Pre-stimulus feedback connectivity biases the content of visual experiences

Elie Rassi[1*], Andreas Wutz[1], Nadia Müller[2,3], & Nathan Weisz[1,3]

[1] Centre for Cognitive Neuroscience, University of Salzburg, Austria
[2] Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy
[3] Department of Neurosurgery, Friedrich-Alexander- Universitätsklinikum, Erlangen-Nürnberg, Germany
* corresponding author

Elie Rassi: elie.elrassi@sbg.ac.at
Andreas Wutz: andreas.wutz@sbg.ac.at
Nadia Müller: nadia.mueller@gmail.com
Nathan Weisz: nathan.weisz@sbg.ac.at

Address for correspondence:

Elie Rassi
University of Salzburg
Centre for Cognitive Neuroscience
Hellbrunnerstr. 34
A-5020 SALZBURG
Austria/Europe
Tel. 0043 662 8044 5148 or 5136
Fax 0043 662 8044 5126
Cell 0043 677 62116323
Email: elie.elrassi@sbg.ac.at

Abstract

Ongoing, spontaneous fluctuations in neural excitability and in network-wide activity patterns before stimulus onset have been proposed to underlie variability in near-threshold stimulus detection paradigms, i.e. whether an object is perceived or not. Here, we investigated the impact of pre-stimulus neural fluctuations on the content of perception, i.e. whether one or another object is perceived. We recorded neural activity with magnetoencephalography before and while participants briefly viewed a bi-stable image, the Rubin face/vase illusion, and required them to report their perceived interpretation on each trial. Using multi-variate pattern analysis, we showed robust decoding of the perceptual report during the post-stimulus period. Applying source localization to the classifier weights suggested early recruitment of V1 and ~160 ms recruitment of category-sensitive FFA. These post-stimulus effects were driven by stronger oscillatory power in the gamma frequency band for face vs vase reports. In pre-stimulus intervals, we found no differences in oscillatory power between face vs. vase reports in V1 nor in FFA, indicating similar levels of neural excitability. Despite this, we found stronger connectivity between V1 and FFA prior to face reports for low-frequency oscillations. Specifically, the strength of pre-stimulus feedback connectivity (i.e. Granger causality) from FFA to V1 predicted not only the category of the upcoming percept, but also the strength of post-stimulus neural activity associated with the percept. Our work identifies pre-stimulus network states that shape future processing in category-sensitive brain regions and in this way determine the content of visual experiences.

Spontaneous, ongoing fluctuations in neural activity interact with perceptual and cognitive processes. They help to explain why repetitions of the same physical stimuli elicit different percepts and responses from trial to trial in both animals (Arieli et al., 1996) and humans (Boly et al., 2007). Both local excitability changes in task-relevant sensory regions (Super et al., 2003; Sadaghiani et al., 2009) and neural network connectivity patterns have been shown to underlie trial-by-trial fluctuations in perception (Weisz et al., 2014; Leske et al., 2015; Frey et al., 2016).

The central paradigm to study the impact of ongoing neural activity on perception typically involves near-threshold stimuli, in which pre-stimulus, neural fluctuations influence their perceptual fate, i.e. whether an object is seen ("Hit") or not ("Miss"; e.g. Ergenoglu et al., 2003; Weisz et al., 2014; Wutz et al., 2014; Leske et al., 2015; Frey et al., 2016; Wutz et al., 2018). Beyond mere stimulus detection, one of the visual system's essential functions is to identify and categorize objects and - by this – construct the content of visual experiences (Logothetis & Sheinberg, 1996; Miller et al., 2003; Wutz et al., 2018). Indeed, neural correlates of object recognition and categorization have been shown to rely on the information flow between occipital and inferior temporal cortical regions (Kreiman et al., 2000; Sigala & Logothetis, 2002; Freedman et al., 2003). Here, we focus on the impact of neural excitability and connectivity patterns before stimulus onset on the content of perceptual operations.

Bi-stable perception paradigms are uniquely suited to address this question (Blake & Logothetis, 2002). In such paradigms, the brain is conflicted between multiple possible interpretations of visual content. Typical examples include the Rubin's face/vase stimulus (Rubin, 1915), the Necker cube (Necker, 1832), and binocular rivalry (Porta, 1593; as cited in Wade, 1996). Recent evidence from fMRI studies has shown that rivalry between two competing percepts is resolved relatively early in the visual hierarchy (e.g. Zou et al., 2016; Leopold & Logothetis, 1996), such as in category-sensitive inferior temporal lobe regions (Tong et al., 1998; but see Sterzer & Kleinschmidt, 2007, for evidence showing an influence of frontal cortex; and see Brascamp et al., 2018, for a recent review). In particular for the Rubin's face/vase illusion, greater BOLD activity has been observed in the fusiform face area (FFA) when participants reported seeing a face rather than a vase (Hasson et al., 2001). Importantly, this BOLD increase in FFA was also observed prior to stimulus onset

(Hesselmann et al., 2008), possibly because the pre-stimulus brain state biased perception towards the “face interpretation”. Still, a more comprehensive, mechanistic account requires means to simultaneously measure neural activity in multiple cortical areas with high temporal resolution, in order to map out the cortical hubs and their inter-areal information flow during bi-stable perception. For example, enhanced BOLD activity in FFA could be a consequence of either increased feedforward activity from earlier visual regions, or of increased feedback activity to earlier visual regions.

In the current study, we used a similar Rubin face/vase paradigm as the aforementioned fMRI study (Hesselmann et al., 2008). Advancing on previous work, we thoroughly characterized neural activity and connectivity patterns with high temporal resolution prior to and during bi-stable perception by means of magnetoencephalography (MEG). We hypothesized that category-sensitive processing region (i.e. FFA) should exhibit differential pre-stimulus connectivity patterns preceding subsequent face vs. vase reports. Based on the “Windows to Consciousness” framework (Weisz et al., 2014; Ruhnau et al., 2014), fluctuating connectivity levels of sensory regions shape upcoming stimulus processing (i.e. whether a stimulus is perceived or not). We extended these predictions to visual object categorization and investigated whether categorical responses to the content of the Rubin stimulus were biased by local excitability, feedforward connectivity, or feedback connectivity between primary visual cortex and FFA.

Results

Twenty participants took part in the MEG experiment. We showed them the Rubin's face / vase stimulus briefly and asked them to report whether they had seen faces or a vase on each trial (see Methods for details). Vase and face reports were equally likely (Face mean: 50.1 %; SD: 12.47%; range 15.2% to 77.4%; t-test against chance (50%) $t(19) = 0.04$, $p = .97$). To ascertain that the reported perception was stochastic trial-by-trial, we analysed the sequences of reported percepts by binning the trials into a range of 0 to 10 repetitions. A binomial distribution accounted well for the binned data for both vase and face trials (goodness-of-fit $R^2 = 0.96$ for face, $R^2 = 0.98$ for vase) indicative of no systematic reporting of either percept. That is, during each trial a participant was equally likely to report a face or vase irrespective of the previous trial.

In a first step, we aimed to investigate to what extent category-specific information was contained in the recorded MEG data and whether source localization would yield the ROIs found in previous work (i.e. V1 and FFA). For this purpose, we trained a classifier in a cross-validation approach and decoded face vs vase reports from the MEG sensor-level data (magnetometers and gradiometers). The analysis was shifted over time, yielding temporally resolved decoding results shown in **Figure 1A**. Decoding performance (operationalized as Area Under Curve, AUC) gradually increased following stimulus onset and reached a peak close to the offset of the stimulus mask, the event which prompted the response query. From there on, decoding performance gradually decreased, reaching chance level after approximately 700 ms. Decoding accuracy was significantly above chance after 100 ms ($p_{cluster}$ = 9.9990e-05; tested over the first 350 ms after stimulus onset to exclude the response epoch).

To find the brain regions that provided informative activity, we adapted a previously reported approach (Marti & Dehaene, 2017), which projects the classifier weights from sensor to source space (for sources see **Figure 1A**). This analysis suggested that the brain regions that provide informative activity changed over time (**Figure 1A**). At earlier (< 120 ms) time intervals, informative activity is predominantly localized in and around V1 (centered on MNI coordinates: [12 -88 0] mm). In the subsequent time interval (120 – 200 ms), informative activity is predominantly localized in and around FFA (centered on MNI coordinates: [28 -64 -4] mm). The location and timing of informative neural activity correspond well with reports on the spatiotemporal dynamics of face perception (Kanwisher & Yovel, 2006; Bötzel et al., 1995). For all subsequent analysis, we used the source-reconstructed data from V1 and FFA.

Next, we performed time-frequency analysis in FFA after bi-stable stimulus onset to reveal the oscillatory patterns that contributed to the decoding results. We contrasted trials on which participants reported seeing a face vs. a vase and corrected for multiple time-frequency samples with a cluster-based permutation approach (Maris & Oostenveld, 2007). We found that face reports showed enhanced post-stimulus gamma activity (p_{cluster} = .029; **Figure 1B**) compared to vase reports, consistent with the functional role of gamma activity for visual perception and specifically for face perception (Engell & McCarthy, 2011; Fisch et al., 2009). Over time, this cluster covered the entire relevant post-stimulus time-range and peaked at around 40ms. In terms of frequencies, the cluster covered a range between 48-93Hz and peaked between 60-70Hz. In the lower frequencies, there were no clusters in the time-frequency maps which contributed to the statistical effect (see **Figure 1B**). Overall, this analysis showed that perceiving the stimulus as face was accompanied by enhanced post-stimulus gamma activity in FFA.

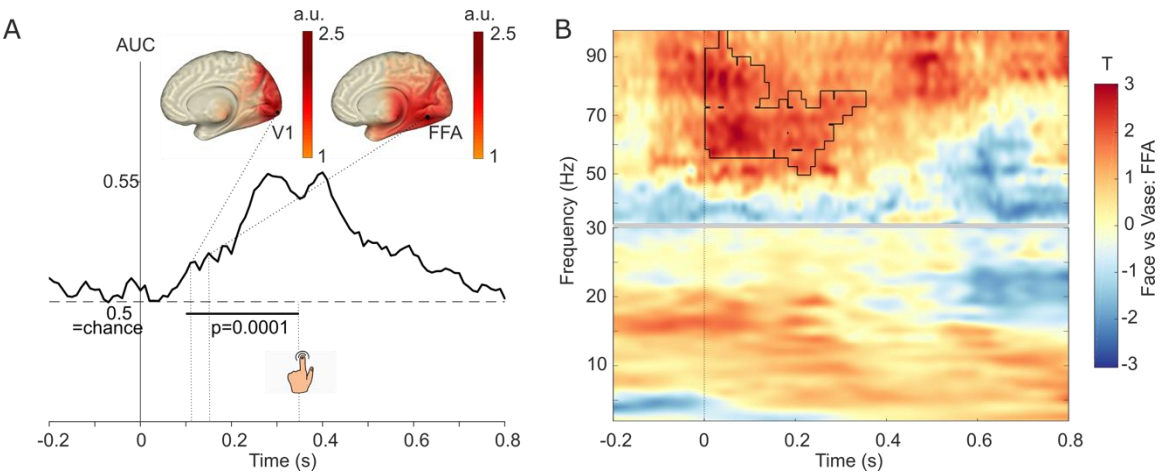

Figure 1: Post-stimulus MEG data contains category-sensitive information with respect to the processing of the Rubin vase stimulus. A) Temporal decoding of face vs vase reports. Inset depicts source reconstruction of the classifier weights (in arbitrary units a.u.), applying a procedure proposed by Marti and Dehaene (2017), at different time points, suggesting temporally changing informative regions (V1 around 100 ms and FFA around 160 ms after stimulus onset). B) Time-frequency contrast of face vs vase reports. Colors represent smoothed T-values obtained from cluster-based permutation testing of the contrast (Face – Vase). Black lines surround the time-frequency gamma-range cluster that drove the significant statistical difference (p_{cluster} = .029).

The MVPA analysis yielded favorable ROIs to test whether pre-stimulus connectivity dynamics between early visual regions (V1) and later category-sensitive regions (FFA) bias the report of upcoming bi-stable percepts (see **Figure 1A**). First, we focused on oscillatory power as an index of local excitability in these regions and tested whether excitability alone predicted the reported categories of upcoming stimuli. Oscillations reflect rhythmic changes in the activity of neural populations and thus reflect phases of high and low excitability (Klimesch et al., 2007). Cluster-based permutation testing revealed no statistical differences in pre-stimulus oscillatory power between face and vase trials, neither in V1 nor in FFA

(**Figure 2A**). Nevertheless, the power spectra in both conditions showed that pre-stimulus oscillatory activity was largely restricted to lower frequencies (5 to 25 Hz, see **Figure 2A**) with a clear peak in the alpha range (~10 Hz). Since frequency-domain measures of connectivity (such as coherence or Granger causality) assume underlying oscillatory activity (i.e. oscillations with high power), we restricted statistical testing for the subsequent connectivity analyses to this frequency range.

Next, we focused on pre-stimulus connectivity between V1 and FFA. Specifically, we hypothesized that increased pre-stimulus coherence between V1 and FFA would precede face reports. A cluster-based permutation test in the frequency range of 5 to 25 Hz revealed that pre-stimulus coherence between V1 and FFA was significantly greater on face vs. vase trials (p_{cluster} = .0036). This increase was most pronounced in a cluster of frequencies ranging from 8.5 to 16.5 Hz (**Figure 2B**). To control for spurious coherence as a result of field spread (Bastos & Schoffelen, 2016), which might explain the high-frequency noise in **Figure 2B**, we repeated the aforementioned analysis using the imaginary part of coherency (Nolte et al., 2004). We obtained qualitatively and quantitatively similar results but with far less high-frequency noise (**Figure S1**).

To further characterize the observed connectivity effect, we used Granger causality to resolve the question of whether increased connectivity prior to face reports represented an increased feedforward drive from V1 to FFA, or an increased feedback drive from FFA to V1. We contrasted face and vase trials separately for the feedforward and feedback directions. The cluster-based permutation test revealed no statistical differences between face and vase reports in the pre-stimulus granger causality estimates in the feedforward direction (V1 to FFA; **Figure 2C, left**); however, for feedback-connectivity we found significantly greater pre-stimulus Granger causality estimates during face trials compared to vase trials (FFA to V1, p_{cluster} = .0115). This increase was most pronounced in a cluster of frequencies ranging from 5 to 10.5 Hz (**Figure 2C, right**).

In sum, we show that face reports (vs. vase reports) were preceded by increased connectivity between V1 and FFA, and that this relative connectivity increase was predominantly driven by an increase in feedback connectivity (FFA to V1).

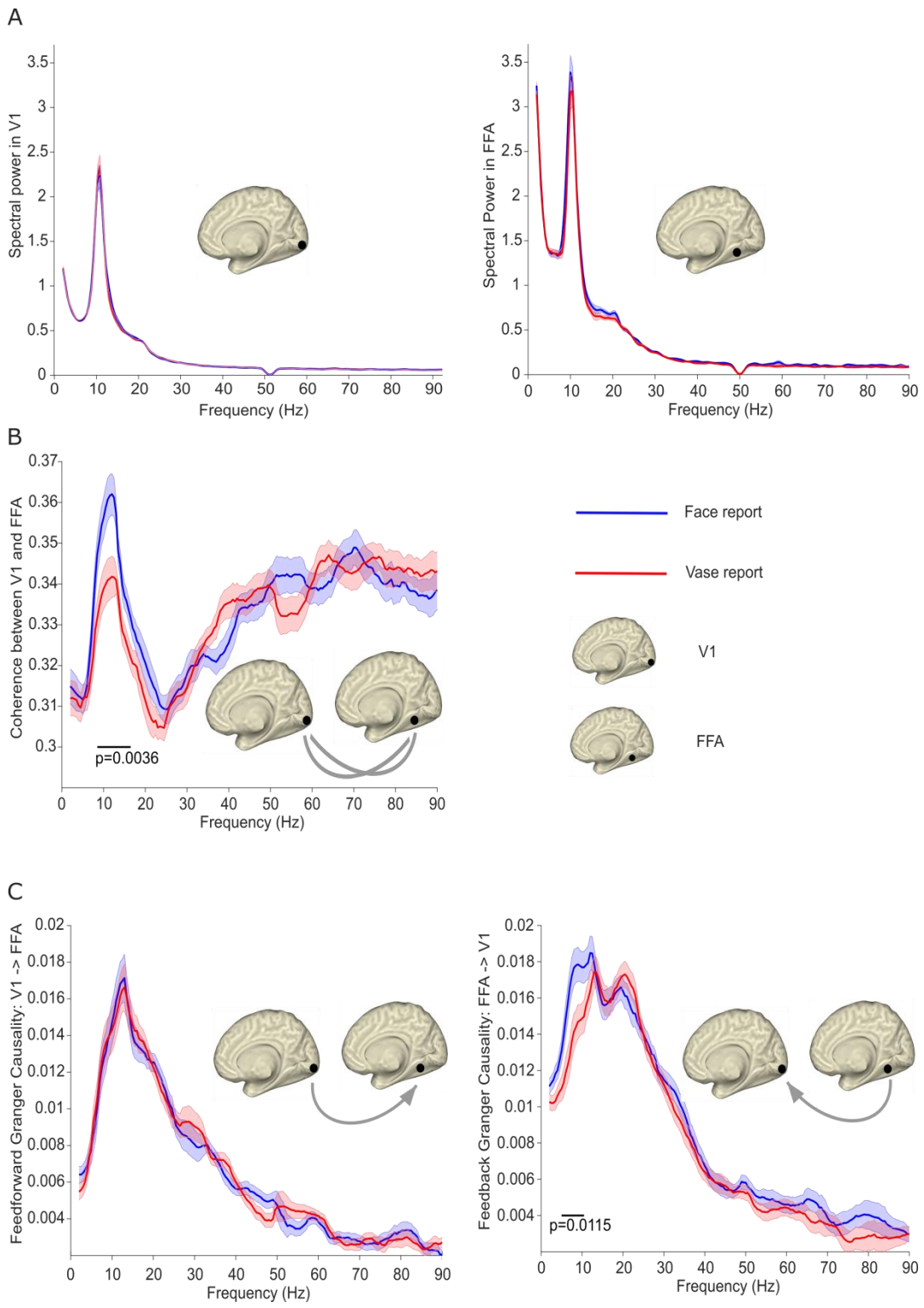

Figure 2: Pre-stimulus MEG connectivity is predictive of upcoming perceptual decision. Shaded error regions represent the standard error of the mean for within-subject designs (Morey, 2008). A) No statistical differences in pre-stimulus spectral power between face and vase trials in either V1 or FFA. B) Compared to vase trials, face trials show increased pre-stimulus coherence between V1 and FF1 in the alpha and beta frequency ranges. C) Compared to vase trials, face trials show increased pre-stimulus feedback connectivity from FFA to V1 in the alpha range, but no differences in pre-stimulus feedforward connectivity from V1 to FFA.

Finally, we focused on the relationship between pre-stimulus connectivity and post-stimulus MVPA category decoding. We extracted for each participant the maximum decoding accuracy (AUC), gamma-band effect, and pre-stimulus feedback connectivity. The maximum gamma effect (max. face – vase power over time and frequencies) and maximum decoding accuracy were correlated (r= .58, p = .008; **Figure 3C**) suggesting that post-stimulus Gamma activity underlies the decoding results. Importantly, we found that the maximum pre-stimulus feedback connectivity was correlated with both the maximum gamma effect (r = .57, p = .009; **Figure 3A**) and maximum decoding accuracy (r = .48, p = .034, **Figure 3B**). In sum, we found that pre-stimulus feedback connectivity strength predicted not only the category of the upcoming percept, but also the strength of post-stimulus neural activity associated with the percept.

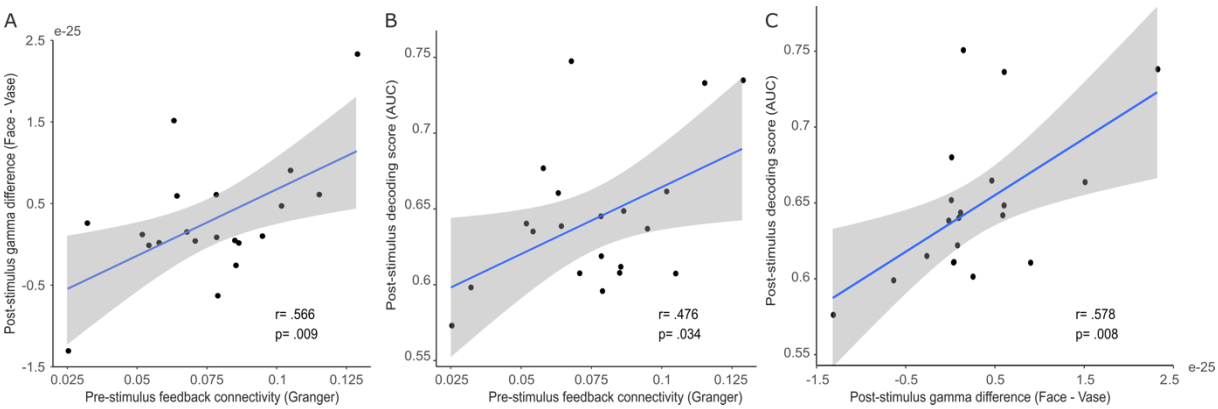

Figure 3: Pre-stimulus connectivity is correlated with post-stimulus activity across participants. A) Maximum pre-stimulus feedback Granger causality estimates are correlated with maximum post-stimulus gamma differences (face – house). B) Maximum pre-stimulus feedback Granger causality estimates are correlated with maximum post-stimulus decoding (AUC) scores. C) Maximum post-stimulus gamma differences (face – house) are correlated with maximum post-stimulus decoding (AUC) scores.

Discussion

While most studies that investigated the effects of pre-stimulus activity on perception were concerned with determining the requisites of successfully detecting stimuli at perceptual threshold (near-threshold paradigms; e.g.: Weisz et al., 2014), our main interest was with the requisites of perceiving one or another content of perception. We found that prior to the Rubin face/vase stimulus onset, FFA was more strongly connected to V1 when face rather than vase was subsequently reported, specifically in the feedback direction of FFA to V1. Connectivity between these two regions was concentrated in the alpha and beta frequency bands (around 5 to 25 Hz). Further, pre-stimulus feedback connectivity strength was correlated with post-stimulus neural activity strength as well as decoding accuracy. Taken together, our findings suggest that spontaneous fluctuations in neural activity in the absence of stimulation can bias

the perceptual content of subsequently presented stimuli by mimicking known dynamics of visual object recognition.

Using MVPA, we first identified the signals that distinguished face from vase responses, and used the source-projected classifier weights to define regions of interest. In line with the substantial body of work which has localized face-responses in time and space using both unambiguous stimuli (for a review see: Kanwisher & Yovel, 2006) and ambiguous stimuli (eg: Sterzer & Rees, 2008; Hasson et al., 2006), we localized the information in our time-resolved classifier to the FFA in a time window centered around 160 ms after stimulus onset. In FFA, we found enhanced post-stimulus gamma activity when a face rather than a vase was reported. This gamma effect correlated with decoding accuracy and thus it likely contributed substantially to the decoder. In addition, we localized early (around 100 ms) information from the decoder to V1. Thus, in line with the work on neural correlates of object recognition and categorization (Kreiman et al., 2000; Sigala & Logothetis, 2002; Freedman et. al, 2003), we established a network consisting of two ROIs, V1 and FFA, in which we then investigated pre-stimulus effects.

We measured pre-stimulus oscillatory power in FFA and V1 and found no differences in either prior to face vs vase reports. We then measured coherence between these two regions and found increased coherence prior to face (vs vase) reports. Next, we separately measured feedback and feedforward Granger causality to determine the directionality of the connectivity effect, and found increased feedback connectivity prior to face (vs vase) reports, but no differences in feedforward connectivity. This is in line with the finding that alpha oscillations subserve feedback connectivity among human (Michalareas et al., 2016) and monkey (Van Kerkoerle et al., 2014) visual cortical areas. Finally, we found that the feedback connectivity estimates (Granger causality) correlated with post-stimulus activity (the gamma effect and decoding accuracy) across participants, strongly suggesting that feedback connectivity biased the subjective reports via its effect on post-stimulus neural processing.

A recent fMRI study employing the Rubin face-vase stimulus (Hesselmann et al., 2008) found that pre- and post-stimulus neural activity was pronounced in the FFA and interpreted the observed pre-stimulus BOLD signal differences as differences in baseline excitability. Indeed, this interpretation is consistent with a large body of work that shows that alpha band activity in task-sensitive sensory regions, an index of neuronal excitability in those regions, predicts behavioral outcome (Frey et al., 2016; Leske et al., 2015; Weisz et al., 2014; Wyart and Tallon-Baudry, 2009). Yet, we found no differences in pre-stimulus alpha activity that could

account for the behavioral outcome. We propose that for a more complete account of the underlying dynamics, the connectivity between regions of interest must also be considered.
Our results are in line with the Windows to Consciousness framework which emphasizes the influence of pre-established connectivity patterns of relevant sensory regions to downstream processing regions on upcoming perceptual processing (Weisz et al., 2014; Ruhnau et al., 2014). We offer a mechanistic account defined in time, space, oscillatory frequency, and directional connectivity. Our account proposes a key role of spontaneous pre-stimulus neural fluctuations in activating connectivity pathways and biasing categorical percepts. Specifically, pre-stimulus feedback connectivity in the alpha range from FFA to V1 represents such a connectivity pathway that biases towards face perception of the Rubin face/vase stimulus.

Conclusion

By recording MEG signals at high temporal resolution before and while people were exposed to a bi-stable stimulus, we showed that the content of visual perception is critically shaped by the ongoing network states, in this case feedback alpha-band connectivity between face-sensitive FFA and early visual cortex. Our work bridges object perception-related pre- and post-stimulus effects and shows how a pre-stimulus network state shapes future processing in a category-sensitive brain region.

Materials and Methods

Participants

20 right-handed volunteers with normal or corrected-to-normal vision participated in this experiment (9 m/11 f, mean age 25.3). Before the experiment, participants filled a form declaring they had no metal objects on their bodies. During the course of the experiment, participants wore non-magnetic clothes. The Ethics Committee of the University of Trento approved the experimental procedure, and all participants gave written informed consent before taking part in the study.

Experimental Procedure

After we placed the Head Position Indicator (HPI) coils on the participants' head, the experiment proper began. Participants sat upright in the MEG system. We instructed them to keep fixation and to avoid eye blinks and movements as best as possible during the experiment. In between the blocks participants took a short break but remained seated in the MEG system. We displayed the visual stimuli via a video projector outside of the MEG chamber. It projected to a back-projection screen in the MEG chamber. We monitored

participants via video camera as they performed the experimental task. At the beginning of each trial, a fixation cross would appear at the centre of the screen for 1 to 1.8 s. After this jittered period, the Rubin vase picture would appear at the centre of the screen for 150 ms (see **Figure 4** for trial example). We presented half of the participants with the original Rubin vase picture (black background) and the other half with a colour inverted version (white background). We did this to ensure that the luminance of the picture did not bias the dominant percept, and post-hoc group analysis revealed no differences between the types of background on the measures reported. After the Rubin vase picture presentation, a mask stimulus would appear for 200 msec. We created this mask by randomly scrambling blocks of pixels of the face-vase picture (see **Figure 4**). After the mask presentation, we asked participants to report whether they saw the face or the vase. The response window was 2 s; if participants did not respond within that window, the next trial would start. Participants responded with the index and middle finger of the left or right hand. We counter-balanced the response hands across subjects. The experiment consisted of 400 trials in total, broken into 4 blocks of 100 trials. We controlled the visual stimulus presentation with Matlab (The MathWorks, Inc) and the Psychophysics Toolbox (Kleiner et al., 2007), and corrected the timings using a photo diode. The procedure of the experiment is illustrated in **Figure 4**.

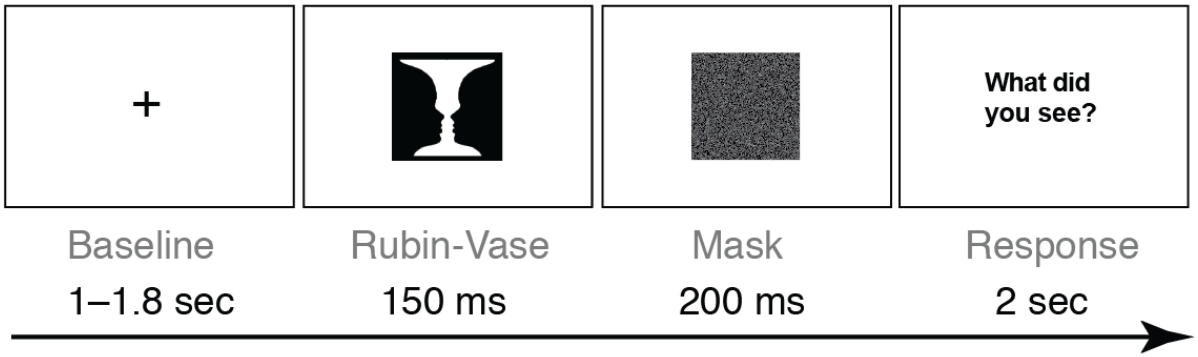

Figure 4: Outline of sample trial.

Behavioural Analysis

We collected behavioural reports after the end of each trial, giving us 400 responses. To test for the stochastic nature of the response we used curve-fitting procedures from the curve-fitting toolbox in Matlab. Specifically, for each participant we binned the data according to how many trials in a row they responded with the same perceptual report. We broke this down in 10 bins with 0 repetitions to 10 repetitions, and then fitted the data to a binomial distribution generated in Matlab across the 10 bins before calculating goodness-of-fit.

MEG Data Acquisition

We carried out the MEG recordings using a 306-channel whole-head VectorView MEG system (Elekta-Neuromag, Ltd., Helsinki, Finland, 204 gradio- and 102 magnetometers) installed in a magnetically shielded chamber (AK3b, Vakuumschmelze Hanau, Germany), with signals recorded at 1000 Hz sample rate. Hardware filters were adjusted to band-pass the MEG signal in the frequency range of 0.01 Hz to 330 Hz. Prior to the recording, we recorded points on the participant's head using a digitizer (Polhemus, VT, USA). These points included the 5 HPI coils, the three fiducials (nasion, left and right pre-auricular points), and over 200 additional points covering the head as evenly as possible. We used the HPI coils to monitor head position during the experiment.

MEG Preprocessing and Source Projection

We pre-processed the data using the Fieldtrip toolbox (Oostenveld et al., 2011). From the raw continuous data, we extracted epochs of 4 seconds lasting from 2.5 seconds before onset of the picture to 1.5 seconds after onset of the picture. This resulted in 400 trials per participant. We applied a high-pass filter on this epoched data at 1 Hz (IIR Butterworth 6-order two-pass filter with 36 dB/oct roll-off), followed by a band-stop filter of 49 – 51Hz to remove power line noise. We then down-sampled the data to 400 Hz. We visually inspected the trials for strong non-physiological artefacts (e.g. channel jumps) and rejected the contaminated trials before computing ICA. We removed components representing typical physiological artefacts (e.g. blinks, ECG) and reconstructed the cleaned raw data. We finally removed the remaining noisy trials by visual inspection. For each participant, we then assigned the trials to the 2 conditions according to the participants' response. Although there was almost equal incidence of face and vase reports to start with, the balance of trial numbers changed after artefact rejection. To ensure a similar signal-to-noise-ratio across conditions, we equalized the trial numbers of face and vase reports by random omission (percentage of trials left in the analysis: M = 79.32%, SD = 15.12%).

We projected the data to source space by applying LCMV (linear constrained mean variance) beamformer filters to the sensor level data (Van Veen et al., 1997). We created anatomically realistic headmodels (Nolte, 2003) using participants' individual structural MRI and the Polhemus digitized scalp shape. For three participants for which we could not obtain an individual MRI, we used a template MRI which was morphed to fit the individuals head shape using an affine transformation. We calculated a three-dimensional source grid (resolution: 8 mm) covering an entire MNI standard brain volume. For each of these points, we computed an LCMV filter using the individual leadfield and the data covariance matrix (estimated separately for the focus of analysis; see below). We used these spatial filters to then project

classifier weights into source space and compute oscillatory and connectivity measures for distinct ROIs (see next sections).

Multivariate Pattern Analysis (MVPA)

We resampled the MEG data to 100 Hz to speed up the MVPA computations. We used MNE Python (Gramfort et al., 2014) which uses Scikit-learn (Abraham et al., 2014; Pedregosa et al., 2011) for the decoding and implemented a 4-fold cross-validation procedure within each subject. For each time-point at each sensor, we Z-normalized the MEG data, trained a Logistic Regression classifier on three folds, and tested on the left-out fold. We operationalized the decoding performance as Area Under the Curve (AUC).

To find out which brain regions contributed to above chance decoding performance the most, we used the classifier weights that the classifier used to separate face from vase reports. To obtain interpretable sensor-level topographies, we multiplied the classifier weights by the data covariance in a first step (Haufe et al., 2014). Then we applied LCMV beamformer filters (using a -.3 to .4 s window to calculate the covariance matrix) to project the weights into source space. At the source level, we abolished polarity differences across voxels by taking the absolute values. This approach is near-identical to the "informative activity" procedure reported in a recent study (Marti and Dehaene, 2017).

Analysis of Post-stimulus Oscillatory Activity

The MVPA provided a clear ROI in right ventromedial temporal cortex corresponding to Fusiform Face Area (FFA) at expected latencies (around 160 ms). We performed time-frequency analysis specifically for this ROI (MNI coordinates: [28 -64 -4] mm) by using the single-trial source-projected time series (using the full epoch length for calculation of the covariance matrix). We estimated power using multitaper Fast Fourier Transform (FFT) with discrete prolate spheroidal sequences (dpss) tapers (Mitra and Pesaran, 1999), with separate window-lengths (.5 s for low frequencies [2-30 Hz in 1 Hz steps] and .3 s for high frequencies [33-99 Hz in 3 Hz steps]). We adapted smoothing to the specific frequencies for which we estimated power (linearly increasing from 2 to 6 Hz for low frequencies and set to +/- 20% of the center frequencies for high frequencies).

Analysis of Pre-stimulus Power, Coherence, and Granger Causality

In addition to the face-sensitive region FFA, the classifier weights source analysis implied the involvement of V1 at earlier time points. To test possible differential communication patterns between V1 and FFA, we first identified the region by averaging the classifier weights between 50 to 120 ms after stimulus onset and applying a 95%-max threshold to mask that

region. This operation resulted in 32 grid points centered around MNI coordinates [12 -88 0] mm. We calculated power, coherence, and Nonparametric Granger causality (Dhamala et al., 2008) in the pre-stimulus period between FFA and V1 in source space (using the full epoch length to calculate the covariance matrix). We used multi-taper frequency transformation with a spectral smoothing of 2 Hz to get Fourier coefficients in the pre-stimulus period (-1 to 0 s), after which we extracted power and computed coherence and bivariate Granger causality. This gave us separate estimates of connection strengths from FFA to V1 ("feedback") and vice versa ("feedforward"). We averaged the results over all grid points within the V1 ROI.

Statistical Analysis of M/EEG Data

For the MVPA analysis, we tested decoding performance against chance level (50%) using a dependent-samples T-test. Since we were interested in periods in which the classifier performs above chance, we used a one-sided T-test.

For all remaining statistical analyses, we used nonparametric cluster permutation tests (Maris and Oostenveld, 2007), comparing a selected test statistic against a distribution obtained from 10000 permutations. We set thresholds for forming clusters at $p < .05$ and considered an effect significant if its probability with respect to the nonparametric distribution was $p < .05$. For the post-stimulus time-frequency contrast in FFA, we tested power in the face-vase contrast in the window of 0 to .35 s, separately for the low and high frequencies, using 2-sided dependent-samples T-tests. We did the same for pre-stimulus power, except that this estimate was not time-resolved. We tested coherence and feedforward and feedback connectivity with 1-sided dependent-samples T-tests as we had hypothesized greater values of these measures on face trials compared to house trials. We restricted the statistical testing window of coherence and Granger to the frequency window 5-25 Hz, which covers the visible peaks in the grand-averaged power and coherence spectra.

References

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. D. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science*, *273*(5283), 1868-1871.

Boly, M., Balteau, E., Schnakers, C., Degueldre, C., Moonen, G., Luxen, A., ... & Laureys, S. (2007). Baseline brain activity fluctuations predict somatosensory perception in humans. *Proceedings of the National Academy of Sciences*, *104*(29), 12187-12192.

Super, H., van der Togt, C., Spekreijse, H., & Lamme, V. A. (2003). Internal state of monkey primary visual cortex (V1) predicts figure–ground perception. *Journal of Neuroscience*, *23*(8), 3407-3414.

Sadaghiani, S., Hesselmann, G., & Kleinschmidt, A. (2009). Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection. *Journal of Neuroscience*, *29*(42), 13410-13417.
Weisz, N., Wühle, A., Monittola, G., Demarchi, G., Frey, J., Popov, T., & Braun, C. (2014). Prestimulus oscillatory power and connectivity patterns predispose conscious somatosensory perception. *Proceedings of the National Academy of Sciences*, *111*(4), E417-E425.
Leske, S., Ruhnau, P., Frey, J., Lithari, C., Müller, N., Hartmann, T., & Weisz, N. (2015). Prestimulus network integration of auditory cortex predisposes near-threshold perception independently of local excitability. *Cerebral Cortex*, *25*(12), 4898-4907.
Frey, J. N., Ruhnau, P., Leske, S., Siegel, M., Braun, C., & Weisz, N. (2016). The tactile window to consciousness is characterized by frequency-specific integration and segregation of the primary somatosensory cortex. *Scientific reports*, *6*, 20805.
Ergenoglu, T., Demiralp, T., Bayraktaroglu, Z., Ergen, M., Beydagi, H., & Uresin, Y. (2004). Alpha rhythm of the EEG modulates visual detection performance in humans. *Cognitive Brain Research*, *20*(3), 376-383.
Wutz, A., & Melcher, D. (2014). The temporal window of individuation limits visual capacity. *Frontiers in psychology*, *5*, 952.
Wutz, A., Melcher, D., & Samaha, J. (2018). Frequency modulation of neural oscillations according to visual task demands. *Proceedings of the National Academy of Sciences*, 201713318.
Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual review of neuroscience*, *19*(1), 577-621.
Miller, E. K., Nieder, A., Freedman, D. J., & Wallis, J. D. (2003). Neural correlates of categories and concepts. *Current opinion in neurobiology*, *13*(2), 198-203.
Kreiman, G., Koch, C., & Fried, I. (2000). Category-specific visual responses of single neurons in the human medial temporal lobe. *Nature neuroscience*, *3*(9), 946.
Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*(6869), 318.
Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, *23*(12), 5235-5246.
Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13.
Rubin, E. (1915). Synsoplevede figurer.
Necker, L. A. (1832). 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*(5), 329-337.
Porta, J. B. (1953). De refractione optices parte: Libri novem. *Naples: Apud Io Iacobum Carlinum & Antonium Pacem*.
Wade, N. J. (1996). Descriptions of visual phenomena from Aristotle to Wheatstone. *Perception*, *25*(10), 1137-1175.
Zou, J., He, S., & Zhang, P. (2016). Binocular rivalry from invisible patterns. *Proceedings of the National Academy of Sciences*, *113*(30), 8408-8413.
Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, *379*(6565), 549.
Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*(4), 753-759.
Sterzer, P. & Kleinschmidt A. (2007). A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences*, *104*(1), 328-8.

Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual review of psychology*, *69*.
Hasson, U., Hendler, T., Bashat, D. B., & Malach, R. (2001). Vase or face? A neural correlate of shape-selective grouping processes in the human brain. *Journal of cognitive neuroscience*, *13*(6), 744-753.
Hesselmann, G., Kell, C. A., Eger, E., & Kleinschmidt, A. (2008). Spontaneous local variations in ongoing neural activity bias perceptual decisions. *Proceedings of the National Academy of Sciences*, *105*(31), 10984-10989.
Ruhnau, P., Hauswald, A., & Weisz, N. (2014). Investigating ongoing brain oscillations and their influence on conscious perception–network states and the window to consciousness. *Frontiers in psychology*, *5*, 1230.
Marti, S., & Dehaene, S. (2017). Discrete and continuous mechanisms of temporal selection in rapid visual streams. *Nature communications*, *8*(1), 1955.
Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *361*(1476), 2109-2128.
Bötzel, K., Schulze, S., & Stodieck, S. R. (1995). Scalp topography and analysis of intracranial sources of face-evoked potentials. *Experimental brain research*, *104*(1), 135-143.
Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.
Engell, A. D., & McCarthy, G. (2010). The relationship of gamma oscillations and face-specific ERPs recorded subdurally from occipitotemporal cortex. *Cerebral cortex*, *21*(5), 1213-1221.
Fisch, L., Privman, E., Ramot, M., Harel, M., Nir, Y., Kipervasser, S., ... & Malach, R. (2009). Neural "ignition": enhanced activation linked to perceptual awareness in human ventral stream visual cortex. *Neuron*, *64*(4), 562-574.
Klimesch, W., Sauseng, P., & Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain research reviews*, *53*(1), 63-88.
Bastos, A. M., & Schoffelen, J. M. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in systems neuroscience*, *9*, 175.
Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., & Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical neurophysiology*, *115*(10), 2292-2307.
Morey, R.D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial in quantitative methods for psychology*, *4*(2), 61-64.
Sterzer, P., & Rees, G. (2008). A neural basis for percept stabilization in binocular rivalry. *Journal of cognitive neuroscience*, *20*(3), 389-399.
Michalareas, G., Vezoli, J., Van Pelt, S., Schoffelen, J. M., Kennedy, H., & Fries, P. (2016). Alpha-beta and gamma rhythms subserve feedback and feedforward influences among human visual cortical areas. *Neuron*, *89*(2), 384-397.
Van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M. A., Poort, J., Van Der Togt, C., & Roelfsema, P. R. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proceedings of the National Academy of Sciences*, *111*(40), 14332-14341.
Bastos, A. M., Vezoli, J., Bosman, C. A., Schoffelen, J. M., Oostenveld, R., Dowdall, J. R., ... & Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron*, *85*(2), 390-401.
Wyart, V., & Tallon-Baudry, C. (2009). How ongoing fluctuations in human visual cortex predict perceptual awareness: baseline shift versus decision bias. *Journal of Neuroscience*, *29*(27), 8715-8725.
Guide, M. U. S. (1998). The mathworks. *Inc., Natick, MA*, *5*, 333.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception*, *36*(14), 1.
Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational intelligence and neuroscience*, *2011*, 1.
Van Veen, B. D., Van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on biomedical engineering*, *44*(9), 867-880.
Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Physics in Medicine & Biology*, *48*(22), 3637.
Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage*, *86*, 446-460.
Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., ... & Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics*, *8*, 14.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. *Journal of machine learning research*, *12*(Oct), 2825-2830.
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J. D., Blankertz, B., & Bießmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage*, *87*, 96-110.
Mitra, P. P., & Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical journal*, *76*(2), 691-708.
Dhamala, M., Rangarajan, G., & Ding, M. (2008). Analyzing information flow in brain networks with nonparametric Granger causality. *Neuroimage*, *41*(2), 354-362.

Supplementary Material

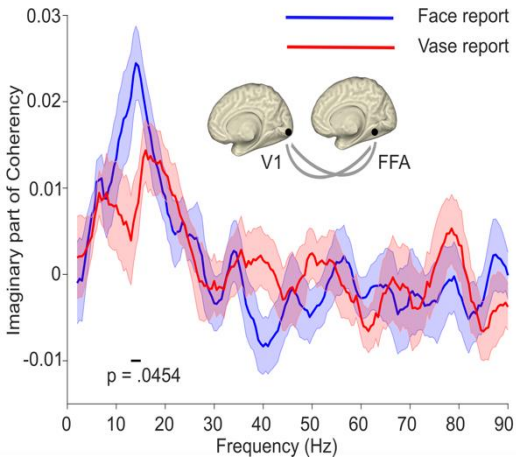

Figure S1: Pre-stimulus contrast (face vs vase) of imaginary part of coherency produces the same result as the coherence contrast, but with less high frequency noise. Shaded error regions represent the standard error of the mean for within-subject designs (Morey, 2008). Compared to vase trials, face trials show increased pre-stimulus imaginary coherency between V1 and FF1 in the alpha/beta frequency range.