseconds. For physical alternations, we selected trials, wherein a stimulus was presented for at least 1000 milliseconds before and after a stimulus switch. The spiking activity was triggered at the beginning of a forward dominance (BR) and stimulus change (PA).

Selectivity was assessed by comparing the distributions of the total number of spike counts across trials where the upward drifting grating was perceived, post (0 to 1000 ms) or pre switch (-1000 to 0 ms), with trials where a downward drifting grating was perceived. We used a Wilcoxon rank sum test and all neurons where $p<0.05$ were considered as selective. For a given transition, spikes were binned in 50ms bins for each selective neuron, and the resultant spike-count histograms were summed across the neurons that make up each selective ensemble to represent a population vector.

To analyse the crossover times between the two competing populations, we computed the trend in these normalised direction-selective population sum PSTH activity for every transition in a 900ms window around the time of the marked smooth pursuit OKN change [-900 to 900] by smoothing the raw ensemble population vectors for the two competing populations using a LOWESS filter. Next, we detected each intersection between these two

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

given vectors using standard interpolation. Where multiple intersection points were detected, only that point was considered which was followed by divergences for a minimum of 200ms before and after the intersection point, denoting distinct encoding of the currently active percept.

Spike-field Coherence

The spike-field coherence (SFC) was computed between the spiking activity of selective ensembles for each transition, and the global LFP for that particular transition averaged over all electrodes. A rate adjustment and a finite-size correction was applied before computing the SFC via a multi-taper method (*67*) (Chronux Toolbox).

Treatment of resting-state activity

LFPs from two continuously-recorded resting state sessions on days when no task-recording was performed, were decimated to 500 Hz as mentioned above. In each channel, the beta bursts were detected using the previously-mentioned LFP event-detection algorithm. The mean of the inter-event interval was used as a threshold to decide which collection of events constituted a phase of sustained activity. These epochs were collected across all channels and pooled across the two monkeys. Both a gamma and an exponential distribution were fit to the observations, with the gamma distribution clearly revealing lower AIC and BIC measures, thereby pointing to this distribution being a better fit than the exponential.

Statistical methods

All statistical comparisons were performed using a Wilcoxon ranksum test (*68*) due to the unknown nature of the underlying distribution from which the data originated. Distributions were fit using the MATLAB statistical toolbox using a Maximum-Likelihood-Estimate

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

method. For model comparisons, the allFitDist.m toolbox was used that also generated metrics for appropriate model selection. For non-parametric fitting of distributions with widely different sample numbers, the kernel density estimate method implemented in the MATLAB Statistics Toolbox was used to generate the best-fit function, which was then normalised for visualisation.

SI Section - Controlling for failed and non-occurrence of switches (SI Figure 8)

Is there a difference in the type and magnitude of low-frequency activity leading to a perceptual transition as compared to when it never happens? To this end, we randomly triggered the LFP activity in observation periods where no such spontaneous transitions occurred, i.e. during periods of sustained predominance. Subtracting the time-frequency decomposition of these randomly triggered (RT) periods from that of BR preserved the pattern observed in the latter, suggesting that weak, low-frequency activity occurs as baseline noise, which only leads to a perceptual change when it is spatiotemporally ramped up in a structured manner (Figure 4B vs Figure 2A). Indeed, we computed a mean rate of 0.015±0.00005 (n = 55026 after considering 100 iterations, i.e. only 550.26 bursts per iteration) bursts per transition; an order of magnitude lower than the corresponding periods during BR (0.17±0.002, n = 9667, bursts/transition, SI Figure 8D). Furthermore, the proportion of sites that displayed low-frequency bursting activity across all BR periods was 100%, compared to only 51% in RT periods. These results further confirm that the low-frequency burst-rate, build-up, and a larger spatial spread of activation is necessary to drive spontaneous transitions.

Furthermore, the low-frequency burst-rate was higher before a clear spontaneous transition compared to the period before transition to a piecemeal percept (0.17±0.002, n = 9667, pre-BR, vs. 0.14±0.004, pre-PM, n = 2486, bursts/transition; $p<10^{-25}$) (SI Figure 8A), with the low-frequency peak rate occurring after the transition to piecemeal (SI Figure 8B).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Furthermore, the burst rate was significantly higher after the transition to a piecemeal period (0.16±0.004, PM, n = 3531, vs 0.14±0.004 pre-PM, n = 2486, bursts/transition; $p<10^{-5}$), while the anti-correlation between low-frequency and beta was significant but weaker compared to clear spontaneous transitions ($r = -0.009$, $p<10^{-137}$ vs $r = -0.05$, $p<10^{-295}$, $p < 10^{-14}$, Figure 3D and SI Figure 8C).

SI Section – PA Switches aligned to the TTL pulses (SI Figure 6)

The PA switches used for analysis in the manuscript were aligned to the manual marks. In addition, we also aligned them to the experimental TTL pulses. The results do not change either qualitatively or quantitatively.

SI Section – Automatic detection of switches after Aleshin et al 2019 (SI Figure 11)

To compare our manually marked switches to those detected automatically, we adapted the Cumulative Smooth Pursuit algorithm designed by Aleshin et al (2019) (*69*). From visual inspection and comparison, manual marking was more robust than the automatic detection, which was contaminated with multiple false positives and negatives. In light of this performance, we continued to use data aligned to the manual marks.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. L. A. Necker, LXI.*Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid*. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. **1**, 329–337 (1832).

2. E. G. Boring, A new ambiguous figure. *Am. J. Psychol.* **42**, 444 (1930).

3. R. Blake, N. K. Logothetis, Visual competition. *Nat. Rev. Neurosci.* **3**, 13–21 (2002).

4. C. W. G. Clifford, Binocular rivalry. *Curr. Biol.* **19**, R1022–3 (2009).

5. R. Blake, A Primer on Binocular Rivalry, Including Current Controversies. *Brain and Mind*. **2**, 5–38 (2001).

6. C. Moutard, S. Dehaene, R. Malach, Spontaneous Fluctuations and Non-linear Ignitions: Two Dynamic Faces of Cortical Recurrent Loops. *Neuron*. **88**, 194–206 (2015).

7. R. Blake, A neural theory of binocular rivalry. *Psychol. Rev.* **96**, 145–167 (1989).

8. D. A. Leopold, N. K. Logothetis, Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*. **379**, 549–553 (1996).

9. N. K. Logothetis, D. A. Leopold, D. L. Sheinberg, What is rivalling during binocular rivalry? *Nature*. **380**, 621–624 (1996).

10. P. Sterzer, A. Kleinschmidt, G. Rees, The neural bases of multistable perception. *Trends Cogn. Sci. (Regul. Ed.)*. **13**, 310–318 (2009).

11. V. A. Weilnhammer, K. Ludwig, G. Hesselmann, P. Sterzer, Frontoparietal cortex mediates perceptual transitions in bistable perception. *J. Neurosci.* **33**, 16009–16015 (2013).

12. E. D. Lumer, K. J. Friston, G. Rees, Neural correlates of perceptual rivalry in the human brain. *Science*. **280**, 1930–1934 (1998).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

13. T. Knapen, J. Brascamp, J. Pearson, R. van Ee, R. Blake, The role of frontal and parietal brain areas in bistable perception. *J. Neurosci.* **31**, 10293–10301 (2011).

14. F. Tong, M. Meng, R. Blake, Neural bases of binocular rivalry. *Trends Cogn. Sci. (Regul. Ed.)*. **10**, 502–511 (2006).

15. T. A. de Graaf, M. C. de Jong, R. Goebel, R. van Ee, A. T. Sack, On the functional relevance of frontal cortex for passive and voluntarily controlled bistable vision. *Cereb. Cortex*. **21**, 2322–2331 (2011).

16. J. F. Hipp, A. K. Engel, M. Siegel, Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron*. **69**, 387–396 (2011).

17. B. Odegaard, R. T. Knight, H. Lau, Should a few null findings falsify prefrontal theories of conscious perception? *J. Neurosci.* **37**, 9593–9602 (2017).

18. M. Boly *et al.*, Are the neural correlates of consciousness in the front or in the back of the cerebral cortex? clinical and neuroimaging evidence. *J. Neurosci.* **37**, 9603–9613 (2017).

19. S. Frässle, J. Sommer, A. Jansen, M. Naber, W. Einhäuser, Binocular rivalry: frontal activity relates to introspection and action but not to perception. *J. Neurosci.* **34**, 1738–1747 (2014).

20. M. C. de Jong *et al.*, Intracranial recordings of occipital cortex responses to illusory visual events. *J. Neurosci.* **36**, 6297–6311 (2016).

21. C. Koch, M. Massimini, M. Boly, G. Tononi, Neural correlates of consciousness: progress and problems. *Nat. Rev. Neurosci.* **17**, 307–321 (2016).

22. J. Brascamp, R. Blake, T. Knapen, Negligible fronto-parietal BOLD activity accompanying unreportable switches in bistable perception. *Nat. Neurosci.* **18**, 1672–1678 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

23. T. I. Panagiotaropoulos, G. Deco, V. Kapoor, N. K. Logothetis, Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. *Neuron*. **74**, 924–935 (2012).

24. H. Gelbard-Sagiv, L. Mudrik, M. R. Hill, C. Koch, I. Fried, Human single neuron activity precedes emergence of conscious perception. *Nat. Commun.* **9**, 2057 (2018).

25. N. K. Logothetis, J. D. Schall, Binocular motion rivalry in macaque monkeys: eye dominance and tracking eye movements. *Vision Res.* **30**, 1409–1419 (1990).

26. N. Tsuchiya, M. Wilke, S. Frässle, V. A. F. Lamme, No-Report Paradigms: Extracting the True Neural Correlates of Consciousness. *Trends Cogn. Sci. (Regul. Ed.)*. **19**, 757–770 (2015).

27. R. Fox, S. Todd, L. A. Bettinger, Optokinetic nystagmus as an objective indicator of binocular rivalry. *Vision Res.* **15**, 849–853 (1975).

28. C. Chandrasekaran, A. Trubanova, S. Stillittano, A. Caplier, A. A. Ghazanfar, The natural statistics of audiovisual speech. *PLoS Comput. Biol.* **5**, e1000436 (2009).

29. M. A. A. van der Meer, A. D. Redish, Low and High Gamma Oscillations in Rat Ventral Striatum have Distinct Relationships to Behavior, Reward, and Spiking Activity on a Learned Spatial Decision Task. *Front Integr Neurosci*. **3**, 9 (2009).

30. G. Schwarz, Estimating the Dimension of a Model. *Ann. Statist.* **6**, 461–464 (1978).

31. V. A. F. Lamme, Towards a true neural stance on consciousness. *Trends Cogn. Sci. (Regul. Ed.)*. **10**, 494–501 (2006).

32. K. Sandberg, S. Frässle, M. Pitts, Future directions for identifying the neural correlates of consciousness. *Nat. Rev. Neurosci.* (2016), doi:10.1038/nrn.2016.104.

33. M. Boly *et al.*, Consciousness in humans and non-human animals: recent advances and future directions. *Front. Psychol.* **4**, 625 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

34. B. van Vugt *et al.*, The threshold for conscious report: Signal loss and response bias in visual and frontal cortex. *Science*. **360**, 537–542 (2018).

35. C. Libedinsky, M. Livingstone, Role of prefrontal cortex in conscious visual perception. *J. Neurosci.* **31**, 64–69 (2011).

36. C. H. M. Brunia, G. J. M. van Boxtel, K. B. E. Böcker, *Negative Slow Waves as Indices of Anticipation: The Bereitschaftspotential, the Contingent Negative Variation, and the Stimulus-Preceding Negativity* (Oxford University Press, 2011), *Oxford Handbooks Online*.

37. A. Schurger, Specific Relationship between the Shape of the Readiness Potential, Subjective Decision Time, and Waiting Time Predicted by an Accumulator Model with Temporally Autocorrelated Input Noise. *Eneuro*. **5** (2018), doi:10.1523/ENEURO.0302-17.2018.

38. H. H. Kornhuber, L. Deecke, [changes in the brain potential in voluntary movements and passive movements in man: readiness potential and reafferent potentials]. *Pflugers Arch. Gesamte Physiol. Menschen Tiere*. **284**, 1–17 (1965).

39. B. Libet, C. A. Gleason, E. W. Wright, D. K. Pearl, 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 (1983).

40. C. Tzagarakis, N. F. Ince, A. C. Leuthold, G. Pellizzer, Beta-band activity during motor planning reflects response uncertainty. *J. Neurosci.* **30**, 11270–11277 (2010).

41. W. J. Ray, H. W. Cole, EEG alpha activity reflects attentional demands, and beta activity reflects emotional and cognitive processes. *Science*. **228**, 750–752 (1985).

42. J. Alayrangues, F. Torrecillos, A. Jahani, N. Malfait, Error-related modulations of the sensorimotor post-movement and foreperiod beta-band activities arise from distinct

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

neural substrates and do not reflect efferent signal processing. *Neuroimage*. **184**, 10–24 (2019).

43. V. Piai, A. Roelofs, J. Rommers, K. Dahlslätt, E. Maris, Withholding planned speech is reflected in synchronized beta-band oscillations. *Front. Hum. Neurosci.* **9**, 549 (2015).
44. F. David, E. Courtiol, N. Buonviso, N. Fourcaud-Trocmé, Competing mechanisms of gamma and beta oscillations in the olfactory bulb based on multimodal inhibition of mitral cells over a respiratory cycle. *Eneuro*. **2** (2015), doi:10.1523/ENEURO.0018-15.2015.
45. A. K. Engel, P. Fries, Beta-band oscillations--signalling the status quo? *Curr. Opin. Neurobiol.* **20**, 156–165 (2010).
46. T. I. Panagiotaropoulos, V. Kapoor, N. K. Logothetis, Desynchronization and rebound of beta oscillations during conscious and unconscious local neuronal processing in the macaque lateral prefrontal cortex. *Front. Psychol.* **4**, 603 (2013).
47. B. Spitzer, S. Haegens, Beyond the status quo: A role for beta oscillations in endogenous content (re)activation. *Eneuro*. **4** (2017), doi:10.1523/ENEURO.0170-17.2017.
48. E. K. Miller, M. Lundqvist, A. M. Bastos, Working Memory 2.0. *Neuron*. **100**, 463–475 (2018).
49. Y. Zerlaut, A. Destexhe, Enhanced Responsiveness and Low-Level Awareness in Stochastic Network States. *Neuron*. **94**, 1002–1009 (2017).
50. O. Jensen, M. Bonnefond, Prefrontal α- and β-band oscillations are involved in rule selection. *Trends Cogn. Sci. (Regul. Ed.)*. **17**, 10–12 (2013).
51. H. Lau, D. Rosenthal, Empirical support for higher-order theories of conscious awareness. *Trends Cogn. Sci. (Regul. Ed.)*. **15**, 365–373 (2011).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

52. S. Dehaene, L. Naccache, Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*. **79**, 1–37 (2001).

53. In the theatre of consciousness. Global Workspace Theory, a rigor...: Ingenta Connect, (available at https://www.ingentaconnect.com/content/imp/jcs/1997/00000004/00000004/776).

54. A. Del Cul, S. Dehaene, P. Reyes, E. Bravo, A. Slachevsky, Causal role of prefrontal cortex in the threshold for access to consciousness. *Brain*. **132**, 2531–2540 (2009).

55. Program Planner, (available at http://www.abstractsonline.com/pp8/index.html#!/4649/presentation/11577).

56. H. Xu *et al.*, Rivalry-Like Neural Activity in Primary Visual Cortex in Anesthetized Monkeys. *J. Neurosci.* **36**, 3231–3242 (2016).

57. L. Mazzucato, A. Fontanini, G. La Camera, Dynamics of multistable states during ongoing and evoked cortical activity. *J. Neurosci.* **35**, 8214–8231 (2015).

58. M. Tsodyks, T. Kenet, A. Grinvald, A. Arieli, Linking spontaneous activity of single cortical neurons and the underlying functional architecture. *Science*. **286**, 1943–1946 (1999).

59. M. J. McGinley *et al.*, Waking state: rapid variations modulate neural and behavioral responses. *Neuron*. **87**, 1143–1161 (2015).

60. E. M. Maynard, C. T. Nordhausen, R. A. Normann, The Utah intracortical Electrode Array: a recording structure for potential brain-computer interfaces. *Electroencephalogr. Clin. Neurophysiol.* **102**, 228–239 (1997).

61. C. R. Hussar, T. Pasternak, Flexibility of sensory representations in prefrontal cortex depends on cell type. *Neuron*. **64**, 730–743 (2009).

62. S. Safavi *et al.*, Nonmonotonic spatial structure of interneuronal correlations in prefrontal microcircuits. *Proc. Natl. Acad. Sci. USA*. **115**, E3539–E3548 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

63. R. Q. Quiroga, Z. Nadasdy, Y. Ben-Shaul, Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* **16**, 1661–1687 (2004).

64. S. N. Kadir, D. F. M. Goodman, K. D. Harris, High-dimensional cluster analysis with the masked EM algorithm. *Neural Comput.* **26**, 2379–2394 (2014).

65. L. Hazan, M. Zugaro, G. Buzsáki, Klusters, NeuroScope, NDManager: a free software suite for neurophysiological data processing and visualization. *J. Neurosci. Methods*. **155**, 207–216 (2006).

66. N. K. Logothetis *et al.*, Hippocampal-cortical interaction during periods of subcortical silence. *Nature*. **491**, 547–553 (2012).

67. H. Bokil, P. Andrews, J. E. Kulkarni, S. Mehta, P. P. Mitra, Chronux: a platform for analyzing neural signals. *J. Neurosci. Methods*. **192**, 146–151 (2010).

68. J. L. Hodges, E. L. Lehmann, Estimates of location based on rank tests. *Ann. Math. Statist.* **34**, 598–611 (1963).

69. S. Aleshin, G. Ziman, I. Kovács, J. Braun, Perceptual reversals in binocular rivalry: Improved detection from OKN. *J. Vis.* **19**, 5 (2019).

Acknowledgements

We thank Dr. Yusuke Murayama and the other technical and animal care staff for excellent technical assistance, Prof. Nicho Hatsopoulos for help with the implantations of the Utah arrays, Prof. Stanislas Dehaene for his inputs and insights, and finally Mr. Akshat Jain for his assistance in writing code for the resting state analysis and preparing publishable quality figures.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Funding

Funding was provided by the Max Planck Society for this research project.

Author contributions

Conceptualisation: AD, VK, TIP (lead), NKL; Data curation: AD (lead), VK and JW; Formal analysis: AD (lead), VK, JW, LAF; Funding acquisition: NKL; Investigation: AD (equal), VK (equal), TIP (supporting); Methodology: AD (equal), VK (equal), JW & SS (supporting), TIP (equal); Project administration: TIP; Resources: JW, NKL (lead); Software: AD (lead), VK, JW, LAF & SS (supporting); Supervision: TIP; Visualisation: AD (lead), TIP (supporting); Writing - original draft: AD, TIP (lead); Writing - review & editing: AD, VK, LAF, SS, TIP (lead), NKL.

Competing interests

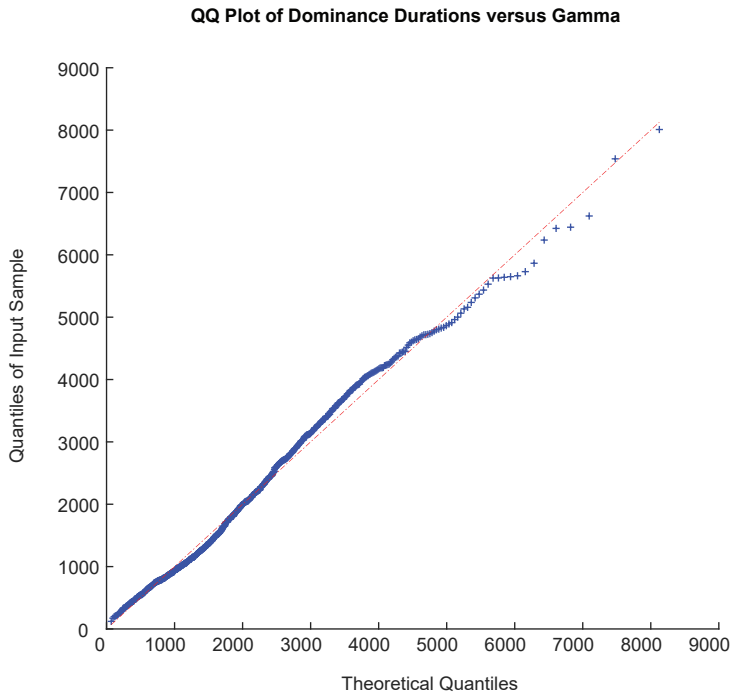
The authors declare no competing interests.

Data and materials availability

Raw data and analysis codes are available upon request.

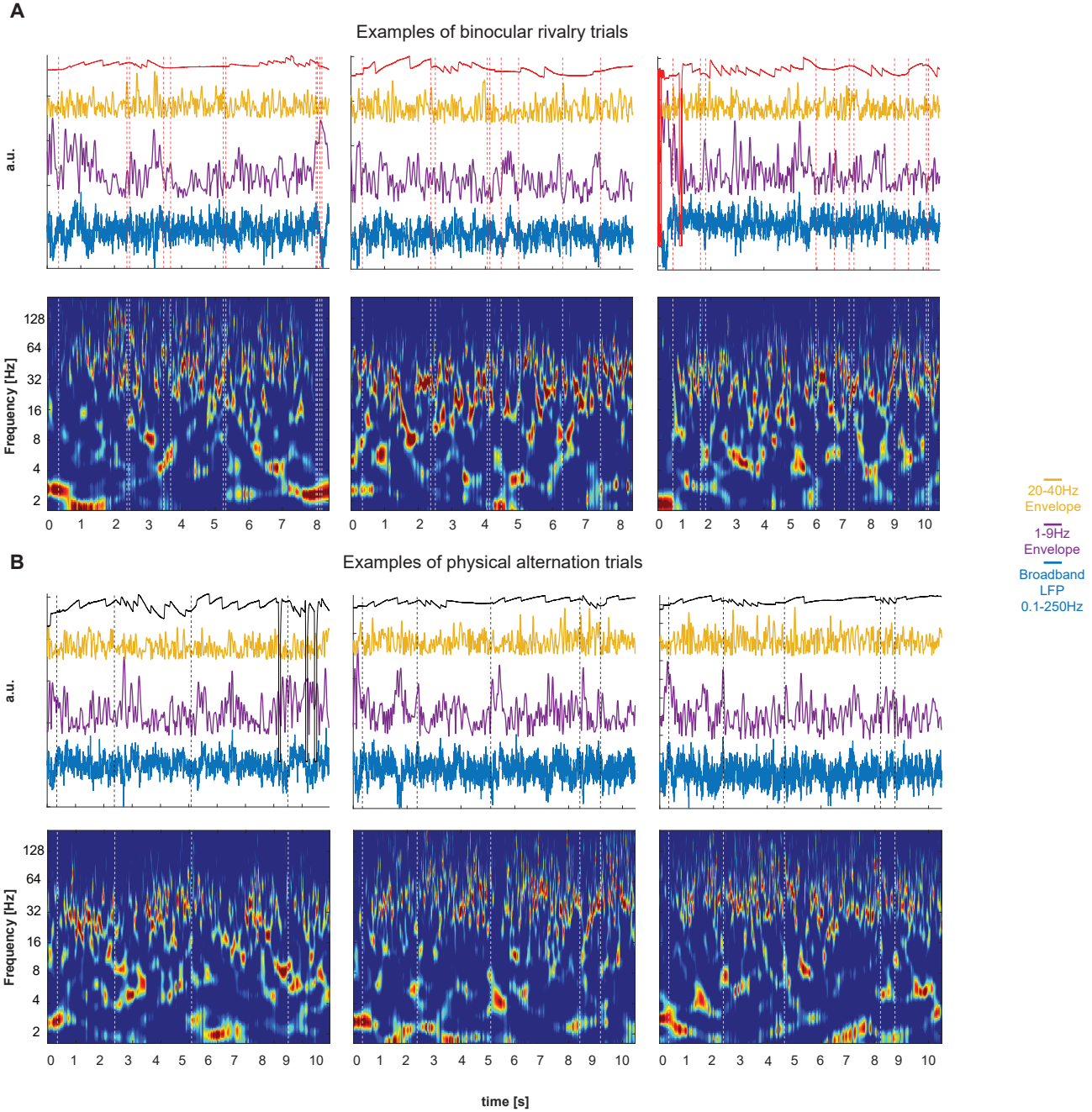
Supplementary materials

Methods, Supporting Information, SI Figures 1-9, References #61-69



SI Figure 1 | Quantile-Quantile diagnostics

We qualitatively confirmed that the gamma distribution was the best fit for the dominance durations using a Quantile-Quantile plot. The red dashed line shows the theoretical distribution whereas the blue crosses show the fidelity of the data to the theoretical distribution. Furthermore, a comparison of the Bayesian Information Criterion (BIC) yielded the lowest value for a gamma distribution (1.92x10^4) as compared to an exponential (1.96x10^4) or a logistic function (1.962x10^4), pointing to the gamma distribution as the best-fit model.

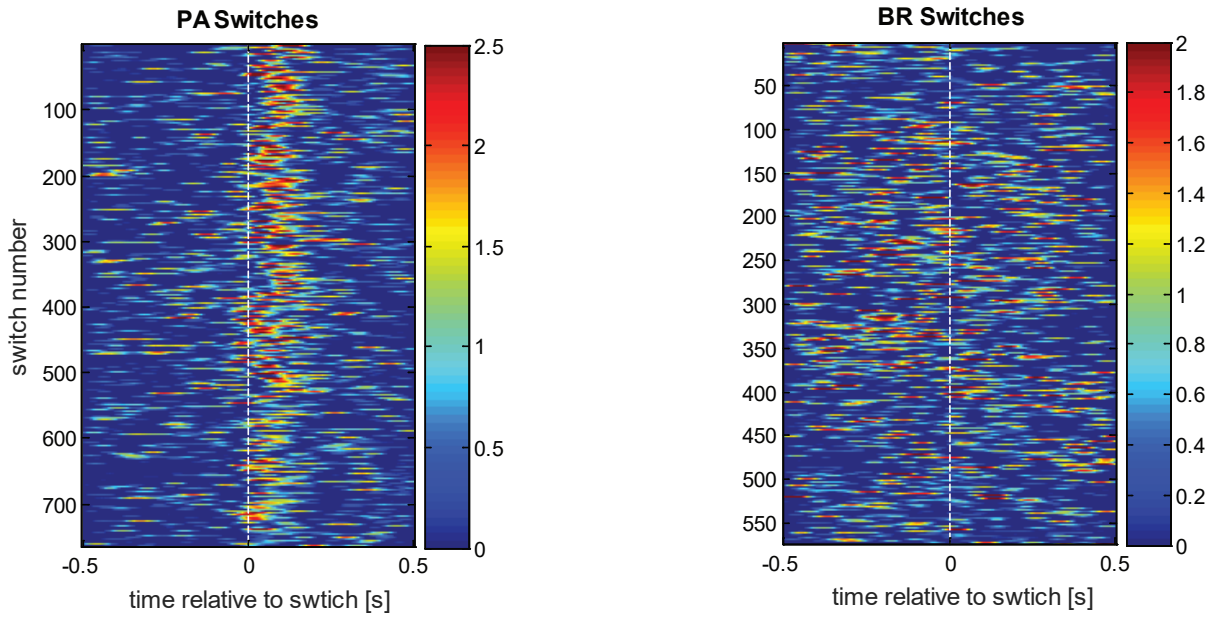


SI Figure 2 | Low Frequency and Beta Instataneous Amplitudes corresponding to a percept switch

A. The low frequency activity and the concomitant beta activity averaged across all the 96 channels for each trial in BR shows a large amplitude deflection and putative climbing activity just before a spontaneous percept switch (marked by the change in the polarity of the Optokinetic Nystagmus(OKN)), or during a piecemeal period before an upcoming transition

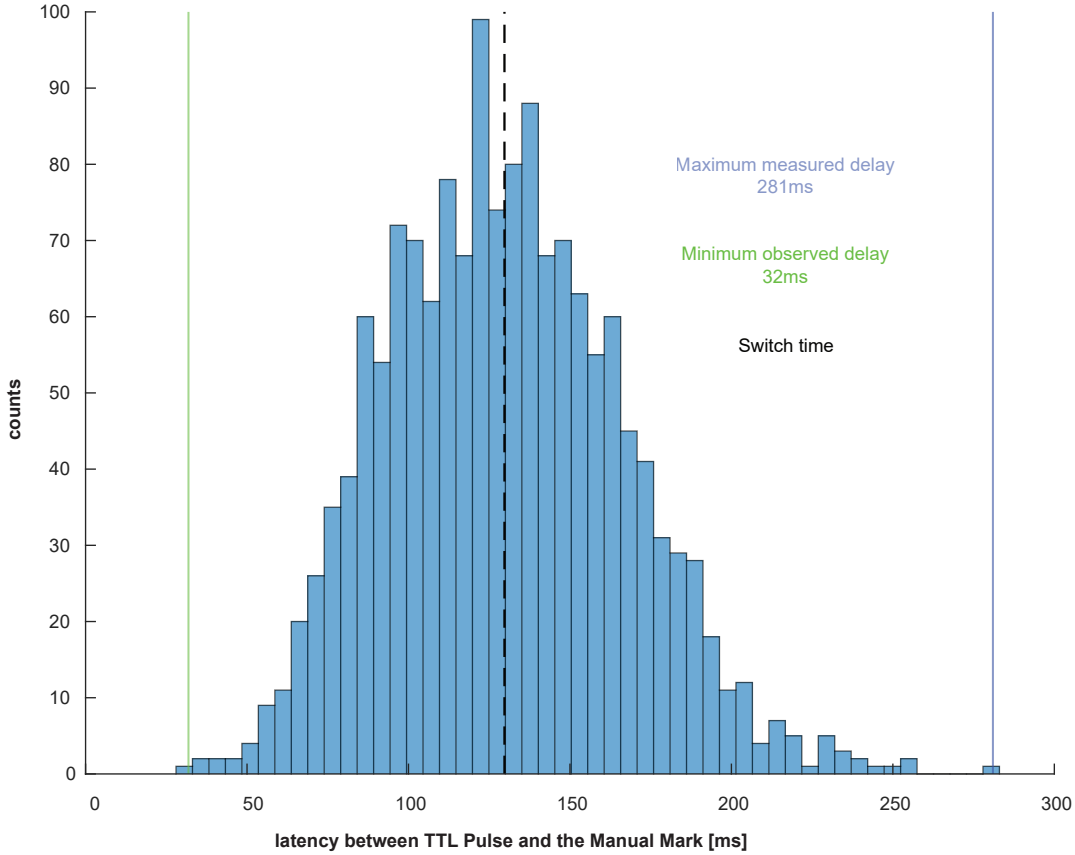
B. The low frequency activity and the concomitant beta activity in Physical Alternation(PA) trials shows the same kind of large amplitude deflection but this time at the point of or after the stimulus has been switched (manually marked transitions succeed TTL pulses). The amplitude deflection is even greater in magnitude because it is a visually evoked potential.

Random deviations from baseline are also observed; however, they are far fewer in number and lesser in magnitude



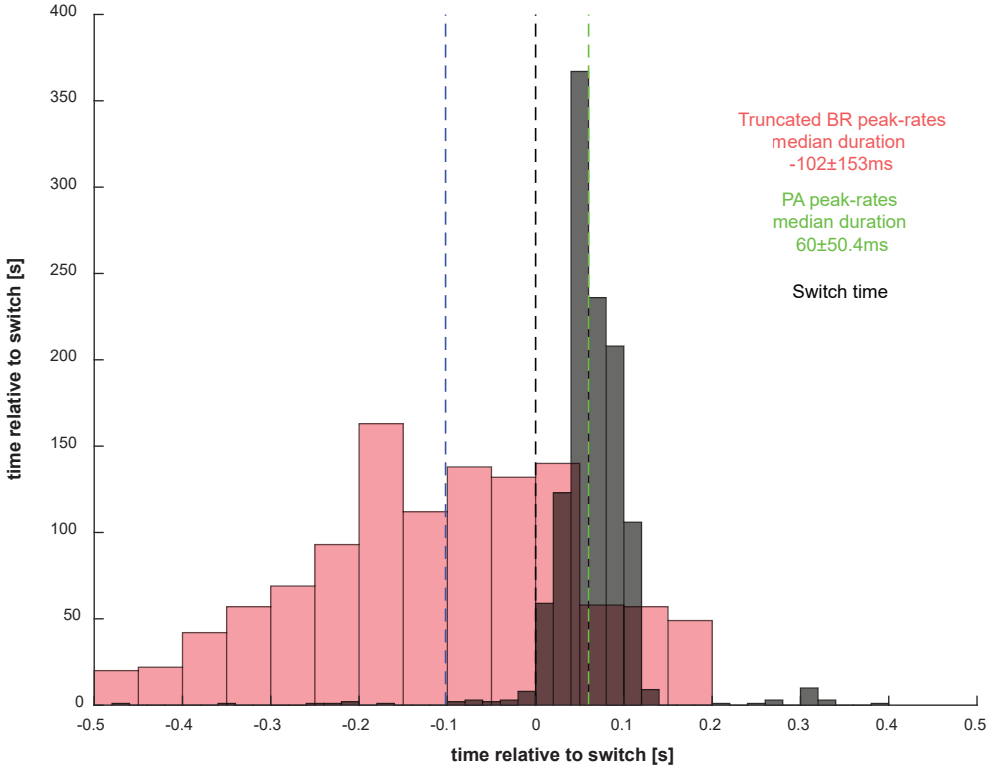
SI Figure 3 | Qualitative analysis of the temporal jitter in the 1-9Hz instantaneous amplitude

This figure depicts the 1-9Hz instantaneous amplitude extracted around clean switches (i.e. a minimum dominance of 500ms before and after a transition) in both the physical alternation (left) and binocular rivalry (right) conditions. While after the transition in PA, a strong and consistent visually-evoked potential is observed, the activity around a spontaneous switch is rather diffuse, yet shows a preponderance of low-frequency activity rising and concentrated in the pre-switch period.



SI Figure 4 | Comparison of the difference between the TTL pulse and the manually marked change in Physical Alternation

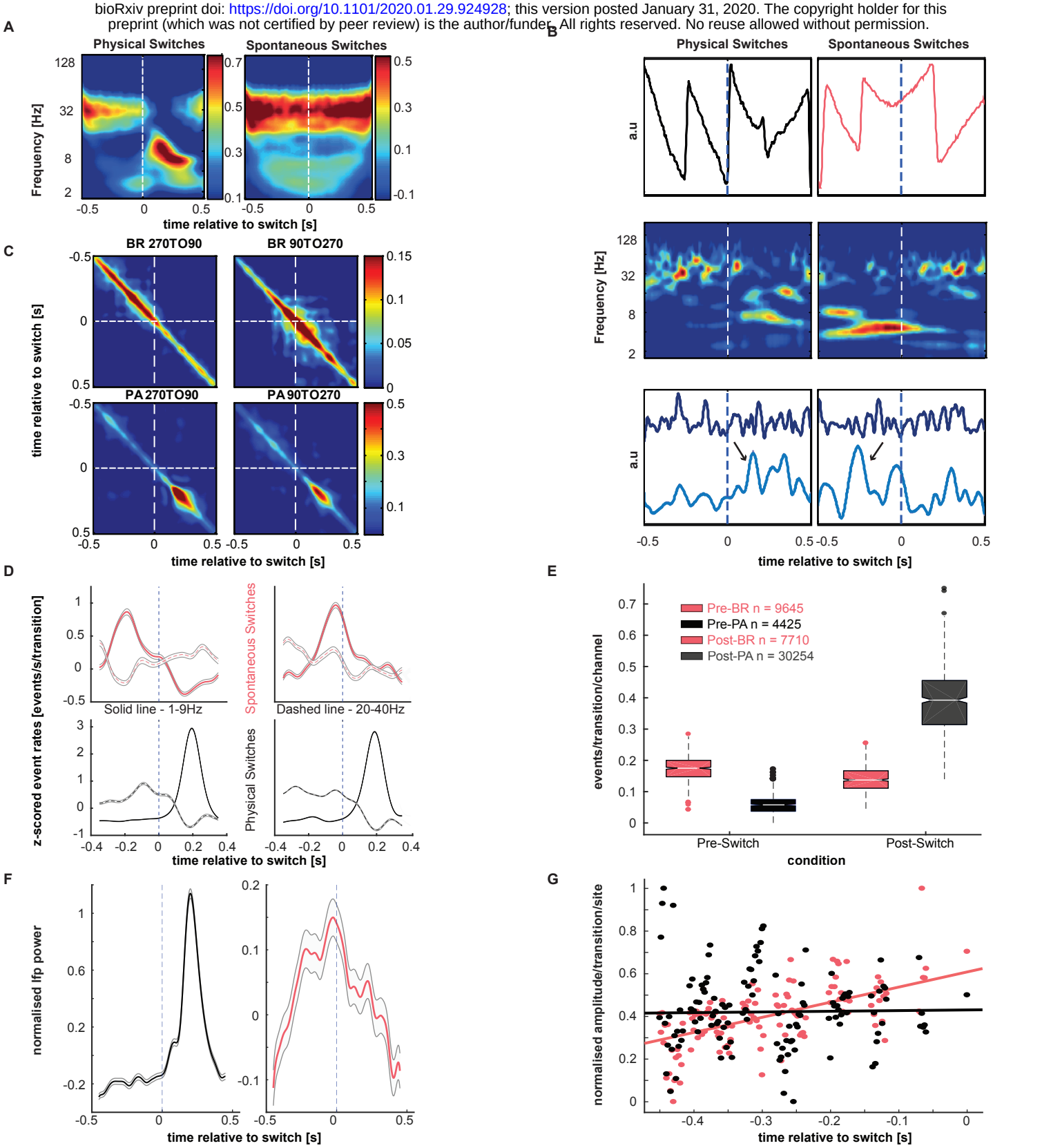
We computed the latency between the onset of the subsequent stimulus and the succeeding change in the polarity of the induced OKN for all PA trials. We found an average latency of 129.4±36.5 ms (mean±SD), which indicates that the change in the eye-movement is induced within a very short interval. The minimum latency observed was 32ms whereas the maximum observed latency was 281ms.



SI Figure 5 | Truncated low-frequency peak rate comparison

Because events towards the end of the post-switch window in BR could signal an upcoming transition, we discarded these events that occurr after 150ms (timing of the end of the VEP in PA). The difference between the median timing of the peak rate in BR and the VEP in PA was further enhanced.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



SI Figure 6 | Statistics for pulse aligned switches

A. Grand average time-frequency analysis of all physical (left) and spontaneous (right) perceptual transitions. Spectrograms are aligned (t=0) to the experimental TTL pulse for PA and to the OKN change for BR for periods of stable perceptual dominance before and after the switch.

B. Upper panel: OKN traces around a single physical (black) and spontaneous (red) transition. Middle panel: channel-averaged normalised spectrograms around the single transition event for the two conditions. Lower panel: Normalised instantaneous amplitudes of the two modulated frequency bands (i.e 1-9Hz and 20-40Hz) identified from the spectrograms. Low-frequency bursts occur after the physical switch but before the spontaneous transition.

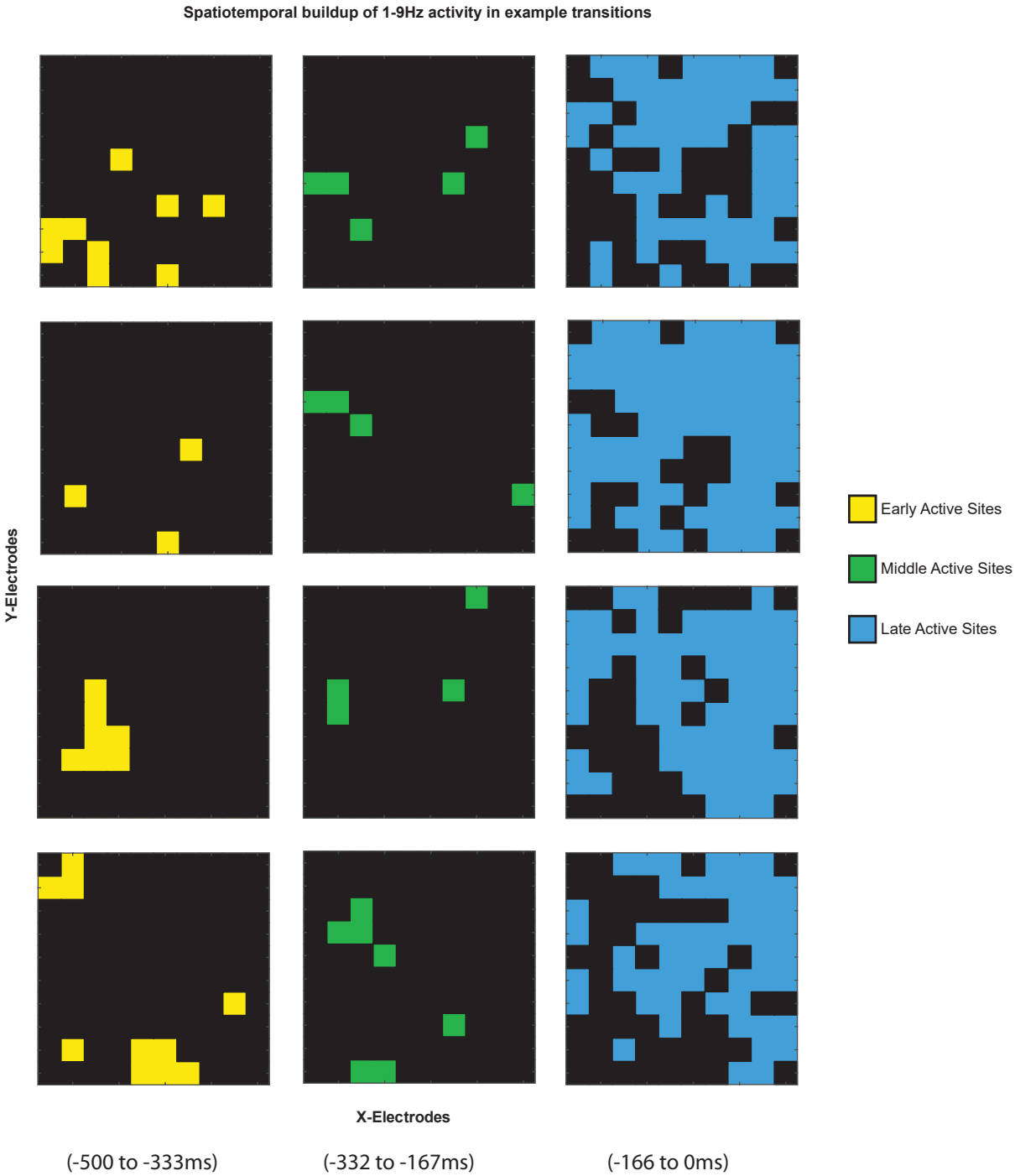
C. Differences in the onset of the low-frequency activity across physical and spontaneous transitions are reflected in the temporal auto-covariance of the low-frequency envelopes across the array recorded simultaneously for every transition type (270 to 90, left and 90 to 270, right). Most of the similarity is observed after a transition in PA (bottom panel) but before a transition in BR (upper panel).

D. Normalised (z-score) event rate in time (events/s/transition) during BR (red lines) and PA trials (white lines) for low-frequency (solid lines) and beta activity (dashed lines).

E. Burst rate per transition per channel for periods before and after transitions in PA (red) and BR (black). In order from left to right: Pre-BR (dark pink), Pre-PA (black), Post-BR (light pink), Post-PA (grey). The whiskers of the box-plots show the dispersion of the data. The white line depicts the median. Outliers are represented by appropriately coloured dots. A significantly higher number of low-frequency bursts occur before spontaneous, but after physical switches. Burst rate before a physical switch is very low, suggesting noise levels. This baseline burst-rate needs to be ramped up for a switch to occur. PA: (0.39±0.004, n = 30254, post-transition, vs. 0.06±0.001, pre-transition, n = 4425; p<10-188 median±SEM) BR: (0.18±0.002, n = 9645, pre-transition vs 0.14±0.002, n = 7710, post-transition, p<10-43 median±SEM).

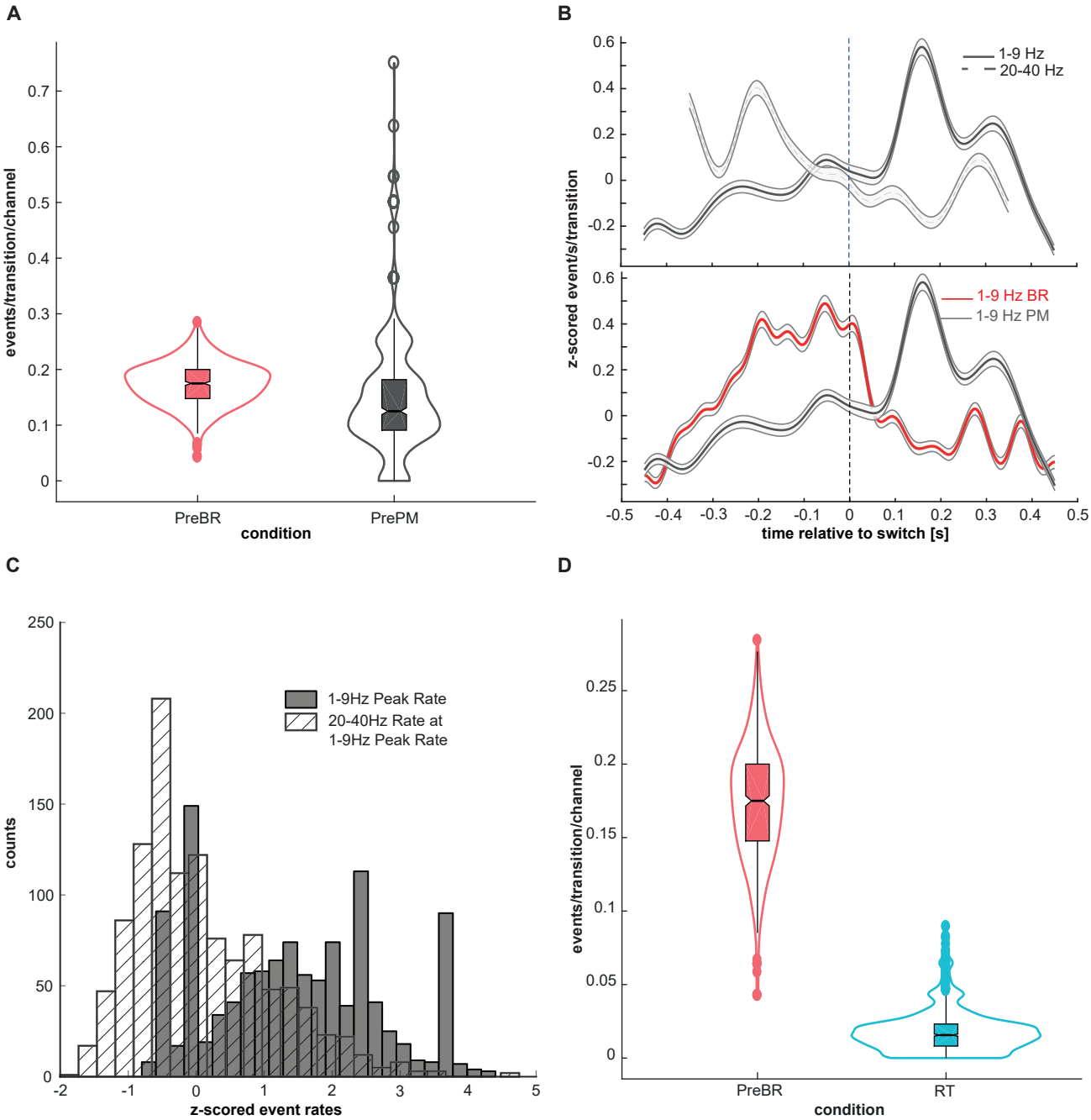
F. Low-frequency instantaneous amplitude shows a slow climbing activity before a perceptual transition (right) but not before a physical transition (left). Curves reflect an average across transitions of the channel-averaged activity for each transition

G. Average build-up of low-frequency activity in time is shown by fitting a line to the pooled and averaged amplitude at every time-point (red – BR, black – PA) across all measured neuronal sites. While before a spontaneous transition, the recorded PFC area ramps up its low-frequency activity in time, before a physical transition, it remains flat.



SI Figure 7 | Spatial buildup of low-frequency activity.

This figure shows the spatial buildup of 1-9Hz activity in 3 temporal windows viz. in an early window ([-500 to -333ms]), a middle window ([-334 to -166ms]) and a late window ([-167 to 0ms]) preceding four typical spontaneous switches. Progressively, more sites are activated approaching a switch.



SI Figure 8 | Controlling for failed and non-occurrence of switches

a. The low frequency burst rate per transition per channel before a spontaneous transition (Pre-BR, n= 9667, 0.17±0.002) when compared to the period before transition to the piecemeal percept (Pre-PM, n= 2486, 0.147±0.004) is higher, but not significantly (p=0.08).

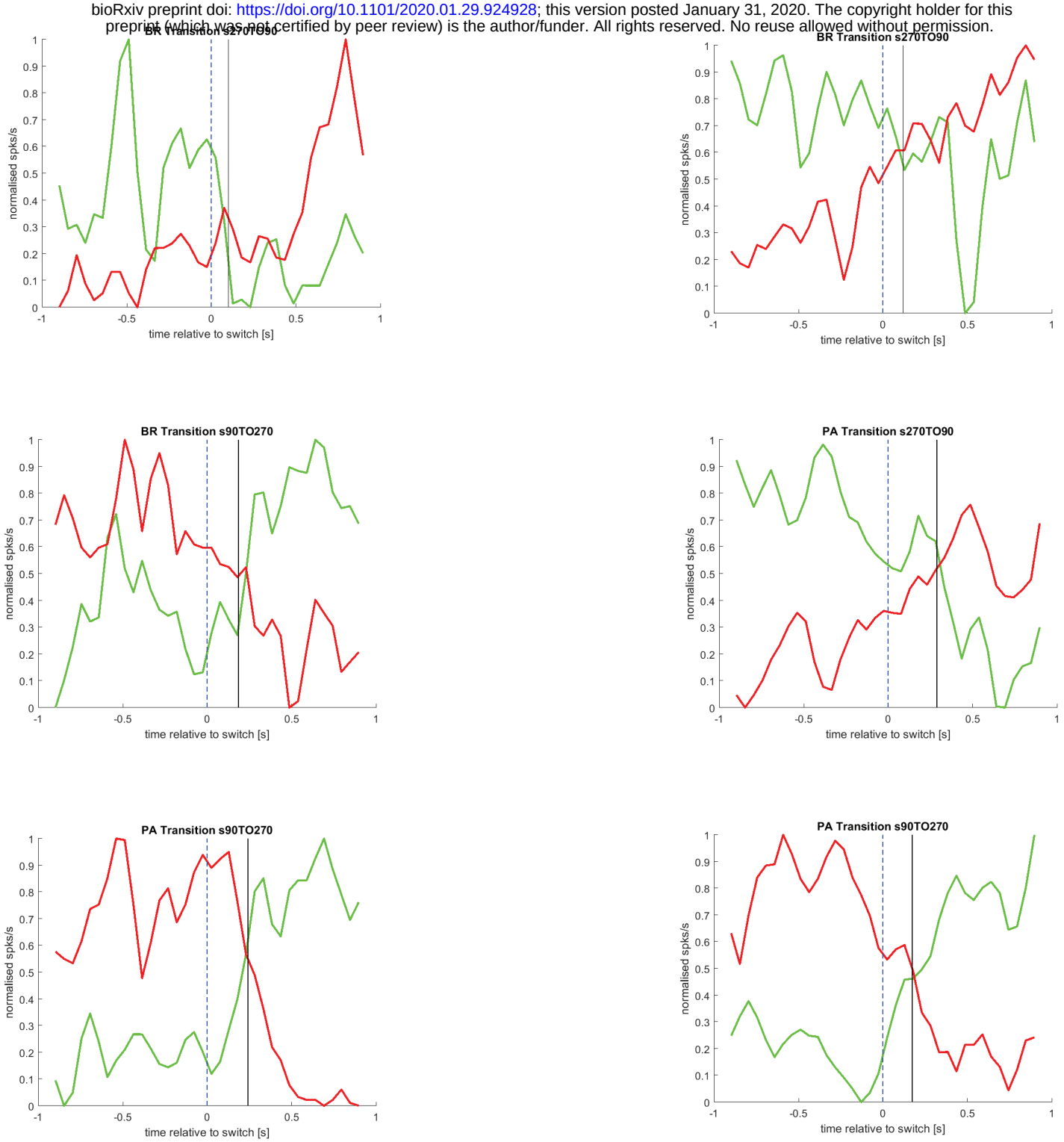
b. Top row panel: z-scored event rates in time (events/transition/s) before and after the transition to piecemeal percepts, the solid line being the low-frequency component and the dashed line being the beta-band activity. While before the transition to piecemeal, the beta dominates, signalling the active percept, after the transition the low-frequency inhibits the beta thus signalling an upcoming percept.

Bottom row panels: z-scored event rates in time (events/transition) for the low frequency activity for BR(red) and PM(grey). The low frequency activity burst rate peaks before transition to another clear dominance during BR (red) whereas it peaks after the transition to a piecemeal percept (grey).

c. The distribution of the peak low-frequency rates vis a vis the rate of the beta activity at the timing of the low-frequency peak reveals no significant antagonism before a transition to a piecemeal (r = -0.0073, p=0.805) as compared to before clear spontaneous transitions, where a significant decoupling is observed (r = -0.08, p=0.0071).

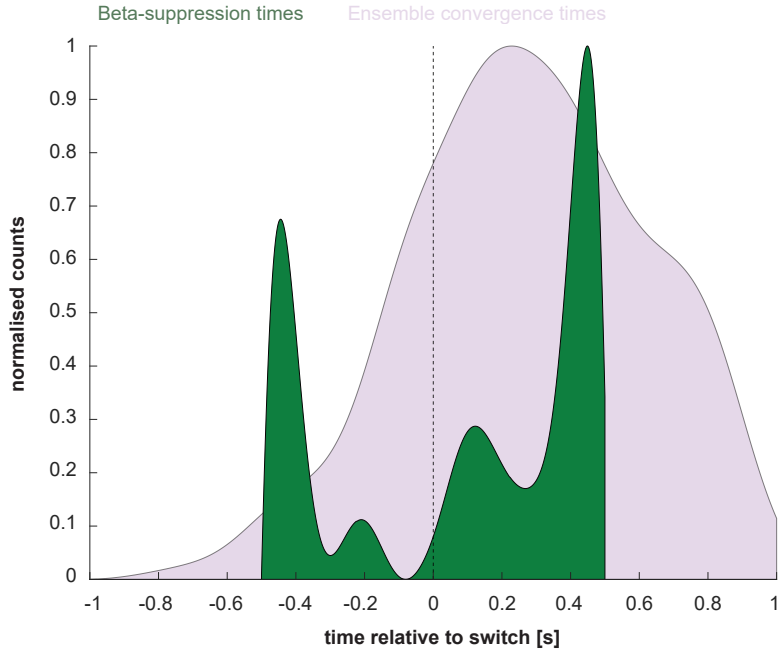
d. Low-frequency event rate per transition per channel for the two conditions, spontaneous switch and randomly triggered periods. The burst rate before a spontaneous switch (pre-BR, 0.17±0.002, n = 9667) is significantly higher than during randomly triggered epochs (RT, 0.018±0.00007, n = 55026, 100 iterations, 550.26 bursts per iteration).

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



SI Figure 9 | Estimation of change in encoding in selective ensembles

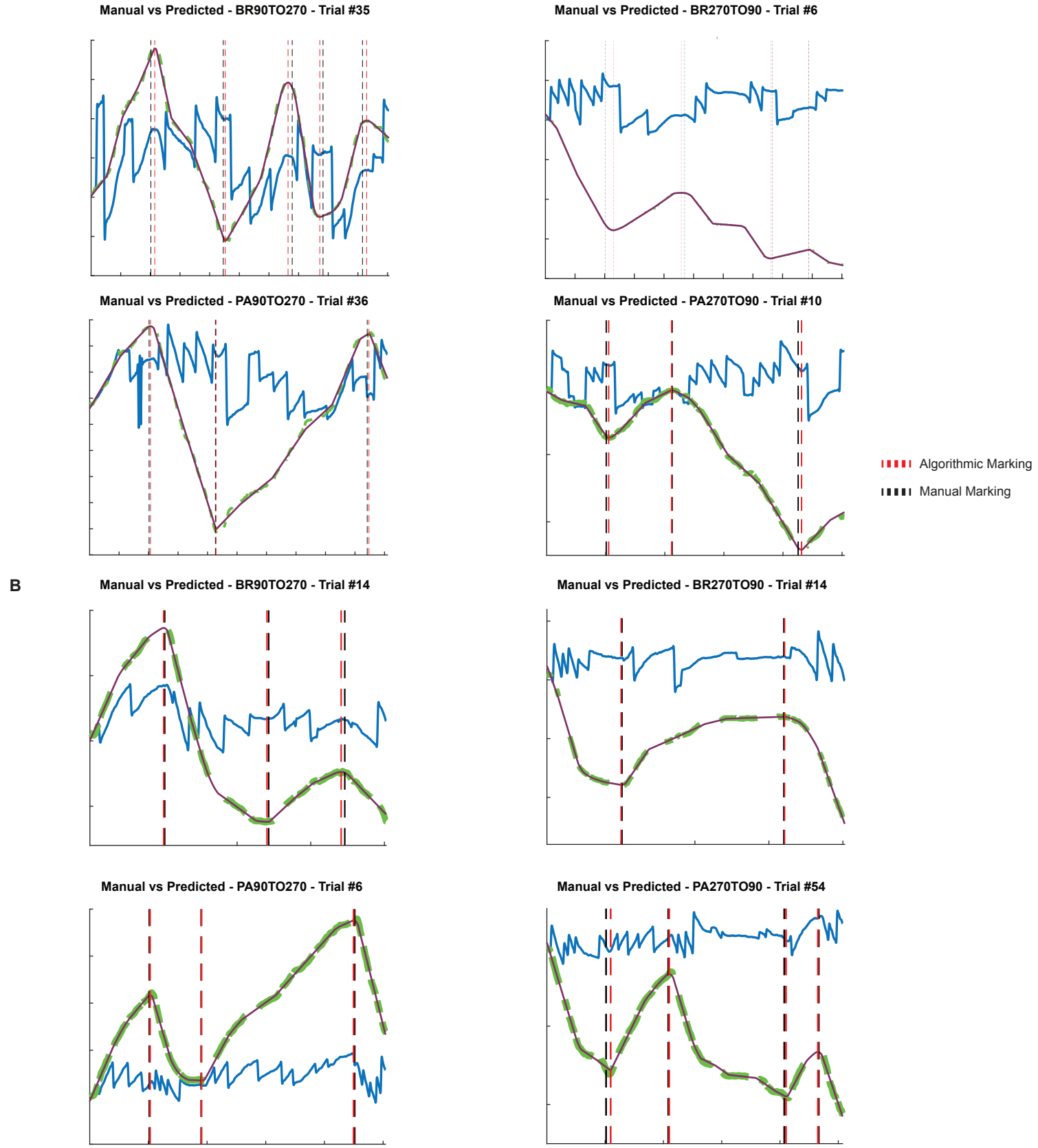
Cross-over points of the normalised population vectors (summed spike counts per bin) describe the point of change in the encoding of the dominant percept by the selective ensembles of neurons. These points are computed by first estimating the best-fit trend of the population vectors LOWESS smoothing, and then estimating the intersection using interpolation. If multiple intersections are estimated, then that intersection point is chosen such that for 200ms before and after it, there are no other crossings, i.e. stable divergence is observed. Three examples of BR and PA transitions each are shown above (green = downward selective population; red = upward selective population). These cross-over points are then collected across all transitions.



SI Figure 10 | Comparison of Beta-Suppression and Ensemble convergence times

The green probability density function shows the distribution of times of beta-activity while the purple function shows the ensemble convergence times of the two competing populations around a switch. The ongoing beta is suppressed even before the a spontaneous transition happens inferred by the change in the polarity of the OKN.

bioRxiv preprint doi: https://doi.org/10.1101/2020.01.29.924928; this version posted January 31, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



SI Figure 11 | Comparison of Automatically detected and manually marked percept switches from the OKN readout for BR and PA trials in 2 example sessions

The percept switches algorithmically detected (red) using extremum points in the the computed cumulative smooth pursuit were compared to the manually marked (black) ones using the inflection point of the polarity reversal of the OKN readout. The results were inconsistent across across sessions and conditions, only the manual markings were used to align all the data.

A. The automatic marking method in this dataset inaccurately marks switches determined by falling within a threshold window of the manually marked switches (black), along with a significant number of outright false positives.
B. The automatically marked switches in this dataset are equally accurate or slightly worse in accuracy than manual marked switches.