Supplementary Information for

Boosting Brain Signal Variability Underlies Liberal Shifts in Decision Bias

Authors: Niels A. Kloosterman[1,2*], Julian Q. Kosciessa[1,2], Ulman Lindenberger[1,2], Johannes Jacobus Fahrenfort[3], Douglas D. Garrett[1,2*]

***Correspondence to:**

Niels Kloosterman
Email: kloosterman@mpib-berlin.mpg.de,

This PDF file includes:

Supplementary text
Figures S1 to S5
SI References

Supplementary Text

Perceptual sensitivity during the visual detection task and its relation to bias

Participants varied widely not only in terms of decision bias, but also in their ability to detect targets, as quantified by the SDT sensitivity measure d' (range 0.26 to 3.97). Distributions of d' in the two conditions largely overlapped (Figure S1A) suggesting that sensitivity was not affected by the experimental decision bias manipulations. However, on average, sensitivity was slightly lower in the liberal (d' = 2.0, SD 0.9) than in the conservative condition (d' = 2.31, SD 0.82, $p<0.001$). Although participants indeed reached similar sensitivity in the two conditions, as indicated by a strong positive correlation in d' between conditions (rho = 0.97, $p < 0.001$)(Figure S1B), all but one participant showed a slight drop in d' in the liberal condition (Figure S1C). This suggests that participants committed relatively many false alarms in the liberal condition to avoid the penalized target misses. In line with this, liberal–conservative criterion and liberal–conservative d' were weakly positively correlated (rho = 0.44, p = 0.09), suggesting that some participants achieved a larger negative (more liberal) criterion shift at the cost of reduced perceptual sensitivity (Figure S1D, right). In contrast, condition-averaged d' and condition-averaged criterion were only weakly correlated (rho = –0.13, p = 0.63), highlighting the independence of overall decision bias and perceptual sensitivity in SDT (Arazi et al., 2017) (Figure S1E). Taken together, participants' perceptual sensitivity was only weakly linked to their decision bias. Further below, we report the correlation between mMSE shifts and criterion shifts while controlling for sensitivity. Please see our previous paper for additional behavioral results, such as reaction times and drift diffusion modeling of the behavioral data (Kloosterman et al., 2019).

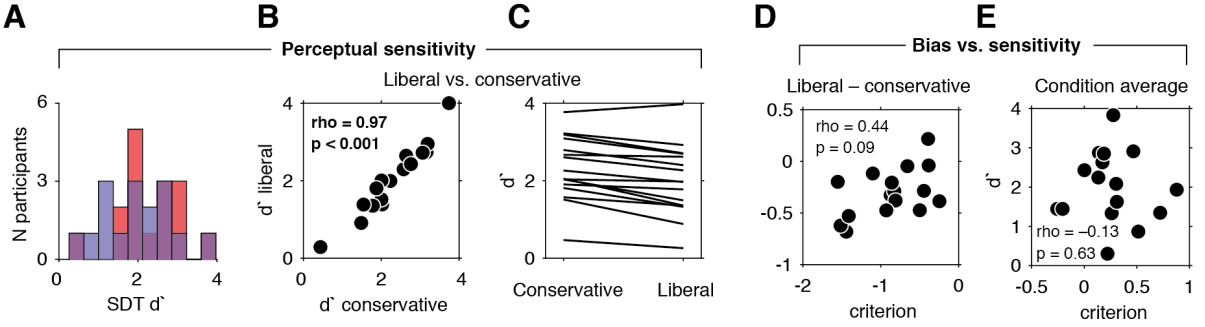

Figure S1 I Perceptual sensitivity and relation between decision bias and sensitivity. A. Distributions of participants' d' in both conditions. A higher d' indicates higher sensitivity. **B.** Scatter plot depicting the correlation between participant's d' in both conditions. **C.** Corresponding individual participant slopes. **D.** Correlation between d' and criterion averaged across conditions. **E.** Correlation between liberal – conservative d' and liberal – conservative criterion.

Correlation between shifts in mMSE and criterion is reliable in split data halves

Correlation analysis with a low number of observations can be unreliable, depending on the amount of data underlying each observation (Yarkoni, 2009). To test whether the observed correlation between liberal – conservative shifts in mMSE and decision bias was reliable, we performed an arbitrary split of the data into odd and even trials, and performed the brain-behavior correlation separately in both halves. We used the correlation cluster reported in Figure 3 of the main paper as a mask to obtain a scalar

mMSE value per participant per data half to correlate with the criterion computed within each data half. We found significant correlations in both halves (odd, r = –0.61, p = 0.013; even, r = –0.64, p = 0.009)(Figure S2A), indicating that the observed relationship between shifts in entropy and bias was reliable. In addition, both the criteria and the mMSE computed in the two halves were positively correlated (Figure S2B), again suggesting reliability. The same pattern occurred when correlating mMSE and criterion from the full data with the estimates from the odd and even trials (Figure S2C and S2D). Taken together, these results suggest that the across participant correlation reported in the main paper is reliable despite the relatively low number of participants, presumably due to the large amounts of data available per participant.

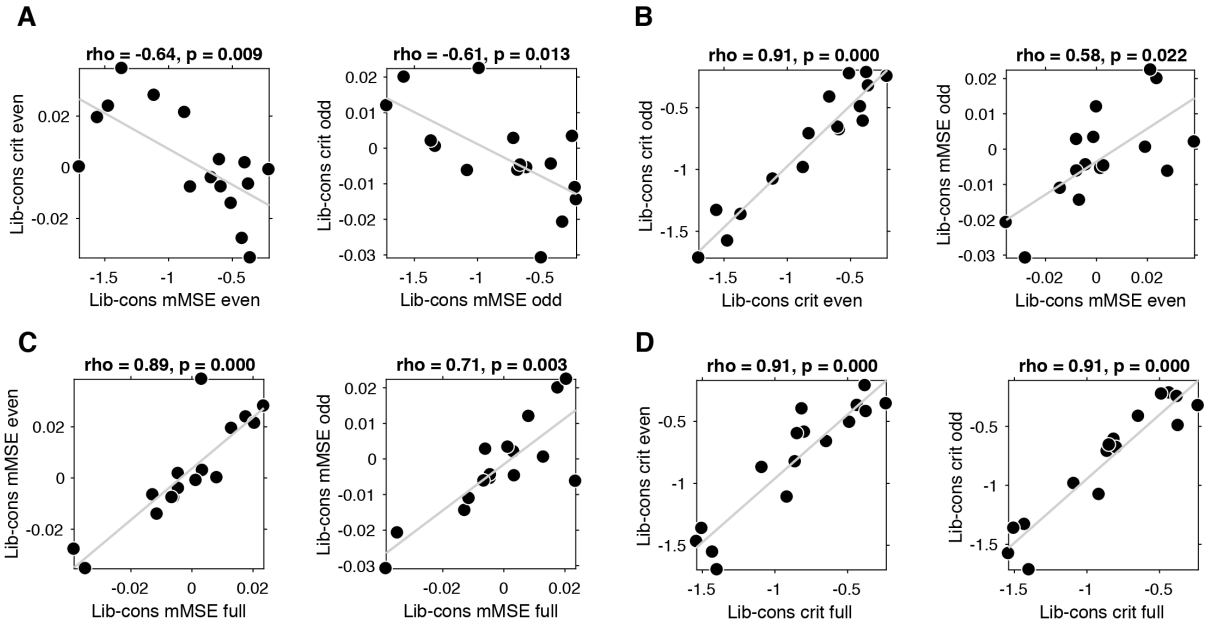

Figure S2 | Correlation between liberal – conservative mMSE and bias shift is reliable in split data halves. A. Spearman correlations across participants in both data halves after arbitrary split into even (left) and odd (right) trials. **B.** Correlations between criteria computed separately in both data halves. **C.** Correlations between mMSE computed using the full data set versus the odd and even data halves. **D.** As C. but for criterion.

Correlation between shifts in mMSE and criterion is not explained by EEG signal standard deviation and spectral power

Both the signal SD and delta showed a negative (liberal–conservative) change-change correlation with mMSE (SD, Spearman's rho = –0.77, p = 0.001; delta, rho = –0.79, p < 0.001, Figure S3A,B), whereas the other frequency bands correlated only weakly (theta, rho = –0.31; alpha, rho = 0.18; beta, rho = 0.26; gamma, rho = 0.29, all p > 0.24. Signal SD and delta were also highly correlated with each other (rho = 0.93, p < 0.001, Figure S1C), indicating that delta contributed strongly to the signal SD, as expected based on the relatively larger amplitude of lower versus higher frequencies in EEG (known as the scale-free or 1/f relationship between frequency and power). Interestingly, however, neither liberal–conservative SD nor delta correlated as strongly with the criterion shift as liberal–conservative mMSE (delta, rho = 0.49, p = 0.059; delta; rho = 0.65, p = 0.008)(Figure S3D,E), suggesting that liberal–conservative mMSE provides the best readout of the decision bias shift. Crucially, we found that the mMSE vs. criterion change-change relationship remained strong and significant when controlling for signal SD (partial rho = –0.82, p < 0.001, Figure S3F), or when

controlling for all major power bands at once (delta, theta, alpha, beta, gamma, partial rho = –0.68, p = 0.02, Figure S3G). In contrast, the correlation was weak when correlating signal SD to criterion while controlling for mMSE (SD, rho = –0.14, p = 0.63, Figure S3H). This indicates that mMSE uniquely explained a substantial portion of the variance in criterion beyond what SD could explain, whereas SD explained almost no unique variance in criterion over mMSE. Finally, we obtained qualitatively similar results when equating the number of trials per condition within participants before computing mMSE, and also when excluding trials that were shorter due to our trial extraction procedure (data not shown).

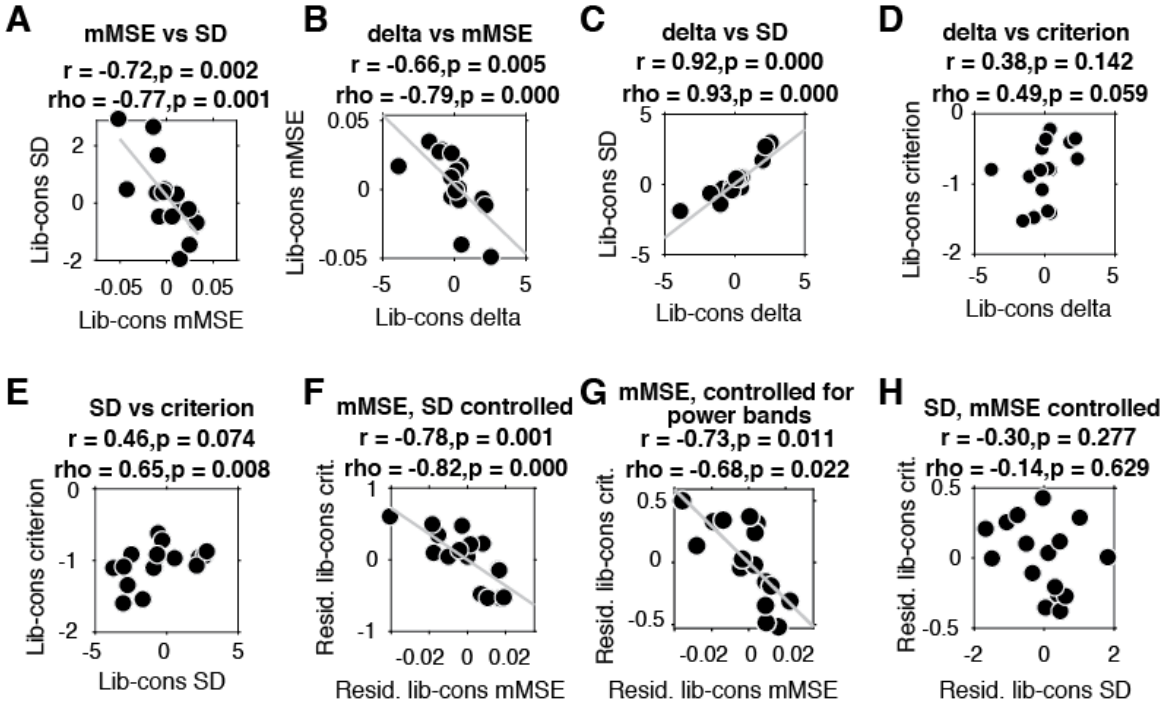

Figure S3 I Correlations between mMSE shifts, criterion and potentially confounding factors in terms of the liberal–conservative difference. **A. – H.** Scatter plots depicting correlations between various factors. See titles and axes labels of each plot for details. Least squares lines (gray) were plotted when the Pearson correlation reached significance ($p<0.05$).

Correlation between mMSE and criterion shift is not observed for neural variability quenching

In contrast to the hypothesis that task performance benefits from increased neural variability, perceptual sensitivity has also been proposed to improve as a function of stimulus-related *decrease* of neural variability, a phenomenon referred to as variability "quenching" (Churchland et al., 2010). Quenching is directly predicted by attractor models of brain organization (Wang, 2002), and is consistent with SDT's main principle that suppression of neural noise enhances perception (Green and Swets, 1966). Quenching has also been reported in the human EEG in a variance reduction across trials in visual cortex following stimulus onset (Arazi et al., 2017), although this type of quenching can be attributed to the well-known suppression of low-frequency spectral power (see e.g. (Kloosterman et al., 2015)). We asked whether quenching can also be observed in mMSE, and whether it is linked to the decision bias shift and/or low-frequency power suppression. We computed the transient mMSE modulation following trial onset by converting the mMSE values into percent signal change with respect to the prestimulus baseline, separately for each timescale. We found that mMSE in

longer timescales was transiently boosted in midfrontal and lateral posterior electrodes (Figure S4A), while being simultaneously suppressed in shorter timescales in posterior electrodes, indicating quenching (Figure S4B). Poststimulus mMSE modulation in frontal cortex was significantly stronger in the liberal than in the conservative condition (cluster $p = 0.042$), which was mostly limited to one midfrontal electrode (Figure S4C). Controlling for signal SD (which is most strongly affected by low-frequency power due to the 1/f nature of EEG signals) completely abolished the occipital negative, but not the frontal positive cluster, indicating that mMSE quenching could indeed be explained by low-frequency spectral power. To test which exact spectral power band could explain the transient mMSE modulation, we correlated mMSE across participants in the positive and negative clusters with theta (3-7 Hz) and beta (12-30 Hz) power. We found that the positive cluster was uncorrelated with these power bands (Figure S4D), but we did indeed observe a significant positive correlation between quenching and beta suppression ($rho = 0.74$, $p = 0.002$) (Figure S4E), indicating that participants with stronger low-frequency suppression quenched more. We found no significant correlation between quenching and alpha, presumably because we did not observe transient alpha suppression after trial onset in our continuous visual stimulation paradigm. Finally, when correlating liberal-conservative mMSE modulation and the criterion shift, we found a negative cluster (cluster $p=0.05$) as found for raw mMSE, but the cluster was smaller than in the principal analysis (see Figure 3A), located only in left temporal and parietal electrodes, and it appeared scattered across shorter and longer timescales during the poststimulus period (Figure S4F). Importantly, we did not find significant relationships between liberal–conservative mMSE quenching in posterior electrodes and the decision bias shift ($rho = 0.06$. $p = 0.83$), and also not for the shift in perceptual sensitivity (d', $rho = –0.20$. $p = 0.45$). Taken together, we indeed observed transient mMSE quenching upon stimulus onset that could be attributed to low-frequency power suppression, but it was not linked to behavioral changes.

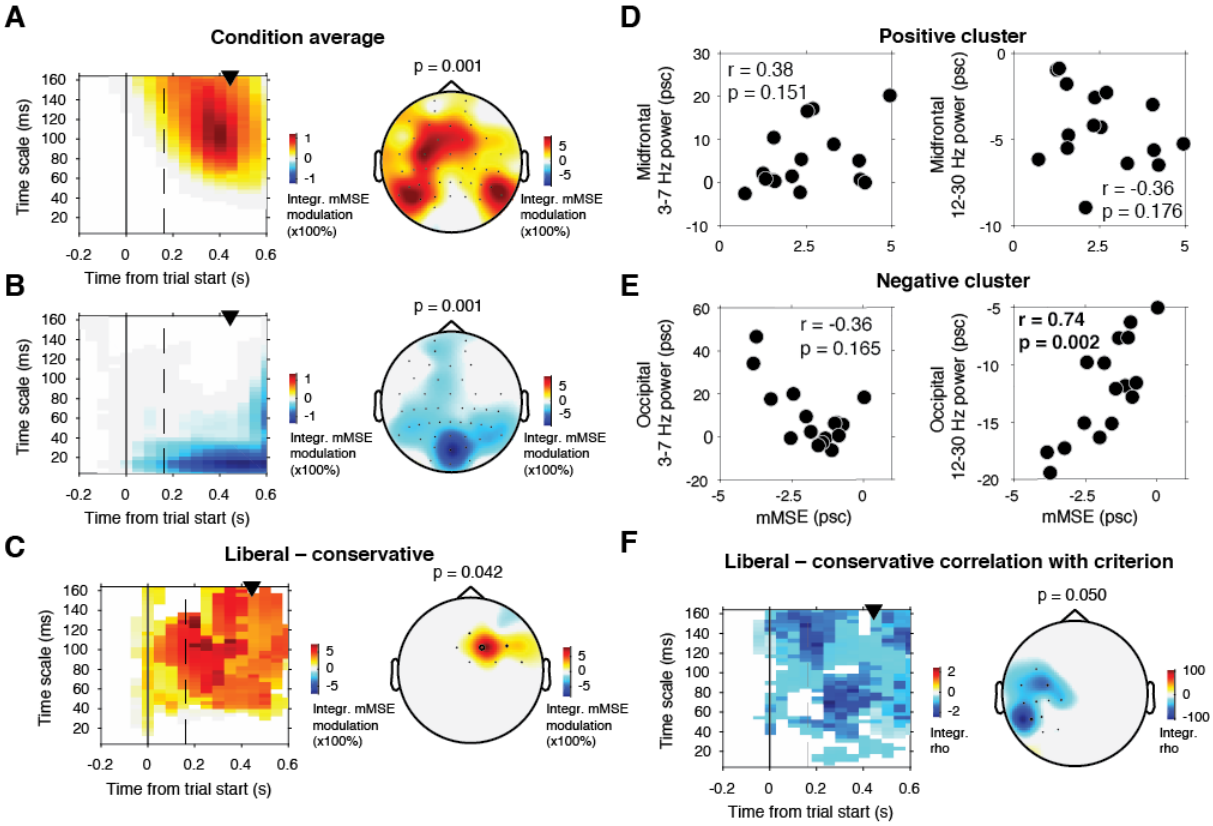

Figure S4 I mMSE modulation with respect to pre-trial baseline. A. Significant positive cluster observed in longer timescales after normalizing mMSE values to percent signal change (psc) units with respect to the pre-trial baseline (–0.2 to 0 s) and averaging across conditions. **B.** Negative ("quenching") cluster observed in shorter timescales in posterior electrodes. Conventions as in Figure 3. **C.** Significant positive cluster observed in mid-frontal electrodes in the liberal–conservative contrast. **D.** Correlation between mMSE modulation in the positive cluster depicted in A. and spectral power modulation in midfrontal electrodes. Left panel, 3-7 Hz; right panel, 12-30 Hz. **E.** As in D. but for the occipital quenching cluster. **F.** Significant cluster resulting from the correlation between liberal–conservative mMSE modulation with liberal–conservative SDT criterion. Conventions as in Figure 3 of the main paper.

No relationship between bias shift and transient spectral power modulation

To test whether transient spectral power modulation following stimulus onset related to individual criterion shifts, we finally repeated the three-dimensional clustering analysis for spectral power modulation normalized with respect to the prestimulus baseline. In the condition average, we observed increased power modulation in posterior electrodes in the 60–100 Hz (gamma) frequency range, reflecting target stimulus processing (Figure S5A)(Kloosterman et al., 2019). Further, we observed a power increase below 7 Hz in lateral occipital and midfrontal electrodes (Figure S5B). The topography of this modulation resembled the baseline corrected mMSE results in Figure S4. At the same time, power modulation was suppressed in the 42–58 Hz and 12–30 Hz range (Figure S5C and S5D). We neither found significant power modulation clusters in the liberal–conservative contrast itself, nor in its change-change correlation with criterion. The same was true for raw spectral power. We note that the narrow-band ~12 Hz, ~25 Hz and ~50 Hz Hz power suppression in Figures S5C and S5D are plausibly related to the removal of EEG activity phase-locked to stimulus onset due to subtracting the ERP (Klimesch et al., 1998) prior to spectral and mMSE analysis. These modulations reflect the the Steady-State Visual Evoked Potential (SSVEP) and

its (sub-)harmonics that were evoked by the 25 Hz presentation rate of the visual stimuli (Methods). Taken together, these results provide no evidence for an inter-individual association between shifts in spectral power modulation and decision bias shifts.

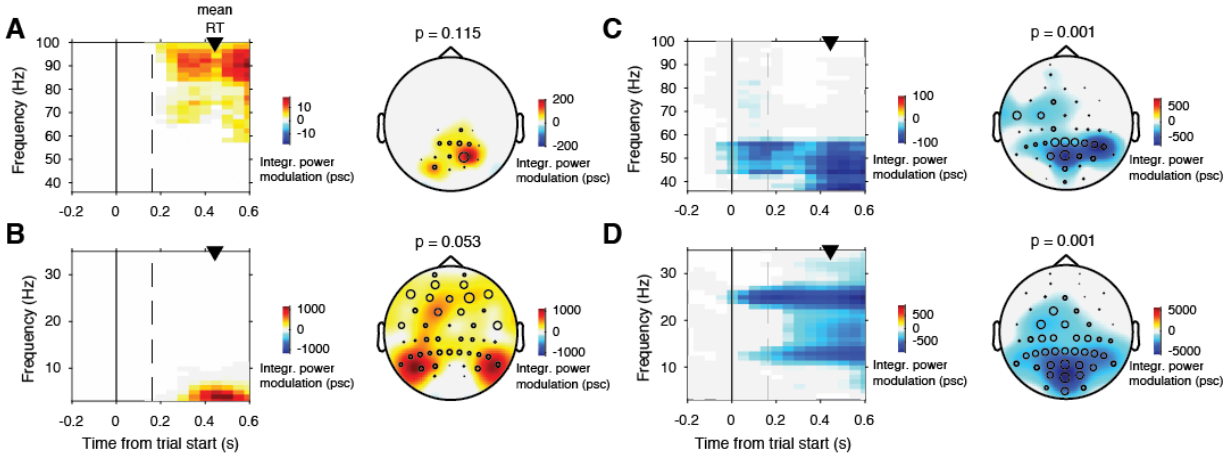

Figure S5 | Spectral power normalized with respect to the pre-trial baseline. A. Significant clusters observed after normalizing spectral power to percent signal change units with respect to the pre-trial baseline (–0.2 to 0 s) and averaging across conditions. **A.** Positive cluster in frequencies > 60 Hz (gamma) in posterior electrodes. **B.** Positive cluster below 7 Hz in lateral posterior and midfrontal electrodes extending globally across the scalp. **C.** Negative 42–58 Hz cluster in posterior and left central electrodes. **D.** Negative cluster in frequencies between ca. 12 and 30 Hz in posterior electrodes. Conventions as in Figure 4, except that frequencies instead of timescales are plotted on the y-axis.

SI References

Arazi A, Censor N, Dinstein I (2017) Neural Variability Quenching Predicts Individual Perceptual Abilities. J Neurosci 37:97 109.

Churchland MM et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369 378.

Green D, Swets J (1966) Signal detection theory and psychophysics. Society 1:521.

Klimesch W, Russegger H, ppelmayr, Pachinger T (1998) A method for the calculation of induced band power: implications for the significance of brain oscillations. Electroencephalogr Clin Neurophysiology Evoked Potentials Sect 108:123 130.

Kloosterman NA, de Gee J, Werkle-Bergner M, Lindenberger U, Garrett DD, Fahrenfort J (2019) Humans strategically shift decision bias by flexibly adjusting sensory evidence accumulation. Elife 8:e37321.

Kloosterman NA, Meindertsma T, Hillebrand A, van Dijk BW, Lamme VA, Donner TH (2015) Top-down modulation in human visual cortex predicts the stability of a perceptual illusion. J Neurophysiol 113:1063–1076.

Wang X-J (2002) Probabilistic Decision Making by Slow Reverberation in Cortical Circuits. Neuron 36:955–968.

Yarkoni T (2009) Big Correlations in Little Studies: Inflated fMRI Correlations Reflect Low Statistical Power-Commentary on Vul et al. (2009). Perspect Psychol Sci 4:294 298.