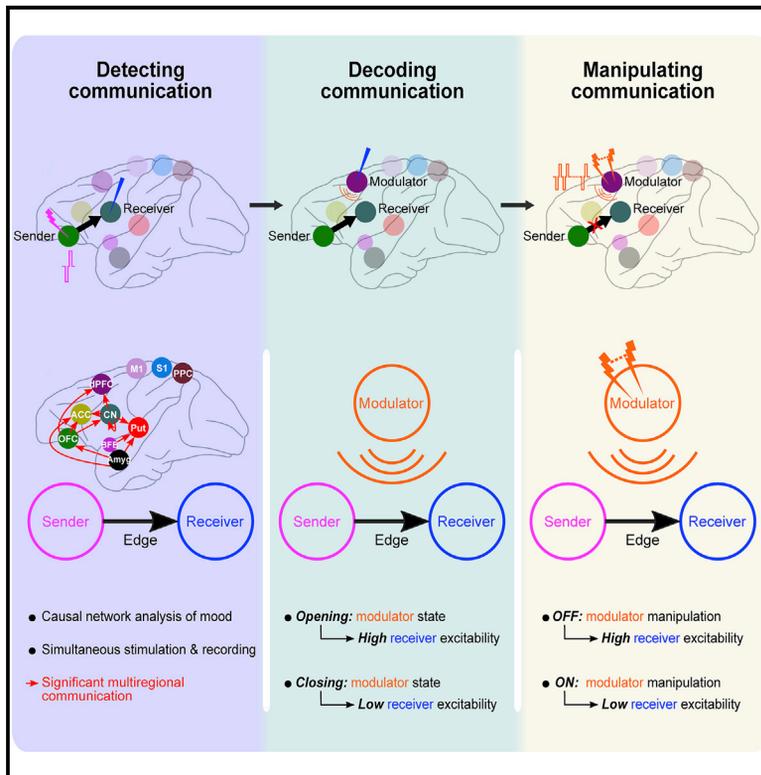


A Causal Network Analysis of Neuromodulation in the Mood Processing Network

Graphical Abstract



Authors

Shaoyu Qiao, J. Isaac Sedillo,
Kevin A. Brown, Breonna Ferrentino,
Bijan Pesaran

Correspondence

bijan@nyu.edu

In Brief

Qiao et al. performed a novel causal network analysis that reveals mechanisms of excitability-based multiregional communication in the large-scale primate mood processing network. The results suggest next-generation brain therapies that can combine neuromodulation with neural decoding in brain-machine interfaces to correct disordered multiregional communication.

Highlights

- Causal network analysis reveals excitability-based multiregional communication
- Modulator decoding predicts network excitability of multiregional communication
- Modulator sbTetMS modulation temporarily disrupts multiregional communication
- Suggesting a novel therapeutic strategy combining modulator decoding and modulation

Article

A Causal Network Analysis of Neuromodulation in the Mood Processing Network

Shaoyu Qiao,¹ J. Isaac Sedillo,¹ Kevin A. Brown,¹ Breonna Ferrentino,^{1,4} and Bijan Pesaran^{1,2,3,5,*}

¹Center for Neural Science, New York University, New York, NY 10003, USA

²Neuroscience Institute, New York University Langone Health, New York, NY 10016, USA

³Department of Neurology, New York University Langone Health, New York, NY 10016, USA

⁴Present address: Graduate School of Biomedical Sciences, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

⁵Lead Contact

*Correspondence: bijan@nyu.edu

<https://doi.org/10.1016/j.neuron.2020.06.012>

SUMMARY

Neural decoding and neuromodulation technologies hold great promise for treating mood and other brain disorders in next-generation therapies that manipulate functional brain networks. Here we perform a novel causal network analysis to decode multiregional communication in the primate mood processing network and determine how neuromodulation, short-burst tetanic microstimulation (sbTetMS), alters multiregional network communication. The causal network analysis revealed a mechanism of network excitability that regulates when a sender stimulation site communicates with receiver sites. Decoding network excitability from neural activity at modulator sites predicted sender-receiver communication, whereas sbTetMS neuromodulation temporarily disrupted sender-receiver communication. These results reveal specific network mechanisms of multiregional communication and suggest a new generation of brain therapies that combine neural decoding to predict multiregional communication with neuromodulation to disrupt multiregional communication.

INTRODUCTION

Brain-machine interfaces (BMIs) seek to achieve therapeutic effects by decoding recorded neural activity (Bensmaia and Miller, 2014; Shanechi, 2019; Shenoy and Carmena, 2014). Neuromodulation technologies offer therapies by intervening to alter neural activity (Johnson et al., 2013). The combination of neural decoding and neuromodulation technologies has broad and wide-ranging potential to treat refractory neurological and neuropsychiatric disorders (Shanechi, 2019). Both technologies can be performed using implantable devices to stimulate, as in deep brain stimulation (DBS; Lozano et al., 2019) and direct cortical surface stimulation (Borchers et al., 2011) and recording (Shanechi, 2019). Ultimately, next-generation neuromodulation- and decoding-based therapies may be able to repair and even cure the disordered brain by recruiting neural plasticity (Braun et al., 2018; Rajasethupathy et al., 2016).

Despite the therapeutic promise and recent advances (Ezzyat et al., 2017, 2018; Inman et al., 2018a; Rao et al., 2018; Suthana et al., 2012; Titz et al., 2017; Widge et al., 2019; Wu et al., 2018), significant hurdles exist, in particular for treating mood disorders (Shanechi, 2019). The primate mood network spans the orbitofrontal cortex (OFC), anterior cingulate cortex (ACC), dorsal prefrontal cortex (dPFC), striatum, amygdala (Amyg), hippocampus, nucleus accumbens (NAc), basal forebrain (BFB), and ventral tegmental area. Functional interactions between these brain re-

gions play a central role in regulating mood state variations (Drevets, 2001; Drysdale et al., 2017; Price and Drevets, 2010; Sani et al., 2018) and modulating distributed value computations in decision-making (Hunt and Hayden, 2017; Nestler and Carlezon, 2006). Dysfunction in the mood network alters motivation to guide behavior, underlies mood disorders such as major depression and anxiety (Williams, 2017), and often involves multiple brain regions distributed across the frontostriatal limbic network (Price and Drevets, 2010; Williams, 2017). Neuromodulation has often been based on focal stimulation (Dougherty et al., 2015; Inman et al., 2018b; Mayberg et al., 2005; Rao et al., 2018), and recent neural decoding approaches have decoded mood states to target neuromodulatory interventions (Sani et al., 2018). Although clinical targets for therapeutic stimulation, such as for mood, are increasingly viewed within a circuit-based model (Braun et al., 2018; Shanechi, 2019) in which dysfunction arises from disordered multiregional communication, how specific mechanisms of multiregional communication can be decoded and manipulated remains unknown.

Neuromodulation may suppress or facilitate activity at specific sites as well as weaken or strengthen multiregional communication. Disambiguating network mechanisms is a grand challenge in modern neuroscience. Communication is often estimated from correlations in activity at each site, or node (Bassett et al., 2018), but correlations are confounded by changes in activity at each site and common inputs to both sites (Pesaran et al.,

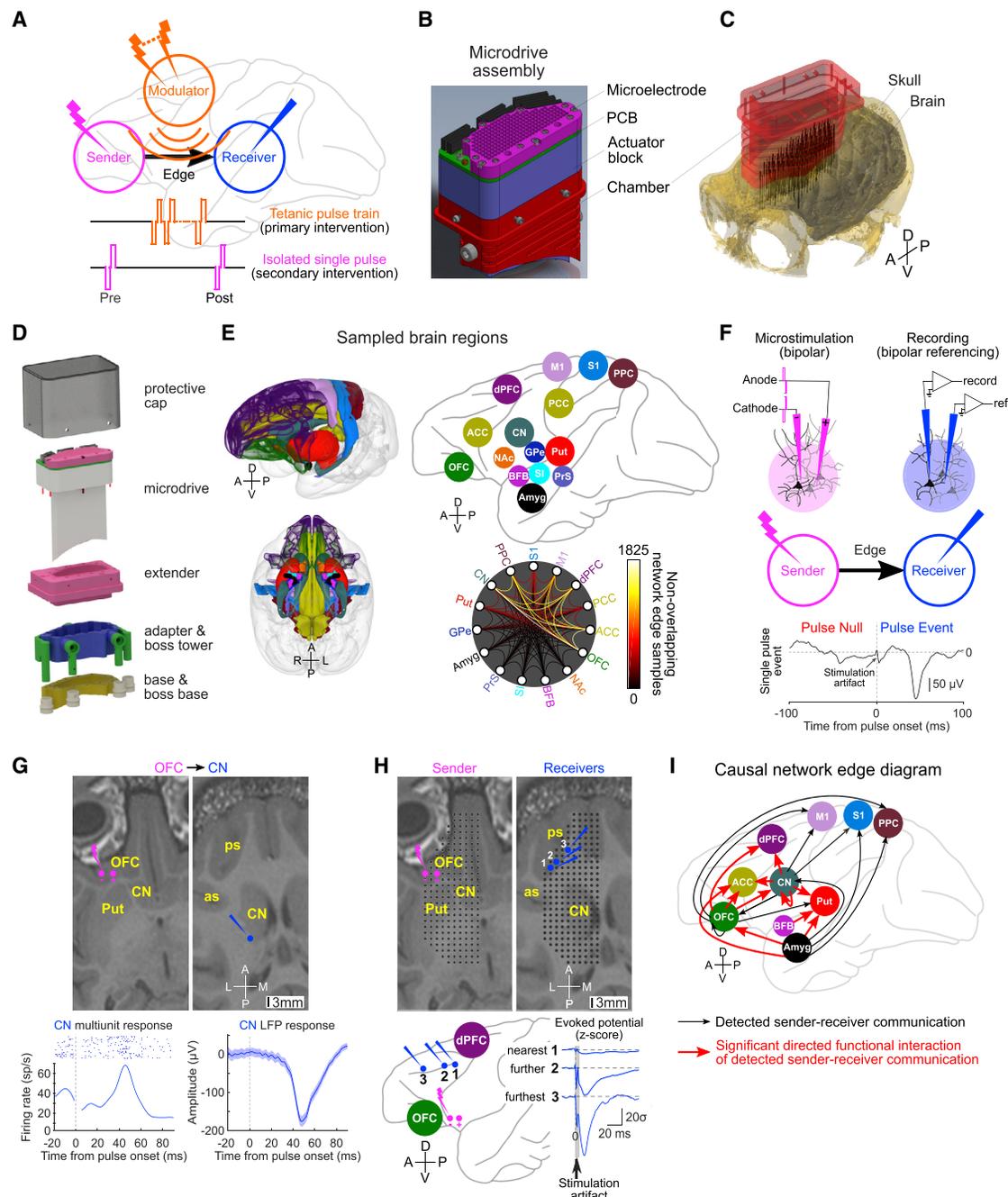


Figure 1. Targeting and Sampling the Primate Mood Processing Network

(A) Experimental design using multisite spatiotemporal patterned microstimulation for decoding and manipulating multiregional communication and analyzing neuromodulation hypotheses of network mechanisms of action of sbTetMS-based intervention (primary)—edge modulation versus node modulation.

(B) Computer-aided design model of the customized large-scale microdrive assembly.

(C) Co-registered chamber with MRI-based skull and brain surface reconstruction. Vertical black lines, co-registered electrode tracks.

(D) Computer-aided design model of the modular chronic implant system.

(E) Sampled brain regions (network nodes), illustrated using a rhesus macaque brain atlas (Calabrese et al., 2015) from the Scalable Brain Atlas. Non-overlapping network edge samples between sampled network nodes are shown.

(F) Single, bipolar, biphasic, charge-balanced microstimulation pulses at the sender to causally identify sender-receiver communication (black arrow). Shown is an example of the CN receiver response to a single microstimulation pulse at the OFC sender (30 μ A, 100 μ s/phase).

(G) Single bipolar microstimulation pulses (10 μ A, 100 μ s/phase) at the OFC sender drove multiunit and LFP responses at the CN receiver (response onset, \sim 20 ms; response peak, \sim 46 ms after pulse onset). Data around the pulse onset were omitted for multiunit activity analysis because of stimulation artifacts. Shading, ± 1 SEM (n = 52).

(legend continued on next page)

2018). To better identify network mechanisms, causal approaches based on simultaneous stimulation and recordings from multiple, interconnected brain regions are required but have been technically prohibitive in the awake primate brain. New approaches are therefore needed to identify, decode, and manipulate multiregional communication to realize next-generation brain therapies.

Here we address the need for novel neural decoding- and neuromodulation-based therapies by analyzing multiregional communication. We present a neural decoding approach to test the network mechanism of action of a neuromodulatory intervention, short-burst tetanic microstimulation (sbTetMS; ≤ 2 s, 100 Hz or 200 Hz), by stimulating and recording across the large-scale mood processing network in two awake rhesus macaque monkeys. Our experiments analyze multiregional communication by extending the state of the art with stimulation-recording coverage that spans the OFC, dPFC, primary motor cortex (M1), primary sensory cortex (S1), parietal cortex, cingulate cortex, striatum, pallidum, BFB, and Amyg. We use the concept of neural excitability (Keller et al., 2018; Volgushev et al., 1998), which describes how each node responds to input according to the weight of an edge, to causally estimate multiregional communication. The resulting causal network analysis permits rigorous dissection of network communication hypotheses. Our results provide a new basis for developing next-generation therapies in which neuromodulation specifically disrupts multiregional communication revealed by decoding network excitability.

RESULTS

Figure 1A illustrates the experimental design for decoding and altering multiregional communication. We first measure neural excitability by measuring neural responses across the network to a secondary intervention, an isolated single microstimulation pulse. Decoding of the neural responses is then used to identify modulators that predict network excitability moment by moment. We then deliver the primary intervention, a sbTetMS pulse train, at different sites and repeat the secondary intervention.

Targeting and Sampling the Mood Processing Network

To decode multiregional communication and analyze neuromodulation across the mood processing network, we developed a novel implantable microdrive to target and sample a large-scale mood processing network for causal network analysis (STAR Methods; Figures 1B–1D; Figure S1). With this device, we successfully targeted and repeatedly sampled a network spanning 15 cortical and subcortical brain regions across two monkeys. The sampled network included seven cortical regions (OFC,

ACC, posterior cingulate cortex [PCC], dPFC, M1, S1, and posterior parietal cortex [PPC]) and eight subcortical regions (caudate nucleus [CN], putamen [Put], external *globus pallidus* [GPe], Amyg, presubiculum [PrS], *substantia innominata* [SI], BFB, and NAc).

The causal network analysis let us study neural interactions between sets of recording and stimulation sites located at nodes in the network. Across all 15 nodes, we sampled from a grand total of 4,646 sites in monkey M and 4,897 sites in monkey A (Figure 1E). Across both monkeys, we obtained 15,734 non-overlapping samples from 21 cortico-cortical edges, 11,276 non-overlapping samples from 56 cortical-subcortical edges, and 1,087 non-overlapping samples from 28 subcortical-subcortical edges (STAR Methods; Figure 1E). We targeted our analysis to 9,548 edge samples from the mood processing network connecting the OFC, dPFC, ACC, CN, Put, and Amyg.

Causal Network Analysis Reveals Multiregional Communication

We causally estimated multiregional communication with a secondary intervention delivering a single, low-amplitude, bipolar, biphasic, charge-balanced microstimulation pulse at a given node (sender) while recording evoked responses at other sampled nodes (receiver) (STAR Methods). Responses to isolated bipolar microstimulation pulses appear to reflect responses by neurons in the receiver because of synaptic inputs driven by stimulation at the sender (sender \rightarrow receiver; Figure 1F). Consistent with this interpretation, single bipolar microstimulation pulses drove spiking activity from neurons in the receiver as well as the evoked local field potential (LFP) activity (Figure 1G). Thus, bipolar re-referenced LFP responses reflected a local neuronal source. Moreover, analyzing the spread of signals from the sender site to the receiver sites confirmed that using bipolar microstimulation and recording configurations revealed focal responses at receiver sites distant to the sender (Figure 1H). Comparison with control recordings during monopolar microstimulation (Figure S2) confirmed that bipolar microstimulation was less likely to contaminate receiver site responses because of current spread. In the following, we refer to each recording from a pair of bipolar re-referenced electrodes as a site sampled in the brain.

To detect stimulation pulse-evoked responses, we used an optimal signal detection algorithm (Banerjee et al., 2010), which we termed stimulation-based accumulating log-likelihood ratio (stimAccLLR; STAR Methods; Figure S3). Across 6,910 pairs of sites tested in 132 causally sampled network edges, we detected and characterized 110 directed functional interactions (sender-receiver communication) in 21 network edges (Figures S4 and S5; Tables S1 and S2). On average, we detected sender-receiver communication between a pair of sites with a

(H) Single bipolar microstimulation pulses (50 μ A, 100 μ s/phase) at the OFC sender revealed significant focal evoked responses (Z-scored LFP activity, $n = 529$) in the principal sulcus (ps) distant from the stimulating site, not nearby. Also shown is a horizontal MRI slice at OFC stimulation sites (left, magenta dots) and ps recording sites (right, blue dots).

(I) Causal network edge diagram showing detected sender-receiver communications (black thin arrows) and statistically significant directed functional interaction of detected sender-receiver communication (thick red arrows; $p < 0.05$, binomial test).

Anterior (A), posterior (P), left (L), right (R), medial (M), lateral (L), dorsal (D), and ventral (V) directions are shown.

See also Figures S1–S5 and Tables S1 and S2.

probability of 1.6% (110 of 6,910). Because not all pairs of sites are expected to functionally interact, we tested whether some pairs of sites from some network edges interacted more than others (Table S2). A subset of network edges (11 of 21) revealed significant directed functional interactions ($p < 0.05$, binomial test with 1.6% chance; Figure 1I): OFC→ACC (2.59%, 15 of 579 samples, $p = 0.026$), OFC→CN (3.44%, 19 of 553 samples, $p = 7.3 \times 10^{-4}$), OFC→dPFC (2.65%, 13 of 490 samples, $p = 0.028$), CN→ACC (3.14%, 9 of 287 samples, $p = 0.018$), CN→dPFC (4.1%, 8 of 195 samples, $p = 4.36 \times 10^{-3}$), CN→Put (6.12%, 3 of 49 samples, $p = 7.7 \times 10^{-3}$), Amyg→ACC (4.71%, 4 of 85 samples, $p = 0.012$), Amyg→OFC (5.56%, 2 of 36 samples, $p = 0.019$), Amyg→Put (5.56%, 1 of 18 samples, $p = 0.033$), and BFB→Put (25%, 1 of 4 samples, $p = 1.49 \times 10^{-3}$). We also observed significant interactions within a single node, such as CN→CN (3.43%, 8 of 233 samples, $p = 0.013$). Thus, the causal network analysis defined a cortical-subcortical limbic network (OFC-dPFC-ACC-CN-Put-Amyg-BFB).

We next set out to determine how sbTetMS neuromodulation alters or modulates multiregional communication and how such communication can be decoded from activity at modulator sites.

sbTetMS Suppresses Excitability of Sender-Receiver Communication

We first asked whether the network mechanism of sbTetMS neuromodulation is consistent with either or both of the edge and node modulation hypotheses. We tested the edge modulation hypothesis by asking whether stimulating the sbTetMS modulator site alters the receiver evoked potential used to measure the strength of sender-receiver communication (STAR Methods; Figure 2A). We did so by first probing the sender-receiver pair before delivering sbTetMS at a modulator site that, critically, was not at the sender or receiver site, and then probing the sender-receiver pair again immediately after sbTetMS (median latency between sbTetMS offset and post-tetanic single pulse onset [τ_{Post}] = 50 ms).

Figure 2B illustrates potential outcomes for the edge modulation hypothesis as suppression or facilitation of neural excitability, as revealed by pulse-triggered evoked responses in the pre-tetanic and post-tetanic epochs. We also tested the node modulation hypothesis for neural excitability by asking whether the sbTetMS modulator alters activity within the receiver node alone. For analyzing receiver node activity before and after sbTetMS, we focused the analysis on LFP activity because it is clinically important and yielded reliable, long-term recordings across the sampled network. We quantified node activity 100 ms immediately before (pre-state) and 200–300 ms after (post-state) the pulse-sbTetMS-pulse sequence (Figure 2A) using the root mean square (RMS) power of LFP activity in the beta (β ; 10–40 Hz) frequency band, a key signature of node activity used in DBS studies (Little and Brown, 2014), as well as in the high-gamma (high- γ , 70–150 Hz) frequency band, a key signature of node activity that reflects local spiking of populations of neurons (Manning et al., 2009). Because the receiver response to the post-tetanic pulse is due to the sender, we defined the post-state as activity shortly after the pulse-triggered evoked response epoch. Figure 2C illustrates potential outcomes for the receiver node modulation hypothesis.

Figures 2D–2G present an example of a sbTetMS modulator in the same network node as the receiver. Here we stimulated an Amyg sender and observed OFC receiver responses (Figure 2E; Figures S6A–S6C). Fluctuations in the evoked response were present pulse by pulse (Figure 2F). sbTetMS delivered at an OFC sbTetMS modulator site suppressed post-tetanic OFC receiver responses (Figure 2E). The OFC sbTetMS modulator did not significantly change the OFC receiver β power ($p = 0.069$, Wilcoxon rank-sum test) or high- γ power ($p = 0.74$) (Figure 2G). The results for this Amyg-sender, OFC-receiver, and OFC-sbTetMS modulator are consistent with the edge modulation mechanisms of neuromodulation.

Figures 2H–2K present an example sbTetMS modulator in a network node different from the receiver. Here we stimulated an OFC sender and observed ACC receiver responses pulse by pulse (Figures 2I and 2J; Figures S6D–S6F). Following sbTetMS delivered to the dorsal CN body (Figure 2H), ACC receiver responses were suppressed (Figure 2I). The CN sbTetMS modulator did not significantly change the ACC receiver β ($p = 0.098$) or high- γ power ($p = 0.45$) (Figure 2K). Hence, this sbTetMS modulator also reveals edge modulation but not node modulation.

We tested 46 detected edge samples by delivering sbTetMS at six sites in nodes of the OFC, CN, Amyg, and BFB (Figure S4; Table S1). These four sbTetMS nodes have been reported recently as potential targets for modulating processing in mood (Rao et al., 2018), value-based decision making (Santacruz et al., 2017), emotion (Inman et al., 2018b), and cognition (Hasselmo and Sarter, 2011). We found that sbTetMS showed clear edge modulation and suppressed excitability across each of the tested edge samples (Table S1). In contrast, sbTetMS modulators did not significantly alter receiver node β power (4 of 46 nodes, $p = 0.079$, binomial test with 5% significance level) or high- γ power (3 of 46 nodes, $p = 0.2$). Therefore, sbTetMS exerts a widespread neuromodulatory influence and acts to specifically suppress neural excitability across the detected edges in the mood processing network, temporarily shutting down multiregional communication.

Decoding Sender-Receiver Communication from Network Excitability

Because sbTetMS modulates neural excitability across the detected network edges, excitability may reflect a network mechanism and involve other brain sites whose activity modulates sender-receiver communication (Figure 3A). This suggests that activity at other network sites (i.e., modulators) may be associated with strengthening or weakening the receiver response to sender synaptic input pulse by pulse. If so, decoding sender-receiver communication may be possible by decoding modulator neural activity to predict receiver responses and, by doing so, reveal how the modulator alters sender-receiver excitability-based communication.

Figures 3B–3D present example OFC→CN communication. We detected the response to each pre-tetanic pulse (Figure S3). Impressively, along with “hit” events, we observed “miss” events when the receiver did not respond to sender stimulation (STAR Methods; Figures 3C and 3D). Of all edge samples detected, the missing responses matched the failure of

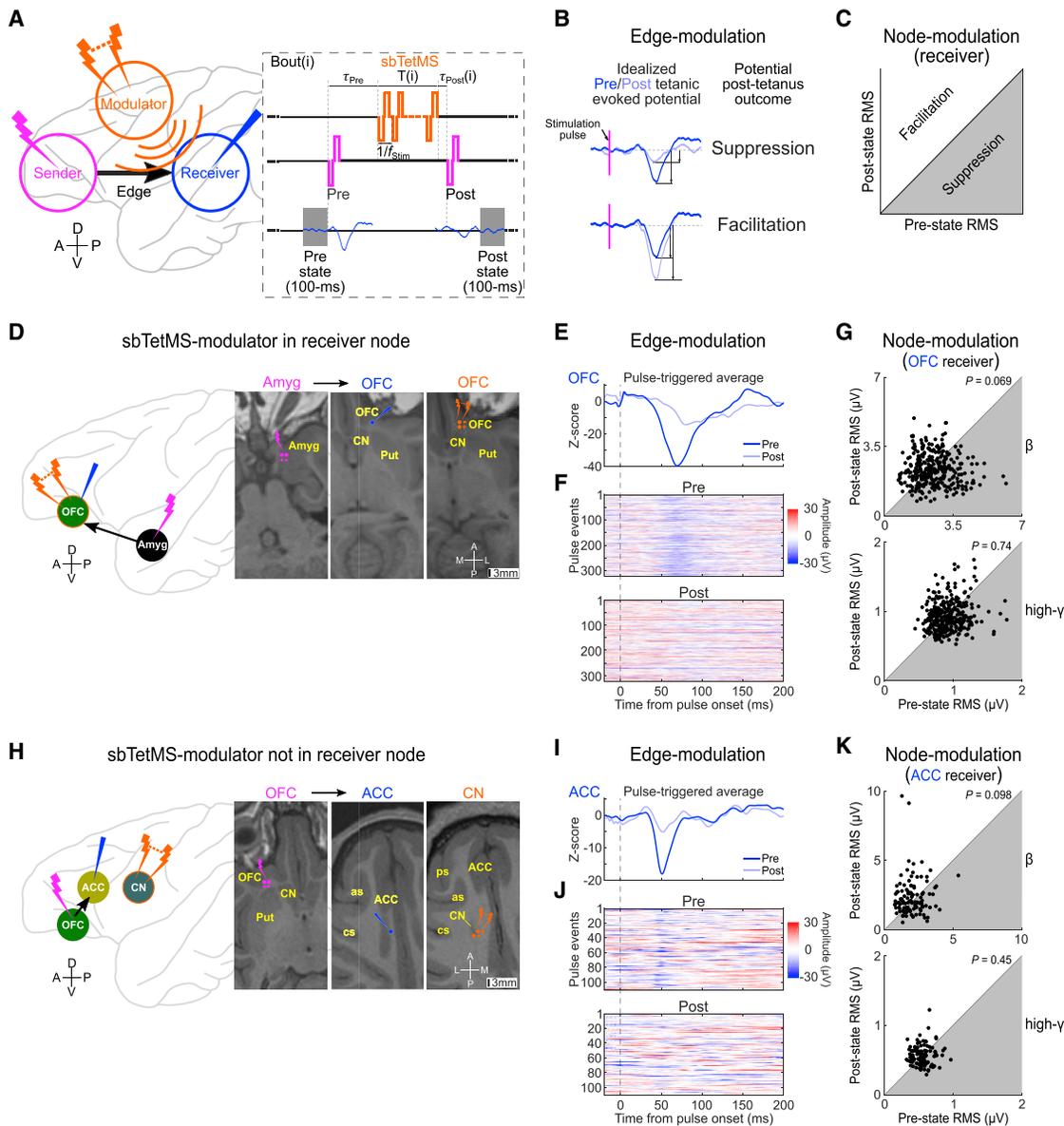


Figure 2. sbTetMS Suppresses Sender-Receiver Excitability

(A) Illustration of how the sbTetMS modulator alters edge weight and receiver node activity. For each stimulation bout, τ_{Pre} is the post Pre-tetanic pulse epoch before the sbTetMS onset, T is the duration of sbTetMS with a given frequency (f_{stim}), and τ_{Post} is the latency between sbTetMS offset and post-tetanic single pulse onset. The gray-shaded rectangles are the 100-ms epochs for computing pre-state RMS and post-state RMS of the neural activity recorded at the receiver, respectively. RMS, root mean square.

(B) Illustration of testing the edge modulation hypothesis by comparing pre-tetanic and post-tetanic evoked potentials.

(C) Illustration of testing the receiver node modulation hypothesis by comparing the RMS power of pre-state and post-state neural activity at the receiver.

(D) The OFC sbTetMS modulator ($T = 50\text{--}500$ ms, $f_{stim} = 100$ Hz, $30 \mu\text{A}$, $100 \mu\text{s}/\text{phase}$; orange dots) suppressed Amyg \rightarrow OFC excitability, as revealed by the evoked OFC receiver (blue dot) response to single pulses ($40 \mu\text{A}$, $100 \mu\text{s}/\text{phase}$; $\tau_{Pre} = 500$ ms and $\tau_{Post} = 25\text{--}100$ ms) delivered at an Amyg sender (magenta dots).

(E) Z-scored pulse-triggered average OFC receiver responses ($n = 325$) to pre-tetanic and post-tetanic single pulses (Amyg \rightarrow OFC).

(F) Evoked OFC receiver responses pulse by pulse during pre-tetanic and post-tetanic epochs.

(G) Scatterplots of pre-state RMS versus post-state RMS of OFC receiver β and high- γ neural activity.

(H) The CN sbTetMS modulator ($T = 250$ ms, $f_{stim} = 200$ Hz, $20 \mu\text{A}$, $100 \mu\text{s}/\text{phase}$; orange dots) suppressed OFC \rightarrow ACC excitability, as revealed by the evoked ACC receiver (blue dot) response to single pulses ($30 \mu\text{A}$, $100 \mu\text{s}/\text{phase}$; $\tau_{Pre} = 500$ ms and $\tau_{Post} = 0, 50$ ms) delivered at an OFC sender (magenta dots). as, arcuate sulcus; cs, central sulcus.

(I) Z-scored pulse-triggered average ACC receiver responses ($n = 110$) to pre-tetanic and post-tetanic single pulses (OFC \rightarrow ACC).

(J) Evoked ACC receiver responses pulse by pulse during pre- and post-tetanic epochs.

(K) Scatterplots of pre-state RMS versus post-state RMS of ACC receiver β and high- γ neural activity.

A, P, L, M, D, and V directions are shown. See also [Figure S6](#).

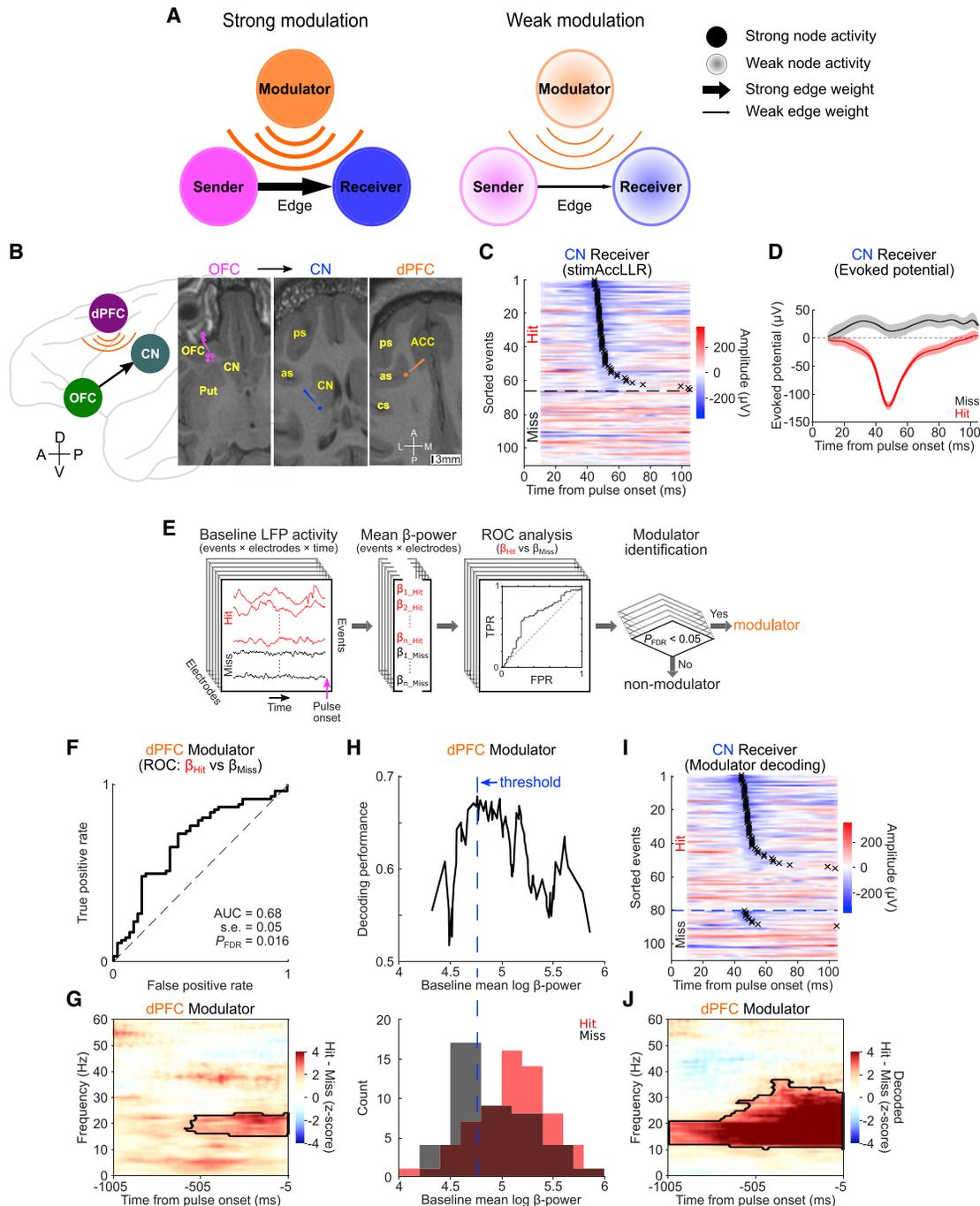


Figure 3. Modulator Activity Covaries with Sender-Receiver Excitability

(A) Schematic of strong and weak modulation of sender-receiver excitability by modulator sites.

(B) OFC → CN communication modulated by the dPFC modulator.

(C) Evoked CN receiver responses sorted by pulse-response latency (crosses). 60% hit events (66 of 110), mean latency = 54 ms.

(D) Pulse-triggered average evoked responses for hit (red) and miss events (black). Shaded, ± 1 SEM.

(E) Schematic of the modulator decoding algorithm using baseline β -activity (10–40 Hz). TPR, true positive rate; FPR, false positive rate; P_{FDR} , p value corrected by false discovery rate analysis for multiple comparisons.

(F) ROC analysis for decoding CN receiver responses from the dPFC modulator (baseline mean log β -power; area under the ROC curve [AUC] = 0.68, SE = 0.05, $P_{FDR} = 0.016$). Black dashed line, chance level.

(G) Modulation spectrograms of dPFC modulator baseline LFP activity, computed as z-score magnitude of power difference between ‘Hit’ and ‘Miss’ events in CN receiver responses. Significance contour from permutation test (cluster-corrected; $n = 10,000$, $p < 0.05$, two-tailed test).

(legend continued on next page)

sender-receiver communication following sbTetMS at sbTetMS modulator sites. Multiunit activity, when observable, was also strongly driven in hit events, but not in miss events (Figure S7). Unsupervised receiver response clustering provided convergent evidence for hit and miss events (STAR Methods; Figure S8), consistent with the action of a modulatory network mechanism.

To identify the modulatory network mechanism, we built a neural decoder to decode sender-receiver communication. We characterized how decoded modulator neural activity predicted hit and miss events from simultaneously recorded baseline neural activity using a receiver-operating characteristic (ROC) analysis (Figure 3E; STAR Methods). We focused on decoding activity in the β -frequency band between 10–40 Hz (Figures 3F and 3G) because it reflects large-scale brain network mechanisms (Pesaran et al., 2008; Siegel et al., 2012; Von Stein and Sarnthein, 2000; Uhlhaas and Singer, 2006). Figure 3G shows that increased dPFC modulator β -power significantly predicted stronger OFC→CN communication. We successfully decoded OFC→CN communication from the baseline β -power of the dPFC modulator activity (p value corrected by false discovery rate analysis for multiple comparisons [P_{FDR}] = 0.016) with 68% performance (Figures 3H–3J). We also showed, as a control, that the baseline β -power of the CN receiver itself did not predict OFC→CN communication (P_{FDR} = 0.62; Figures S9A and S9B). Moreover, by conducting the modulator analysis on the Pulse Null activity for CN receiver, we found that the dPFC modulator baseline β -power did not predict the false alarms for CN receiver baseline activity either (P_{FDR} = 0.67; Figures S9C and S9D). Taken together, our results suggest that the effective opening and closing of the OFC→CN communication channel was not simply due to intrinsic fluctuations of the CN receiver itself but reflected dPFC modulator baseline β -activity. Of 110 detected sender-receiver pairs, we identified 92 significant modulator sites of a total of 1,663 sites from 31 sender-receiver pairs across 14 network edges (Figure 4A; Tables S1, S3, and S4). Thus, neural excitability can be decoded from neural activity at modulator sites, consistent with network excitability mechanism for sender-receiver communication.

Sender-Receiver-Modulator Network Motifs

The neural decoding results suggest that modulation of multiregional communication is dynamically orchestrated by networks of neurons. If so, then network modulation motifs may exist with different mechanistic implications (Figure 4B). If the mechanism involves receiver input-specific modulation, then, when a sender projects to multiple receivers, modulator activity should only covary with excitability for a specific sender-receiver edge (receiver-specific modulator [RSM] motif). In contrast, if the modulation mechanism is shared across multiple projections, then modulator activity should covary with excitability across

multiple sender-receiver edges (receiver network modulator [RNM] motif). Similarly, when multiple senders project to a single receiver, a sender-specific modulator (SSM) should only covary with excitability across a specific sender-receiver edge (SSM motif), whereas mechanisms of modulation shared across multiple projections should be revealed by a sender network modulator (SNM) motif.

Of 92 modulator sites identified, we observed receiver-based modulators but no sender-based modulators (69 RSM and 23 RNM motifs; Tables S3 and S4). Figure 4C presents an RSM motif example. Increased β -power of M1 modulator baseline activity significantly predicted hitting responses at a CN receiver from an OFC sender (OFC→CN), whereas the baseline β -power of the same M1 modulator did not modulate excitability of a dPFC receiver from the same OFC sender (OFC→ACC). Figure 4D presents an RNM motif example. Increased β -power of Put modulator baseline activity significantly predicted hit responses at an ACC receiver from a CN sender (CN→ACC) as well as hit responses at a CN receiver from the same CN sender (CN→CN). Thus, sender-receiver excitability-based communication can be modulated at the receiver in an input-specific manner and as part of a distributed network.

Combining Modulator Decoding with Modulator sbTetMS Modulation

Our data show that sbTetMS modulators and decoded modulators were reflected temporary suppression of sender-receiver communication. We next sought to demonstrate whether targeting sbTetMS modulation specifically at the decoded modulator whose baseline β -power predicted communication suppresses sender-receiver communication (Figure 5A). Figures 5B–5D present an identified BFB modulator whose baseline β -power predicted communication across a CN→ACC pathway. We then targeted sbTetMS at the decoded BFB modulator. Communication across the CN→ACC pathway was temporarily disrupted (Figures 5E–5G). We also successfully suppressed communication across the Amyg→CN pathway by targeting sbTetMS at a decoded OFC modulator (Table S1). These findings indicate that sbTetMS modulation targeted to modulator sites identified by modulator decoding can temporarily disrupt multiregional communication.

DISCUSSION

Here we present a large-scale causal network analysis to decode multiregional communication based on network excitability and identify how sbTetMS alters communication across the pathways connecting the OFC, dPFC, ACC, CN, Put, Amyg, and BFB within the mood processing network. Our work shows how to go beyond focal electrical stimulation

(H) Performance of decoding CN receiver responses as a function of dPFC modulator baseline mean log β -power. Blue vertical dashed line, decoding threshold.

(I) Waveforms classified into decoded hit and miss categories at the CN receiver from the dPFC modulator baseline β -activity. In each category, waveforms were sorted by pulse-response latency (crosses).

(J) Modulation spectrograms of dPFC modulator baseline LFP activity, computed as Z score magnitude of power difference between decoded hit and miss events in CN receiver responses. Also shown is the significance contour from the permutation test (cluster-corrected; n = 10,000, p < 0.05, two-tailed test).

A, P, L, M, D, and V directions are shown. See also Figures S3 and S7–S9.

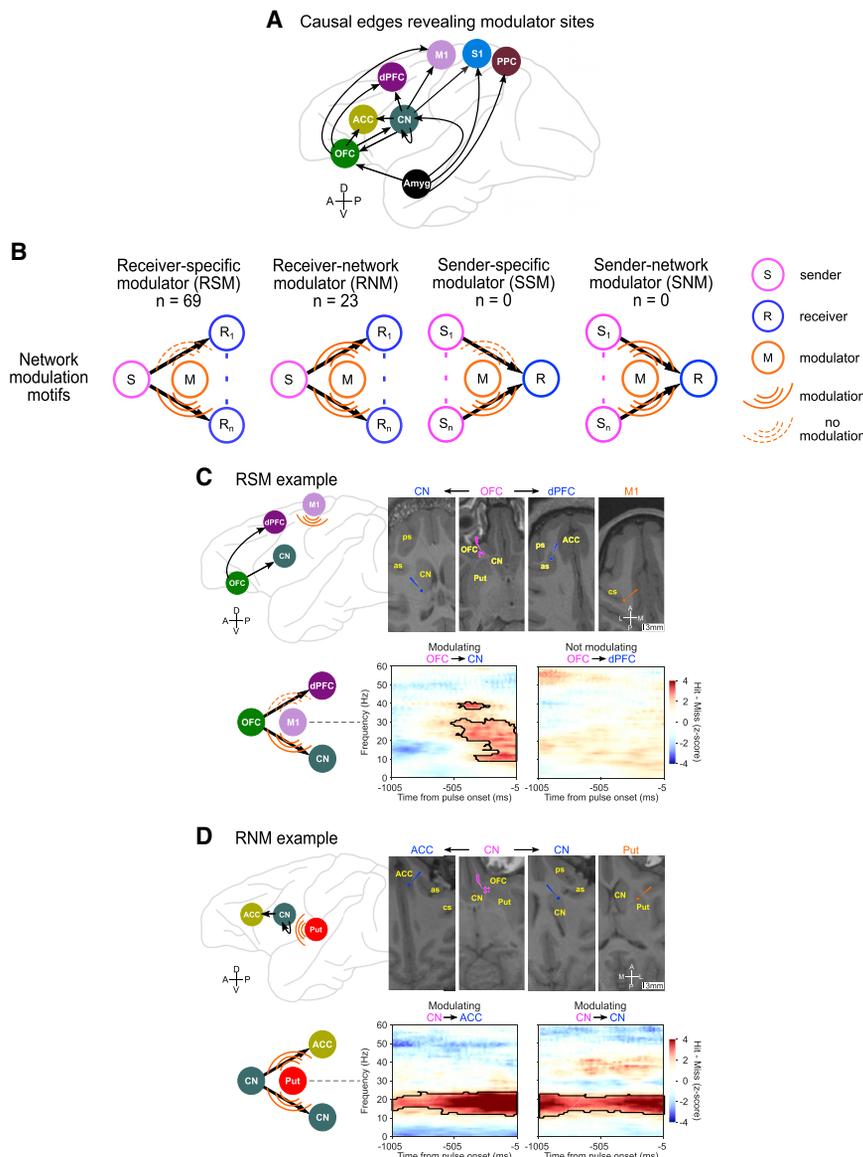


Figure 4. Network Modulation Motifs

(A) Causal network edges that revealed modulator sites modulating the sender-receiver communications.

(B) RSM, RNM, SSM, and SNM network modulation motifs.

(C) RSM example: β power of the M1 modulator site predicted OFC \rightarrow CN excitability (AUC = 0.64, SE = 0.052, $P_{FDR} = 0.023$) but not OFC \rightarrow dPFC excitability (AUC = 0.52, SE = 0.08, $P_{FDR} = 0.53$). Modulation spectrogram contours show significance (permutation test, cluster-corrected; $n = 10,000$, $p < 0.05$, two-tailed test).

(D) RNM example: β -power of the Put modulator predicted CN \rightarrow ACC excitability (AUC = 0.59, SE = 0.032, $P_{FDR} = 0.012$) and CN \rightarrow CN excitability (AUC = 0.58, SE = 0.032, $P_{FDR} = 0.031$). Modulation spectrogram contours show significance (permutation test, cluster-corrected; $n = 10,000$, $p < 0.05$, two-tailed test).

A, P, L, M, D, and V directions are shown. See also [Tables S3 and S4](#).

excitability-based BMIs can be developed to treat disordered brain networks.

We first discuss the implications of our work for mood regulation and how modulator decoding can inform BMI technology, and then discuss the broader mechanistic implications for multiregional communication.

Regulating the Mood Processing Network

The biology of mood processing as well as other emotional and cognitive functions emphasizes the heterogeneous nature of the underlying brain mechanisms. Work in human subjects in particular has emphasized the heterogeneity of mood processing with different biological mechanisms, giving rise to overlapping,

([Dougherty et al., 2015](#); [Inman et al., 2018b](#); [Mayberg et al., 2005](#); [Rao et al., 2018](#)) by analyzing multiregional communication in brain networks. We show that multiregional communication involves a mechanism of network excitability that modulates how the receiver responds to input and is suppressed by sbTetMS. Demonstrating that network modulators can predict multiregional communication moment by moment allows us to decode network excitability from modulator sites. Our finding that neuromodulation perturbs network mechanisms regulating information flow between brain regions has major implications for understanding and treating functional brain networks. By showing that neural decoders can track variations in network excitability that underlie multiregional communication and that neuromodulation specifically alters sender-receiver communication, we develop a novel connection between neuromodulation and BMI technologies. We propose that next-generation stimulation-recording therapies featuring

heterogeneous clinical presentations ([Clementz et al., 2016](#); [Drysdale et al., 2017](#); [Hill et al., 2013](#)). Co-occurrence of clinical symptoms has been used to demonstrate the existence of distinct biological phenotypes. Neuroimaging evidence also points to the existence of different underlying biotypes that can be resolved with high sensitivity and specificity based on shared signatures in resting-state networks ([Drysdale et al., 2017](#)). Mood processing depends on a distributed network whose interconnections maintain functional organization across multiple cortico-cortical, cortico-subcortical, and subcortico-subcortical pathways ([Haber, 2016](#)). These analyses provide biomarkers that reveal different subtypes of depression and predict responses to therapeutic transcranial magnetic stimulation ([Drysdale et al., 2017](#)).

Thus, mood function and dysfunction are understood as being due to expression of multiple overlapping network mechanisms. Despite this mechanistic emphasis, functional interactions

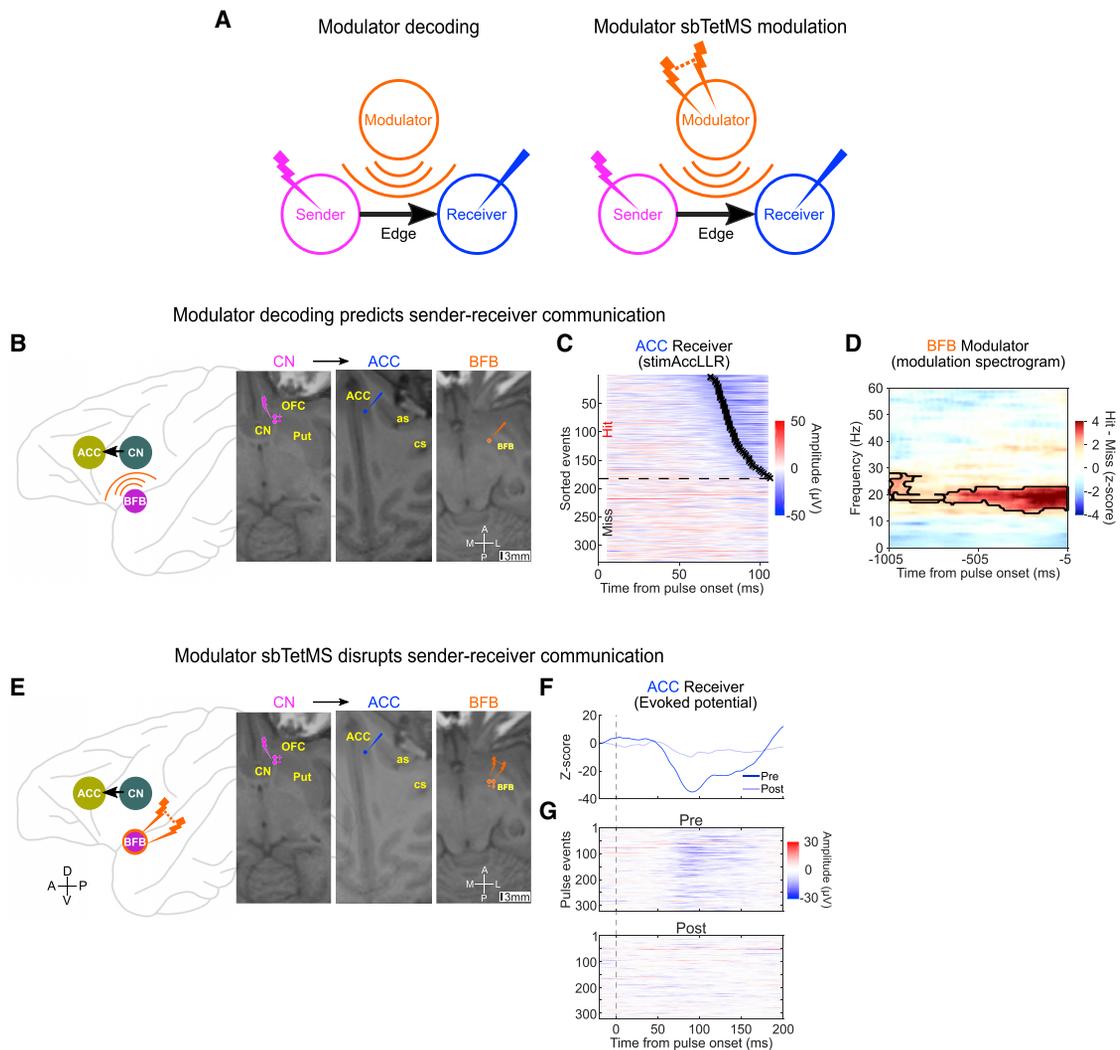


Figure 5. Combining Modulator Decoding with Modulator sbTetMS Modulation

(A) Schematic of decoding sender-receiver communication to identify the modulator site targeted for sbTetMS modulation to alter the communication.

(B) CN→ACC communication modulated by BFB modulator activity.

(C) Evoked ACC receiver responses sorted by pulse-response latency (crosses). 55% hit events (182 of 330), mean latency = 82 ms.

(D) β -Power of the BFB modulator predicted CN→ACC excitability (AUC = 0.58, SE = 0.031, P_{FDR} = 0.017). Shown are modulation spectrograms of BFB modulator baseline LFP activity before the pre-tetanic pulse, computed as Z score magnitude of power difference between hit and miss events in ACC receiver responses. Also shown is the significance contour from the permutation test (cluster-corrected; n = 10,000, p < 0.05, two-tailed test).

(E) The BFB sbTetMS modulator (T = 50–500 ms, f_{stim} = 100 Hz, 30 μ A, 100 μ s/phase; orange dots) suppressed CN→ACC excitability, as revealed by the evoked ACC receiver (blue dot) response to single pulses (60 μ A, 100 μ s/phase; τ_{Pre} = 500 ms and τ_{Post} = 25–100 ms) delivered at a CN sender (magenta dots).

(F) Z-scored pulse-triggered average ACC receiver responses (n = 330) to pre-tetanic and post-tetanic single pulses (CN→ACC).

(G) Evoked ACC receiver responses pulse by pulse during pre-tetanic and post-tetanic epochs.

A, P, L, M, D, and V directions are shown.

underlying mood regulation have been revealed by analyses of pairwise blood oxygen level-dependent (BOLD) functional magnetic resonance imaging (fMRI) signal correlations between brain regions (Drysdale et al., 2017; Tang et al., 2019). However, such functional connectivity analysis lacks spatiotemporal resolution and is relatively indirect. The causal network analysis we report builds on this body of work and provides important mechanistic insights into mood processing. Using more direct electrophysiological evidence from across the frontostriatal limbic network,

we demonstrate that multiregional communication in the mood processing network is subject to a powerful form of modulatory regulation. The mechanism of modulation is rapid and precise and features a mechanism of network excitability modulated at the receiver. The receiver node responds to inputs from the sender node when the modulator node increases receiver excitability. As a result, our work suggests that the heterogeneous nature of mood processing may reflect the action of multiple modulator networks that are necessary for selective communication

between sender and receiver sites processing mood information.

Modulator Decoding for Neuromodulation

BMI technology has successfully restored lost sensory and motor functions by recording from and stimulating neural circuits (Bensmaia and Miller, 2014; Shanechi, 2019; Shenoy and Carmena, 2014). Our results can inform next-generation brain therapies for a wider range of functional goals by using BMI devices to sense the modulator network state and deliver targeted neuromodulation. The inconsistent therapeutic efficacy from prior studies using open-loop focal neuromodulation to treat mood disorders (Dougherty et al., 2015; Inman et al., 2018b; Mayberg et al., 2005; Rao et al., 2018; Widge et al., 2019) may reflect the state dependence of the underlying network mechanisms. Such concerns highlight the need to develop next-generation BMIs that can deliver personalized closed-loop neuromodulation and selectively manipulate multiregional communication and regulate mood. By decoding the specific state of modulator networks associated with disordered multiregional communication (Braun et al., 2018; McCormick et al., 2020; Shanechi, 2019), our results offer a new strategy for restoring disordered multiregional communication in a novel BMI that decodes the state of modulator networks to guide targeted neuromodulation-based therapies.

A key feature of our results is how modulators in one network can also modulate the excitability of other networks (e.g., RNM motifs). Interactions between the different networks integrating sensation, reward, and action to guide motivated behavior may depend on such RNM motifs. These findings suggest novel strategies in which neural decoders to track and predict the state of modulator networks direct neuromodulation-based therapies that can target interventions, control networks, and identify new biomarkers. Neuromodulatory interventions can be targeted in a two-stage approach. A causal network analysis could identify modulators for specific networks that are then targeted for intervention by a neuromodulation system in an open-loop approach. Specifically, if we can decode the state of modulator networks associated with specific clinical symptom profiles of mood disorders (Drysdale et al., 2017), then we may target sbTetMS at the modulator site(s) to effectively alter network excitability of disordered multiregional communication. Over the course of treatment, subject-specific adjustments in how sbTetMS is spatiotemporally optimized (Choi et al., 2016; Khambhati et al., 2019; Muldoon et al., 2016) as modulator networks reconfigure over time may improve treatment efficacy. Causal network analyses may also suggest development of novel mechanistically informed biotypes of neuropsychiatric conditions. Modulator-based biotypes may explain the resting-state network clusters as well as improve their validity and therapeutic efficacy. Modulator decoding could also inform closed-loop neuromodulatory strategies, by triggering neuromodulation when the state of modulator networks changes, or by using feedback control approaches to control modulator network activity directly to reach a target state. Finally, mathematical models are an essential ingredient in BMI technologies, and the causal network analysis we present naturally complements mathematical models, such as state-space models (Kao et al., 2015; Sani

et al., 2018; Yang et al., 2019) and multilayer network models (Bassett and Sporns, 2017; Kivela et al., 2014), and can inform their development.

Network Mechanisms of Modulating Multiregional Communication

Synthesizing evidence from the sender-receiver-modulator motifs we obtained from modulator decoding reveals that modulation of multiregional communication in the sampled mood processing network is widespread and distributed across multiple cortical and subcortical regions. This finding is likely due to how indirect pathways involving cortico-thalamic-cortical and basal ganglia-thalamic projections link cortical to subcortical regions (Haber, 2016). Modulator activity, especially within the β -frequency band, reveals a form of network gating that modulates sender-receiver communication, suppressing directed functional interactions momentarily in time. In particular, we found a consistent example of an RSM motif in both monkeys, where the OFC modulator β -activity gated OFC \rightarrow ACC communication. These results highlight a prominent role of β -activity dynamics in modulating convergence and divergence of cortico-subcortical communication within the mood pathways. The sender-receiver-modulator motifs we observed also support a distributed and recurrent framework (Hunt and Hayden, 2017) for studying value computation in the mood processing network. Further examining such motifs in behavior is needed to understand how decision signals entering particular sender-receiver communication channels can be selectively filtered by other network nodes in the circuit to direct goal-oriented behavior.

Frequency-specific neuronal synchronization offers a putative network mechanism for multiregional communication (Fries, 2015; Pesaran et al., 2008; Siegel et al., 2012). Estimating neural coherence between brain regions has revealed “bottom-up” and “top-down” control for communication across distributed decision-making and mood processing networks. Examples include prefrontal-Amyg interactions modulated by the β - and γ -coherence for social decision preferences (Dal Monte et al., 2020), enhanced frontal-parietal interaction modulated by the β -coherence during free choices in decision-making (Pesaran et al., 2008), and an Amyg-hippocampus interaction modulated by the β -coherence for encoding mood variations (Kirkby et al., 2018). Similar mechanisms have also been observed in the visual attention network (Bastos et al., 2015).

Rhythmic synchronization-based mechanisms of multiregional communication are most often inferred from the structure of correlations in neural activity and from examining responses to large-amplitude and/or trains of stimulation pulses. Using correlations to interpret activity patterns in a receiver as being due to interactions with a sender is subject to significant confounds. Correlations are sensitive to the confounding influence of common inputs from other brain regions, yielding network edges even when the receiver does not receive any input from the sender (Pesaran et al., 2018). Although correlation-based sampling can be used to infer the strength of directed functional interactions (Keller et al., 2011; Khambhati et al., 2019; Solomon et al., 2018), taking a causal sampling approach offers the important advantage of making a more direct measure of functional interactions. Causal responses cannot be due to common input,

which is consistent with their sparse presence in our data. Inferences from large-amplitude stimulation pulses or pulse trains suffer other confounds. Large-amplitude stimulation may recruit network responses because of other mechanisms (Herrington et al., 2016; Lozano et al., 2019), changing the interaction instead of measuring the interaction. Delivering isolated low-amplitude microstimulation pulses offer the opportunity to more directly probe network excitability while avoiding the confounding effects and network responses.

Mechanistic Insights and Network Excitability

Our work provides additional mechanistic insights. Our sbTetMS edge modulation results reinforce prior work showing that high-frequency stimulation can disrupt or block synaptic transmission from the sender to receiver (Grill et al., 2004; Lozano et al., 2019). Such effects are said to induce “information lesions,” highlighting the potential for sbTetMS as a tool to temporarily suppress network excitability. By grouping the network nodes into cortical and subcortical nodes, we found consistent sbTetMS modulation motifs in both monkeys, where sbTetMS at a subcortical node could suppress both cortical→subcortical and cortical→cortical communication. Specifically, we observed that applying sbTetMS to the CN or Amyg node could suppress OFC→CN and OFC→OFC excitability. Our finding supports previous lesion and stimulation studies showing that the CN and Amyg play a causal role in stimulus-action-reward encoding by modulating frontostriatal functional interactions in the reward circuit (Inman et al., 2018b; Rudebeck et al., 2013, 2017; Santacruz et al., 2017; Schultz, 2016).

The fluctuations in sender-receiver excitability we observed suggest orthodromic and *trans*-synaptic stimulation effects (Histed et al., 2009; Tehovnik et al., 2006; Tolias et al., 2005), although we cannot rule out antidromic stimulation effects. The causal network analysis we report, however, extends beyond putatively direct interactions to uncover mechanisms of neuromodulation that are not focal but distributed. The observed β -power suppression at the receiver by the sbTetMS modulator, although not significant across all tested sender-receiver pairs (4 of 46, 7%; Table S1), is consistent with the “synaptic filtering” hypothesis. According to synaptic filtering, a high-frequency stimulation pulse train delivered to the pre-synaptic site induces short-term suppression that selectively suppresses the synaptic transmission of low-frequency content contained in stimulation-induced neural activity at the post-synaptic site (Farokhniaee and McIntyre, 2019; Rosenbaum et al., 2014). However, our experiments did not constrain the mechanism by which the modulator sites modulate neural excitability. For example, we were not able to test whether single-pulse stimulation of modulator sites drove responses in the associated receiver nodes. Therefore, how modulators interact with the receiver and can specifically modulate pre-synaptic inputs from the sender remains open.

Changes in sender-receiver excitability likely reflect additional mechanisms within the receiver node. (1) sbTetMS may change the excitatory-inhibitory balance in the receiver (Froemke, 2015), reducing excitatory and increasing inhibitory potentials, consistent with reductions in firing rate as well as the pulse-evoked LFP responses we observed. (2) sbTetMS may also induce spatial reorganization of the pre-synaptic inputs that changes the

strength of the effective dipole moment generating the observed LFP activity (Mazzoni et al., 2015). (3) The absence of responses following sbTetMS we observed might also reflect short-term depression in the post-synaptic neurons (Froemke, 2015). (4) Momentary failures at the pre- and post-synaptic junction may also contribute to the gating of sender synaptic input to the receiver. Nevertheless, the predictability of neural excitability by modulators cannot simply be explained by mechanisms within the receiver node. The modulator decoding results provide strong evidence of a network excitability-based mechanism for communication.

Prior studies have examined neural excitability by applying primary and secondary interventions at the sender only, perhaps reflecting technical limitations (Keller et al., 2018; Rao et al., 2018). As a result, edge- and node-based mechanisms could not be dissociated. With sender-only interventions, changes in receiver node responses may reflect changes in pre-synaptic activity recruited by the primary intervention or the priming effect of the secondary intervention on the sender (Siebner et al., 2004). Using our large-scale multisite approach, we disambiguated these effects and investigated dynamic changes in excitability across edges connecting different senders and receivers. Critically, we could do so independent of node activity at the receiver to specifically conclude that edge-based responses were disrupted.

Conclusions

In conclusion, we show that performing a minimally perturbative causal network analysis reveals specific network mechanisms of multiregional communication and directly relates sbTetMS neuromodulation to a general mechanism of network excitability that supports multiregional communication. These results reveal multiregional communication within the cortico-subcortical limbic mood processing network and point to how next-generation brain therapies can combine neuromodulation with neural decoding in BMIs to correct disordered multiregional communication. Finally, our causal approach may be generalized to uncover network mechanisms of action for other neuromodulation modalities.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Large-scale microdrive system design and assembly
 - Magnetic resonance imaging and processing
 - Experimental preparation
 - Neural recordings
 - Sampling network nodes and edges from recording
 - Sampling causal network edges from simultaneous microstimulation and recording

● **QUANTIFICATION AND STATISTICAL ANALYSIS**

- Pulse-triggered evoked potentials
- Pulse-triggered multiunit activity
- Detection of pulse-evoked LFP responses using the stimAccLLR method
- Decoding sender-receiver communication to reveal modulators
- Modulation spectrogram of modulator baseline activity

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.neuron.2020.06.012>.

ACKNOWLEDGMENTS

We would like to thank Baldwin Goodell, Charles Gray, Jessica Kleinbart, and Amy Orsborn for assistance with chamber and microdrive system design; Stephen Frey and Brian Hynes for custom modifications to the Brainsight system; Keith Sanzenbach and Pablo Velasco from the NYU Center for Imaging for help with magnetic resonance imaging and diffusion weighted imaging; and Ryan Shewcraft, John Choi, Marsela Rubiano, Yoohee Jang, Octavia Martin, and the NYU Office of Veterinary Resources for help with animal preparation and care. This work was supported in part by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement W911NF-14-2-0043, issued by the Army Research Office contracting office in support of DARPA's SUBNETS program (to B.P.). The views, opinions, and/or findings expressed are those of the author(s) and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government. This work was also supported in part by an award from the Simons Collaboration on the Global Brain (to B.P.) and National Institutes of Health (NIH) BRAIN grant R01-NS104923 (to B.P.).

AUTHOR CONTRIBUTIONS

Conceptualization, S.Q. and B.P.; Methodology, S.Q. and B.P.; Investigation, S.Q., J.I.S., K.A.B., B.F., and B.P.; Formal Analysis, S.Q. and B.P.; Writing – Original Draft, S.Q. and B.P.; Writing – Review & Editing, S.Q. and B.P.; Funding Acquisition, B.P.; Supervision, B.P.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 10, 2019

Revised: April 21, 2020

Accepted: June 9, 2020

Published: July 8, 2020

REFERENCES

Andersson, J.L.R., and Sotiropoulos, S.N. (2015). Non-parametric representation and prediction of single- and multi-shell diffusion-weighted MRI data using Gaussian processes. *Neuroimage* *122*, 166–176.

Banerjee, A., Dean, H.L., and Pesaran, B. (2010). A likelihood method for computing selection times in spiking and local field potential activity. *J. Neurophysiol.* *104*, 3705–3720.

Bassett, D.S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* *20*, 353–364.

Bassett, D.S., Zurn, P., and Gold, J.I. (2018). On the nature and use of models in network neuroscience. *Nat. Rev. Neurosci.* *19*, 566–578.

Bastos, A.M., Vezoli, J., Bosman, C.A., Schoffelen, J.M., Oostenveld, R., Dowdall, J.R., De Weerd, P., Kennedy, H., and Fries, P. (2015). Visual areas exert feedforward and feedback influences through distinct frequency channels. *Neuron* *85*, 390–401.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B Stat. Methodol.* *57*, 289–300.

Bensmaia, S.J., and Miller, L.E. (2014). Restoring sensorimotor function through intracortical interfaces: progress and looming challenges. *Nat. Rev. Neurosci.* *15*, 313–325.

Borchers, S., Himmelbach, M., Logothetis, N., and Karnath, H.-O. (2011). Direct electrical stimulation of human cortex - the gold standard for mapping brain functions? *Nat. Rev. Neurosci.* *13*, 63–70.

Braun, U., Schaefer, A., Betzel, R.F., Tost, H., Meyer-Lindenberg, A., and Bassett, D.S. (2018). From maps to multi-dimensional network mechanisms of mental disorders. *Neuron* *97*, 14–31.

Butovas, S., and Schwarz, C. (2003). Spatiotemporal effects of microstimulation in rat neocortex: a parametric study using multielectrode recordings. *J. Neurophysiol.* *90*, 3024–3039.

Calabrese, E., Badea, A., Coe, C.L., Lubach, G.R., Shi, Y., Styner, M.A., and Johnson, G.A. (2015). A diffusion tensor MRI atlas of the postmortem rhesus macaque brain. *Neuroimage* *117*, 408–416.

Castro-Alamancos, M.A., and Connors, B.W. (1996). Spatiotemporal properties of short-term plasticity sensorimotor thalamocortical pathways of the rat. *J. Neurosci.* *16*, 2767–2779.

Choi, J.S., Brockmeier, A.J., McNeil, D.B., Kraus, L.M., Principe, J.C., and Francis, J.T. (2016). Eliciting naturalistic cortical responses with a sensory prosthesis via optimized microstimulation. *J. Neural Eng.* *13*, 056007.

Clementz, B.A., Sweeney, J.A., Hamm, J.P., Ileva, E.I., Ethridge, L.E., Pearson, G.D., Keshavan, M.S., and Tamminga, C.A. (2016). Identification of distinct psychosis biotypes using brain-based biomarkers. *Am. J. Psychiatry* *173*, 373–384.

Dal Monte, O., Chu, C.C.J., Fagan, N.A., and Chang, S.W.C. (2020). Specialized medial prefrontal-amygdala coordination in other-regarding decision preference. *Nat. Neurosci.* *23*, 565–574.

Dotson, N.M., Hoffman, S.J., Goodell, B., and Gray, C.M. (2017). A large-scale semi-chronic microdrive recording system for non-human primates. *Neuron* *96*, 769–782.e2.

Dougherty, D.D., Rezai, A.R., Carpenter, L.L., Howland, R.H., Bhati, M.T., O'Reardon, J.P., Eskandar, E.N., Baltuch, G.H., Machado, A.D., Kondziolka, D., et al. (2015). A randomized sham-controlled trial of deep brain stimulation of the ventral capsule/ventral striatum for chronic treatment-resistant depression. *Biol. Psychiatry* *78*, 240–248.

Drevets, W.C. (2001). Neuroimaging and neuropathological studies of depression: implications for the cognitive-emotional features of mood disorders. *Curr. Opin. Neurobiol.* *11*, 240–249.

Drysdale, A.T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R.N., Zebley, B., Oathes, D.J., Etkin, A., et al. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* *23*, 28–38.

Ezzyat, Y., Kragel, J.E., Burke, J.F., Levy, D.F., Lyalenko, A., Wanda, P., O'Sullivan, L., Hurley, K.B., Busygin, S., Pedisich, I., et al. (2017). Direct brain stimulation modulates encoding states and memory performance in humans. *Curr. Biol.* *27*, 1251–1258.

Ezzyat, Y., Wanda, P.A., Levy, D.F., Kadel, A., Aka, A., Pedisich, I., Sperling, M.R., Sharan, A.D., Lega, B.C., Burks, A., et al. (2018). Closed-loop stimulation of temporal cortex rescues functional networks and improves memory. *Nat. Commun.* *9*, 365.

Farokhniaee, A., and McIntyre, C.C. (2019). Theoretical principles of deep brain stimulation induced synaptic suppression. *Brain Stimul.* *12*, 1402–1409.

Frey, S., Pandya, D.N., Chakravarty, M.M., Bailey, L., Petrides, M., and Collins, D.L. (2011). An MRI based average macaque monkey stereotaxic atlas and space (MNI monkey space). *Neuroimage* *55*, 1435–1442.

Fries, P. (2015). Rhythms for Cognition: Communication through Coherence. *Neuron* *88*, 220–235.

Froemke, R.C. (2015). Plasticity of cortical excitatory-inhibitory balance. *Annu. Rev. Neurosci.* *38*, 195–219.

- Grill, W.M., Snyder, A.N., and Miodinovic, S. (2004). Deep brain stimulation creates an informational lesion of the stimulated nucleus. *Neuroreport* *15*, 1137–1140.
- Haber, S.N. (2016). Anatomy and connectivity of the reward circuit. In *Decision Neuroscience: An Integrative Perspective*, Jean-Claude Dreher and Léon Tremblay, eds. (Academic Press), pp. 3–19.
- Hasselmo, M.E., and Sarter, M. (2011). Modes and models of forebrain cholinergic neuromodulation of cognition. *Neuropsychopharmacology* *36*, 52–73.
- Häusser, M., Major, G., and Stuart, G.J. (2001). Differential shunting of EPSPs by action potentials. *Science* *291*, 138–141.
- Herrington, T.M., Cheng, J.J., and Eskandar, E.N. (2016). Mechanisms of deep brain stimulation. *J. Neurophysiol.* *115*, 19–38.
- Hill, S.K., Reilly, J.L., Keefe, R.S.E., Gold, J.M., Bishop, J.R., Gershon, E.S., Tamminga, C.A., Pearson, G.D., Keshavan, M.S., and Sweeney, J.A. (2013). Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *Am. J. Psychiatry* *170*, 1275–1284.
- Histed, M.H., Bonin, V., and Reid, R.C. (2009). Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron* *63*, 508–522.
- Hunt, L.T., and Hayden, B.Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. *Nat. Rev. Neurosci.* *18*, 172–182.
- Inman, C.S., Manns, J.R., Bijanki, K.R., Bass, D.I., Hamann, S., Drane, D.L., Fasano, R.E., Kovach, C.K., Gross, R.E., and Willie, J.T. (2018a). Direct electrical stimulation of the amygdala enhances declarative memory in humans. *Proc. Natl. Acad. Sci. USA* *115*, 98–103.
- Inman, C.S., Bijanki, K.R., Bass, D.I., Gross, R.E., Hamann, S., and Willie, J.T. (2018b). Human amygdala stimulation effects on emotion physiology and emotional experience. *Neuropsychologia*. Published online March 15, 2018. <https://doi.org/10.1016/j.neuropsychologia.2018.03.019>.
- Jbabdi, S., Sotiropoulos, S.N., Savio, A.M., Graña, M., and Behrens, T.E.J. (2012). Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems. *Magn. Reson. Med.* *68*, 1846–1855.
- Johnson, M.D., Lim, H.H., Netoff, T.I., Connolly, A.T., Johnson, N., Roy, A., Holt, A., Lim, K.O., Carey, J.R., Vitek, J.L., and He, B. (2013). Neuromodulation for brain disorders: challenges and opportunities. *IEEE Trans. Biomed. Eng.* *60*, 610–624.
- Kao, J.C., Nuyujukian, P., Ryu, S.I., Churchland, M.M., Cunningham, J.P., and Shenoy, K.V. (2015). Single-trial dynamics of motor cortex and their applications to brain-machine interfaces. *Nat. Commun.* *6*, 7759.
- Keller, C.J., Bickel, S., Entz, L., Ulbert, I., Milham, M.P., Kelly, C., and Mehta, A.D. (2011). Intrinsic functional architecture predicts electrically evoked responses in the human brain. *Proc. Natl. Acad. Sci. USA* *108*, 10308–10313.
- Keller, C.J., Huang, Y., Herrero, J.L., Fini, M.E., Du, V., Lado, F.A., Honey, C.J., and Mehta, A.D. (2018). Induction and quantification of excitability changes in human cortical networks. *J. Neurosci.* *38*, 5384–5398.
- Khambhati, A.N., Kahn, A.E., Costantini, J., Ezzayat, Y., Solomon, E.A., Gross, R.E., Jobst, B.C., Sheth, S.A., Zaghoul, K.A., Worrell, G., et al. (2019). Functional control of electrophysiological network architecture using direct neurostimulation in humans. *Netw. Neurosci.* *3*, 848–877.
- Kirkby, L.A., Luongo, F.J., Lee, M.B., Nahum, M., Van Vleet, T.M., Rao, V.R., Dawes, H.E., Chang, E.F., and Sohal, V.S. (2018). An amygdala-hippocampus subnetwork that encodes variation in human mood. *Cell* *175*, 1688–1700.e14.
- Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J.P., Moreno, Y., Porter, M.A., and Estrada, E. (2014). Multilayer networks. *J. Complex Netw.* *2*, 203–271.
- Little, S., and Brown, P. (2014). The functional role of beta oscillations in Parkinson's disease. *Parkinsonism Relat. Disord.* *20 (Suppl 1)*, S44–S48.
- Logothetis, N.K., Augath, M., Murayama, Y., Rauch, A., Sultan, F., Goense, J., Oeltermann, A., and Merkle, H. (2010). The effects of electrical microstimulation on cortical signal propagation. *Nat. Neurosci.* *13*, 1283–1291.
- Lozano, A.M., Lipsman, N., Bergman, H., Brown, P., Chabardes, S., Chang, J.W., Matthews, K., McIntyre, C.C., Schlaepfer, T.E., Schulder, M., et al. (2019). Deep brain stimulation: current challenges and future directions. *Nat. Rev. Neurol.* *15*, 148–160.
- Manning, J.R., Jacobs, J., Fried, I., and Kahana, M.J. (2009). Broadband shifts in local field potential power spectra are correlated with single-neuron spiking in humans. *J. Neurosci.* *29*, 13613–13620.
- Maris, E., Schoffelen, J.-M., and Fries, P. (2007). Nonparametric statistical testing of coherence differences. *J. Neurosci. Methods* *163*, 161–175.
- Mayberg, H.S., Lozano, A.M., Voon, V., McNeely, H.E., Seminowicz, D., Hamani, C., Schwab, J.M., and Kennedy, S.H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron* *45*, 651–660.
- Mazzoni, A., Lindén, H., Cuntz, H., Lansner, A., Panzeri, S., and Einevoll, G.T. (2015). Computing the local field potential (LFP) from integrate-and-fire network models. *PLoS Comput. Biol.* *11*, e1004584.
- McCormick, D.A., Nestvogel, D.B., and He, B.J. (2020). Neuromodulation of Brain State and Behavior. *Annu. Rev. Neurosci.* Published online April 6, 2020. <https://doi.org/10.1146/annurev-neuro-100219-105424>.
- Muldoon, S.F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S.T., Vettel, J.M., and Bassett, D.S. (2016). Stimulation-based control of dynamic brain networks. *PLoS Comput. Biol.* *12*, e1005076.
- Nestler, E.J., and Carlezon, W.A., Jr. (2006). The mesolimbic dopamine reward circuit in depression. *Biol. Psychiatry* *59*, 1151–1159.
- Pesaran, B., Nelson, M.J., and Andersen, R.A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature* *453*, 406–409.
- Pesaran, B., Vinck, M., Einevoll, G.T., Sirota, A., Fries, P., Siegel, M., Truccolo, W., Schroeder, C.E., and Srinivasan, R. (2018). Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. *Nat. Neurosci.* *21*, 903–919.
- Price, J.L., and Drevets, W.C. (2010). Neurocircuitry of mood disorders. *Neuropsychopharmacology* *35*, 192–216.
- Rajasethupathy, P., Ferenczi, E., and Deisseroth, K. (2016). Targeting neural circuits. *Cell* *165*, 524–534.
- Rao, V.R., Sellers, K.K., Wallace, D.L., Lee, M.B., Bijanzadeh, M., Sani, O.G., Yang, Y., Shanechi, M.M., Dawes, H.E., and Chang, E.F. (2018). Direct electrical stimulation of lateral orbitofrontal cortex acutely improves mood in individuals with symptoms of depression. *Curr. Biol.* *28*, 3893–3902.e4.
- Rosenbaum, R., Zimnik, A., Zheng, F., Turner, R.S., Alzheimer, C., Doiron, B., and Rubin, J.E. (2014). Axonal and synaptic failure suppress the transfer of firing rate oscillations, synchrony and information during high frequency deep brain stimulation. *Neurobiol. Dis.* *62*, 86–99.
- Rudebeck, P.H., Mitz, A.R., Chacko, R.V., and Murray, E.A. (2013). Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron* *80*, 1519–1531.
- Rudebeck, P.H., Ripple, J.A., Mitz, A.R., Averbeck, B.B., and Murray, E.A. (2017). Amygdala Contributions to Stimulus-Reward Encoding in the Macaque Medial and Orbital Frontal Cortex during Learning. *J. Neurosci.* *37*, 2186–2202.
- Sani, O.G., Yang, Y., Lee, M.B., Dawes, H.E., Chang, E.F., and Shanechi, M.M. (2018). Mood variations decoded from multi-site intracranial human brain activity. *Nat. Biotechnol.* *36*, 954–961.
- Santacruz, S.R., Rich, E.L., Wallis, J.D., and Carmena, J.M. (2017). Caudate Microstimulation Increases Value of Specific Choices. *Curr. Biol.* *27*, 3375–3383.e3.
- Schultz, W. (2016). Reward functions of the basal ganglia. *J. Neural Transm. (Vienna)* *123*, 679–693.
- Seidemann, E., Arieli, A., Grinvald, A., and Slovin, H. (2002). Dynamics of depolarization and hyperpolarization in the frontal cortex and saccade goal. *Science* *295*, 862–865.
- Shanechi, M.M. (2019). Brain-machine interfaces from motor to mood. *Nat. Neurosci.* *22*, 1554–1564.
- Shenoy, K.V., and Carmena, J.M. (2014). Combining decoder design and neural adaptation in brain-machine interfaces. *Neuron* *84*, 665–680.

- Shewcraft, R.A., Dean, H.L., Fabiszak, M.M., Hagan, M.A., Wong, Y.T., and Pesaran, B. (2020). Excitatory/inhibitory responses shape coherent neuronal dynamics driven by optogenetic stimulation in the primate brain. *J. Neurosci.* *40*, 2056–2068.
- Siebner, H.R., Lang, N., Rizzo, V., Nitsche, M.A., Paulus, W., Lemon, R.N., and Rothwell, J.C. (2004). Preconditioning of low-frequency repetitive transcranial magnetic stimulation with transcranial direct current stimulation: evidence for homeostatic plasticity in the human motor cortex. *J. Neurosci.* *24*, 3379–3385.
- Siegel, M., Donner, T.H., and Engel, A.K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* *13*, 121–134.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* *23* (Suppl 1), S208–S219.
- Solomon, E.A., Kragel, J.E., Gross, R., Lega, B., Sperling, M.R., Worrell, G., Sheth, S.A., Zaghoul, K.A., Jobst, B.C., Stein, J.M., et al. (2018). Medial temporal lobe functional connectivity predicts stimulation-induced theta power. *Nat. Commun.* *9*, 4437.
- Suthana, N., Haneef, Z., Stern, J., Mukamel, R., Behnke, E., Knowlton, B., and Fried, I. (2012). Memory enhancement and deep-brain stimulation of the entorhinal area. *N. Engl. J. Med.* *366*, 502–510.
- Tang, E., Mattar, M.G., Giusti, C., Lydon-Staley, D.M., Thompson-Schill, S.L., and Bassett, D.S. (2019). Effective learning is accompanied by high-dimensional and efficient representations of neural activity. *Nat. Neurosci.* *22*, 1000–1009.
- Tehovnik, E.J., Tolias, A.S., Sultan, F., Slocum, W.M., and Logothetis, N.K. (2006). Direct and indirect activation of cortical neurons by electrical microstimulation. *J. Neurophysiol.* *96*, 512–521.
- Titiz, A.S., Hill, M.R.H., Mankin, E.A.M., M Aghajian, Z., Eliashiv, D., Tchémoudanov, N., Maoz, U., Stern, J., Tran, M.E., Schuette, P., et al. (2017). Theta-burst microstimulation in the human entorhinal area improves memory specificity. *eLife* *6*, e29515.
- Tolias, A.S., Sultan, F., Augath, M., Oeltermann, A., Tehovnik, E.J., Schiller, P.H., and Logothetis, N.K. (2005). Mapping cortical activity elicited with electrical microstimulation using fMRI in the macaque. *Neuron* *48*, 901–911.
- Uhlhaas, P.J., and Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* *52*, 155–168.
- Volgushev, M., Chistiakova, M., and Singer, W. (1998). Modification of discharge patterns of neocortical neurons by induced oscillations of the membrane potential. *Neuroscience* *83*, 15–25.
- Von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: From local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* *38*, 301–313.
- Widge, A.S., Zorowitz, S., Basu, I., Paulk, A.C., Cash, S.S., Eskandar, E.N., Deckersbach, T., Miller, E.K., and Dougherty, D.D. (2019). Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function. *Nat. Commun.* *10*, 1536.
- Williams, L.M. (2017). Defining biotypes for depression and anxiety based on large-scale circuit dysfunction: a theoretical review of the evidence and future directions for clinical translation. *Depress. Anxiety* *34*, 9–24.
- Wu, H., Miller, K.J., Blumenfeld, Z., Williams, N.R., Ravikumar, V.K., Lee, K.E., Kakusa, B., Sacchet, M.D., Wintermark, M., Christoffel, D.J., et al. (2018). Closing the loop on impulsivity via nucleus accumbens delta-band activity in mice and man. *Proc. Natl. Acad. Sci. USA* *115*, 192–197.
- Yang, Y., Sani, O.G., Chang, E.F., and Shانهchi, M.M. (2019). Dynamic network modeling and dimensionality reduction for human ECoG activity. *J. Neural Eng.* *16*, 056014.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Experimental Models: Organisms/Strains		
Rhesus macaque (<i>Macaca mulatta</i>)	Covance (Monkey A) Charles River Laboratories (Monkey M)	N/A
Software and Algorithms		
MATLAB R2017b	Mathworks	https://www.mathworks.com/products/matlab.html
FSL	Analysis Group, FMRIB, Oxford, UK	https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/
FreeSurfer	Laboratory for Computational Neuroimaging, Athinoula A. Martinos Center for Biomedical Imaging, Boston, USA	https://surfer.nmr.mgh.harvard.edu/
Scalable Brain Atlas	Scalable Brain Atlas Composer	https://scalablebrainatlas.incf.org/index.php
roc.m (receiver operating characteristic analysis)	KULeuven-ESAT-SCD	https://www.esat.kuleuven.ac.be/sista/lssvmlab
POLARMAP: polarized colormap	François Beauducel	https://www.mathworks.com/matlabcentral/fileexchange/37099-polarmap-polarized-colormap
Experimental Hardware		
Microelectrodes (recording)	Alpha Omega single electrodes	http://www.ao-neuro.com/index.php?route=product/product&path=65&product_id=63
Microelectrodes (recording and microstimulation)	MicroProbes monopolar electrodes	https://microprobes.com/products/metal-microelectrodes/monopolar-electrodes/platinum-iridium
Neural recordings and amplifier	NSpike NDAQ system, Harvard Instrumentation	http://nspike.sourceforge.net/#Overview
Microstimulator	Blackrock Microsystems CereStim R96	https://www.blackrockmicro.com/neuroscience-research-products/ephys-stimulation-systems/cerestim-96-neurostimulation-system/
Stim-record headstages	Blockrock Microsystems	https://www.blackrockmicro.com/neuroscience-research-products/ephys-headstages/analog-headstages/
Touch screen	ELO Touch Systems	https://www.elotouch.com/
Task controller	Custom LabView software with a real-time embedded system NI PXI-8820	https://www.ni.com/en-us/shop/labview.html
Deposited Data		
Experimental data and code	Custom software and experimental data	https://github.com/pesaranlab/causal_network_analysis_pub

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Bijan Pesaran (bijan@nyu.edu).

Materials Availability

This study did not generate any new unique reagents.

Data and Code Availability

The code and datasets supporting this study are available at GitHub (https://github.com/pesaranlab/causal_network_analysis_pub).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All surgical and experimental procedures were performed in compliance with the National Institute of Health Guide for Care and Use of Laboratory Animals and were approved by the New York University Institutional Animal Care and Use Committee. Two male rhesus macaques (*Macaca mulatta*) participated in the study (Monkey M, 8.4 kg and Monkey A, 7 kg at the beginning of the experiments).

METHOD DETAILS

Large-scale microdrive system design and assembly

We developed a customized large-scale semi-chronic microdrive system for mapping and manipulating the cortico-subcortical mood processing network. The screw-driven actuation mechanism of the microdrive provides bi-directionally independent control of the position of 220 microelectrodes (1.5-mm spacing) along a single axis with a range up to 32 mm (Monkey M) and 40 mm (Monkey A) with 125- μ m pitch (Dotson et al., 2017). Each actuator consists of a lead screw, a teardrop brass shuttle bonded to the electrode tail with conductive epoxy, and a compression spring. Each electrode can be moved with an accuracy of approximately 15 μ m. For Monkey M, we distributed two types of microelectrode in the microdrive, 160 platinum/iridium (Pt/Ir) electrodes (MicroProbes, Gaithersburg, MD) with impedance 0.1–0.5 M Ω for extracellular recording and microstimulation, and 60 tungsten electrodes (Alpha Omega, Israel) for extracellular recording with impedance 0.8–1.2 M Ω . For Monkey A, we loaded 220 Pt/Ir electrodes (MicroProbes, Gaithersburg, MD) with impedance 0.5 M Ω for extracellular recording and microstimulation. Electrode impedances were measured at 1 kHz (Bak Electronics, Umatilla, FL). The tungsten electrode's shank diameter was 125 μ m and its total diameter was 250 μ m with glass insulation. The Pt/Ir electrode's shank diameter was 225 μ m and its total diameter was 304 μ m with parylene C and polyimide insulation.

Magnetic resonance imaging and processing

We reconstructed each monkey's brain, skull and cerebral vasculature (Brainsight®, Rogue Research, Montreal, QC) from anatomical magnetic resonance images (MRIs) and MRIs using ABLAVAR® contrast agent for angiography with a T1-weighted magnetization-prepared rapid acquisition gradient-echo (MPRAGE) sequence. We also performed multishell *high-angular resolution diffusion imaging* (HARDI) tractography (Jbabdi et al., 2012) registered to microdrive implantation. We collected diffusion weighted images to reconstruct white matter tracts connecting key cortical and subcortical areas of interest. During MRI procedures, the monkeys were anesthetized with isoflurane and placed in the scanner in sphinx position. We acquired data with a 3-Tesla (3T) Siemens Allegra (Erlangen, Germany) using 3 elements out of a 4-channel phased array from Nova Medical Inc. (Wilmington, MA) and 64 gradient directions in 1.2 mm² in-plane resolution (TR = 7000 ms; TE = 110 ms; b-values: 0, 750, 1500, 2250 s/mm²; FOV: 80 × 64 pixels; slices: 48; slice thickness: 1.2 mm; DWI to b_0 ratio 65:1). To correct for geometric distortions from field inhomogeneities caused by the non-zero off-resonance fields, we collected data with reversed phase-encode blips, forming pairs of images with distortions going in opposite directions. From these pairs, we estimated the susceptibility-induced off-resonance field using FSL's TOPUP tool (Smith et al., 2004). We combined the two images into a single corrected image. We corrected Eddy currents generated by the 64 gradient directions using FSL's eddy tool (Andersson and Sotiropoulos, 2015).

Experimental preparation

We performed a craniotomy and dura thinning on the targeted brain regions over the left (Monkey M) and right (Monkey A) hemispheres. We implanted a customized large-scale recording chamber (Gray Matter Research, Bozeman, MT) fitted to the skull surface using MR-guided stereotaxic surgical techniques (Brainsight®, Rogue Research, Montreal, QC). We aligned the chamber and registered it within 1 mm of the target coordinates (nominally 0.4 mm) and affixed and sealed it to the skull surface via C&B-METABOND® (Parknell Inc., Edgewood, NY) and dental acrylic. We then mounted the microdrive into the chamber and sealed it with compressed gaskets and room-temperature-vulcanizing (RTV) sealant (734 flowable sealant, Dow Corning, Midland, MI). To target each brain region, we registered electrodes to anatomical magnetic resonance images (MRIs) and magnetic resonance angiograms (MRAs). To limit routes for infection, we customized the shape of the chamber and microdrive to each animal's anatomy and established that the chamber was sealed following implantation. To limit damage to the vasculature, midline, and ventricles when lowering electrodes, we only advanced a subset of electrodes along trajectories considered safe, greater than 2 mm from MR-visible vasculature, ventricles, and midline. Finally, we co-registered the MRIs to the MNI Paxinos labels (Frey et al., 2011) to label each recording and stimulation site (Figure S1). We successfully studied activity from 165 electrodes (50 tungsten and 115 Pt/Ir) in Monkey M for over 24 months and 208 Pt/Ir electrodes in Monkey A for over 12 months.

We tested long-term stability of neural recordings in two ways. First, we examined local field potential (LFP) activity during the movement epoch of a reaching task with an instructed delay (1–1.5 s). Figures S1A–S1C present task-evoked LFP activity for 67 depths along a sample electrode trajectory. Task-evoked LFP activity clearly increased and decreased in magnitude as the electrode transitioned into and out of gray matter, respectively. We also validated localization based on spiking. Increased task-evoked LFP activity was associated with increasing spiking activity (Figures S1D–S1G). Patterns of spiking activity and task-evoked LFP activity were preserved over many months, demonstrating long-term stability of the implanted device.

Neural recordings

The monkeys were awake, head-restrained and seated in a primate chair placed in an unlit sound-attenuated electromagnetically shielded booth (ETS Lindgren). Neural recordings were referenced to a ground screw implanted in the left posterior parietal lobe (Monkey M) and left occipital lobe (Monkey A), respectively, with the tip of the screw just piercing through the dura mater. Neural signals from all channels were simultaneously amplified and digitized at 30 kHz with 16 bits of resolution with the lowest significant bit equal to 0.1 μV (NSpike, Harvard Instrumentation Lab; unit gain headstage, Blackrock Microsystems), and continuously streamed to disk during the experiment with lights switched off in the recording booth.

Sampling network nodes and edges from recording

We consider each brain region to be a network node. We sampled each node by positioning an electrode at a site in the node. For every pair of electrodes that simultaneously recorded activity in two nodes, we obtained one sample of the network edge connecting the two nodes. Repeated node sampling enabled repeated edge sampling. With M electrodes positioned at different sites in Node1, and N less than M electrodes in Node2, we obtained N non-overlapping samples of the Node1-Node2 network edge. In Monkey M, we sampled four cortical nodes at 3521 sites (1506 OFC, 890 ACC, 817 dPFC, and 308 M1 sites) and three subcortical nodes at 1125 sites (915 CN, 122 Put, and 88 GP sites). In Monkey A, we sampled seven cortical nodes at 3900 sites (319 OFC, 469 ACC, 343 PCC, 427 dPFC, 713 M1, 429 S1, and 1200 PPC sites) and eight subcortical nodes at 997 sites (614 CN, 214 Put, 59 GP, 45 Amyg, 25 PrS, 20 SI, 2 BFB, and 18 NAc). We targeted our causal network analysis to 9548 samples from the cortico-subcortical limbic network across two monkeys, specifically for 1359 OFCACC, 1244 OFC-dPFC, 1529 OFC-CN, 336 OFC-Put, 45 OFC-Amyg, 1244 ACC-dPFC, 1359 ACC-CN, 336 ACC-Put, 45 ACC-Amyg, 1244 dPFC-CN, 336 dPFC-Put, 45 dPFC-Amyg, 336 CN-Put, 45 CN-Amyg, and 45 Put-Amyg interactions. Note that OFC combines medial OFC and lateral OFC; dPFC combines inferior, middle, and superior frontal gyri; and posterior parietal cortex (PPC) combining superior parietal lobule (SPL), supramarginal gyrus (SMG), and precuneus (pCun).

Sampling causal network edges from simultaneous microstimulation and recording

We applied microstimulation using a bipolar configuration, made by simultaneously sending a biphasic charge-balanced square wave pulse via a pair of Pt/Ir microelectrodes with the same pulse amplitude, pulse width, and interphase interval, but opposite polarity (e.g., cathode-lead for electrode 1 and anode-lead for electrode 2) (Cerestim R96, Blackrock Microsystems, Salt Lake City, UT). We used the pulse width as 100 μs per phase and interphase interval as 53 μs for all stimulation sessions. The Pt/Ir microelectrodes had a typical tip geometric surface area of $223 \pm 37 \mu\text{m}^2$. For instance, a single pulse, with amplitude of 40 μA and width of 100 μs per phase (4 nC/phase), could yield a charge density of approximately 1800 $\mu\text{C}/\text{cm}^2$. We simultaneously recorded neural signals from all electrodes while stimulating at certain pair of electrodes. No seizure activity due to stimulation was detected during experiments.

We implemented two microstimulation protocols to perform causal network analysis. Primarily used in this study, we designed a novel multisite spatiotemporal patterned microstimulation framework to assess changes in the weight of network edges detected using evoked LFP responses to single microstimulation pulses. For each stimulation bout, the stimulation pattern started with a Pre-tetanic single microstimulation pulse from a pair of electrodes at a network site (sender) to identify the network edges, followed by a sbTetMS from another pair of electrodes at another network site (sbTetMS-modulator), and followed by a Post-tetanic single microstimulation pulse at the sender. The latency between the Pre-tetanic pulse and the onset of the sbTetMS (τ_{Pre}) was either 0.5 or 1 s. The duration (T) of the sbTetMS pulse train and the latency (τ_{Probe}) between the offset of sbTetMS pulse train and the onset of Post-tetanic single pulse varied in a pseudo-random fashion bout by bout (T : range = 50-2000 ms; τ_{Post} : range = 0-200 ms; [Table S1](#)). We also varied the inter-trial interval bout by bout (duration of each stimulation bout plus 2-4 s variation). We defined the neural excitability (edge weight) of detected network edges as the peak magnitude of pulse-triggered averaged evoked LFP response within a 100-ms post single-pulse epoch. We also used a single-pulse microstimulation protocol with either periodic or Poisson pulse train ([Shewcraft et al., 2020](#)) to identify network edges. This protocol consisted of 1 s pulse trains and 1-3 s baseline epochs. Due to stimulation artifact, to preserve a minimum 100-ms post single-pulse epoch for analyzing the response to each of pulses, we used pulse train frequencies of 5 and 10 Hz with corresponding refractory period of 200 and 100 ms, respectively.

Bipolar microstimulation limited the current spread from the stimulation site and corresponding stimulation artifacts. To limit the common-mode confounds of neural recordings, we digitally re-referenced the neural signal of each electrode to its nearest neighbor within 5 mm based on the electrode depth in the same recording headstage. No electrode was shared by more than one pair. We used the locally bipolar re-referenced neural recording throughout the entire study. We used locally bipolar re-referenced recordings to reveal focal evoked responses as well as to limit stimulation artifacts. The secondary intervention was designed to be minimally perturbative. By stimulating the sender above threshold and recording the receiver at threshold, the receiver response to input from the sender measured the weight of the network edge (sender \rightarrow receiver). To do so, we stimulated the sender with different current amplitudes. We then set the stimulation dose at the threshold needed to generate a stimulation-evoked response at receiver sites. Of 152 stimulation site-amplitude combinations tested to establish the threshold microstimulation (10 to 80 μA ; [Table S1](#)), we detected and characterized 110 directed functional interactions (sender-receiver communication) in 21 network edges across 6,910 pairs of sites tested in 132 network edges we causally sampled. Note that we refer a pair of bipolar re-referenced electrodes for recording to a site sampled in the brain. We summarized all stimulation parameters used to detect these sender-receiver pairs in [Table S1](#).

Our inferences of detecting directed functional interactions depend on using a minimally-perturbative secondary intervention, for which we present multiple lines of supporting evidence. First, we used single, bipolar, biphasic pulses of relatively low-amplitude that are not necessarily expected to generate large-scale network responses. Indeed, single microstimulation pulse has not been previously used to map large-scale networks in the awake primate brain. The natural concern is that network responses following such minimal perturbations may not be detectable. Second, following single microstimulation pulses we observed sparse directed functional interactions instead of widespread network effects. Approximately 1.6% of sampled sites showed statistically-significant responses. We should note that sparsity may result from the presence of residual stimulation artifacts that obscure responses at some sites. The degree of sparsity likely also reflects the minimally-perturbative microstimulation approach we used, which is conservative and so may fail to drive detectable responses at some sites where more pre-synaptic inputs might be required to produce excitatory post-synaptic activity (Häusser et al., 2001). Sparsity also likely reflects the underlying anatomical connections and the fact that many recording sites may not receive inputs from the sender. Third, although the responses were sparse, they were strong, visible on a single-pulse basis, and varied moment-by-moment. The presence of strong pulse-evoked responses demonstrates the efficacy of the intervention and mitigates concern about false positives due to multiple comparisons involving the number of tested network edges. The absence of pulse-evoked responses at other times also demonstrates the receiver is perturbed at threshold and hence that the secondary intervention is unlikely to be altering interactions across the network.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pulse-triggered evoked potentials

We obtained LFP activity offline by low-pass filtering the broadband raw recording at 400 Hz using a multitaper filter with time duration of 0.025 s, frequency bandwidth of 400 Hz, and center frequency of 0 Hz, and then down-sampled to 1 kHz from 30 kHz. We removed the events from the analysis if they exceeded 10 standard deviations of from the mean across the stimulation event pool. We also removed noisy or bad channels via visual inspection. Pulse-triggered evoked potentials were computed by averaging bipolar re-referenced LFP signals aligned to the onset time of each single pulse. The z-scored pulse-triggered evoked potentials were then computed using the standard deviation (SD) of the baseline (–30 to –5 ms). We computed the baseline SD for each electrode separately using all time points in the baseline window of the pulse-triggered evoked potential.

We should note that suppression of the evoked response to the Post-tetanic single pulse we observed may be confounded by an effect related to the phase of inhibition after the excitatory response to the Pre-tetanic single pulse (Butovas and Schwarz, 2003; Logothetis et al., 2010; Seidemann et al., 2002) or related to suppressed responses to paired microstimulation pulses (Butovas and Schwarz, 2003; Castro-Alamancos and Connors, 1996). Our experimental design employed Pre-tetanic and Post-tetanic pulses separated by 0.75–3 s, mitigating such concerns.

Pulse-triggered multiunit activity

We obtained multiunit activity (MUA) offline by first band-pass filtering the broadband raw recording from 0.3 to 6.6 kHz with time duration of 0.01 s, frequency bandwidth of 3 kHz, and center frequency of 3.3 kHz. We then applied a 3.5 standard deviation threshold to identify putative spikes (1.6-ms duration). We labeled all waveforms that exceeded this peak as multiunit action potentials. We computed multiunit peristimulus time histograms by aligning MUA to the onset time of each single pulse. Individual spikes were collected in 1 ms bins and the corresponding histogram was smoothed by convolving with a Gaussian function with a standard deviation of 5 ms. Data from –1 ms to 5 (or 10) ms around the single pulse onset were not used for multiunit activity analysis due to stimulation artifact.

Detection of pulse-evoked LFP responses using the stimAccLLR method

To quantify the detection of evoked responses to microstimulation on a single-pulse basis, we used the stimulation-based accumulating log-likelihood ratio (stimAccLLR) method to determine when selectivity in the LFP activity for the null and alternative hypotheses emerged, i.e., the latency of a stimulation response. The latency from single microstimulation pulses was determined by the time at which Pulse Event reached a detection threshold. The threshold was selected based on a trade-off between speed (latency) and accuracy (probability of correct classification) as we increased the level of the detection thresholds from AccLLR equal to zero (Figure S3). We used the detection threshold to determine the response latencies by maximizing difference between probability of ‘Hit’ and ‘Miss’ events.

In the stimAccLLR method, we defined a probabilistic model of the LFP activity for the two alternatives being tested. To determine the latency for a response to single microstimulation pulses, we define the two alternative hypotheses into “condition 1” and “condition 2,” where “condition 1” represents the LFP activity (Pulse Event) in the post-stimulus epoch, while “condition 2” represents LFP activity (Pulse Null) in the pre-stimulus epoch. We performed stimAccLLR on all stimulation response to single microstimulation pulses, combining all Pre-tetanic single pulses using the multisite spatiotemporal patterned microstimulation protocol and all the first pulse of each burst using the Poisson burst microstimulation protocol. We term all the single pulses used in the stimAccLLR procedure Pre-pulses. Pulse Event LFP activity lasted up to 100 ms after Pre-pulse onset plus 5 or 10-ms blanking epoch due to stimulation artifact, while Pulse Null LFP activity lasted up to 100 ms before Pre-pulse onset.

We modeled LFP activity as independent observations from an underlying Gaussian distribution. The signal $x(t)$ at time t is expressed as a mean waveform $\mu(t)$ with additive Gaussian noise, $\varepsilon(t)$. We estimated the mean response $\mu(t)$ by low-pass filtering the raw LFP activity at 100 Hz.

$$x(t) = \mu(t) + \varepsilon(t)$$

The likelihood of observed data $x(t)$ being generated by each model was given by

$$P[x(t) - \mu(t) | \sigma^2] \sim N(0, \sigma^2)$$

The likelihood ratio, $LLR(t)$, at time t for two time-varying Gaussian LFP models, assuming the noise in both models had the same variance, σ^2 , was given by

$$LLR(t) = \log \frac{P[x(t) - \mu_1(t) | \sigma^2]}{P[x(t) - \mu_2(t) | \sigma^2]} = \frac{(x(t) - \mu_2(t))^2 - (x(t) - \mu_1(t))^2}{2\sigma^2}$$

When estimating $LLR(t)$ across events/trials, we used leave-one-out procedure. Finally, we calculated the accumulated log-likelihood ratio by summing log-likelihoods over time:

$$AccLLR(t) = \sum_{t'=0}^t LL(t')$$

To quantify signal selectivity, we performed a receiver-operating characteristic (ROC) analysis on the AccLLR values. Analysis was performed every millisecond to discriminate activity from “condition 1 (Pulse Event)” and “condition 2 (Pulse Null)” events. We used an ideal observer analysis to measure the overall signal selectivity (OSS) for the stimAccLLR procedure. We defined OSS as the choice probability from the ROC analysis at the end of an accumulation interval up to 100 ms. Unlike a threshold-dependent performance metric, the OSS reports the results of an ideal-observer analysis in the form of a threshold-independent choice probability at the end of the accumulation time. To do so, we first converted the OSS to a p value for each of the sampled bipolar-referenced recording sites. To explicitly correct for multiple comparisons, for every recording site across all experimental sessions, we then controlled the P value of the OSS by performing a false discovery rate (FDR) procedure (Benjamini and Hochberg, 1995) with an alpha threshold set at 0.01. Using such procedures, we detected 110 significant sender-receiver pairs (edge samples, $P_{FDR} < 0.01$).

We summarized the detection rate of the stimulation response to a single microstimulation pulses between brain regions (network nodes) in Table S2. We run a binomial test at the significant level of 0.05 to determine if some causally sampled network edges that showed directed functional interactions were more likely selective than others.

Decoding sender-receiver communication to reveal modulators

Performing stimAccLLR reveals ‘Hit’ and ‘Miss’ events in the receiver responses. To further test for the presence of two response classes, we classified pulse-by-pulse evoked responses using a Gaussian mixture model (GMM). The GMM clustered pulse-evoked responses into two classes. For a majority of the detected network edge samples (62/110, 56%), the two analysis methods agreed ($p < 0.05$, Fisher’s exact test, two-tailed test; Figure S8).

To identify modulators for each of detected network edges (sender-receiver pairs), we grouped the Pre-pulse baseline LFP activity (500-ms epoch before pulse onset) of all simultaneously recorded modulator candidate sites in the gray matter into two categories based on ‘Hit’ and ‘Miss’ events detected from the stimAccLLR procedure. We estimated baseline LFP power using multitaper methods with 500-ms sliding window with ± 2 -Hz smoothing. We removed the events from the analysis if they exceeded 5 standard deviations of from the mean across the stimulation event pools. We then performed the ROC analysis to test if baseline LFP power of each tested site in ‘Hit’ events can be significantly differentiated from ‘Miss’ events in the β -frequency band (10-40 Hz). For each of detected sender-receiver pairs, we computed the P value of the area under ROC curve (AUC) of each tested site. To control for multiple comparisons, we performed the FDR procedure on all tested sites for a given detected sender-receiver pair. We identified a tested site as a modulator, where its AUC was significant above chance (one-tail test, $P_{FDR} < 0.05$; Figure 3E). To demonstrate modulator decoding performance, we selected the magnitude of modulator baseline mean log β -power as the threshold when the overall decoding performance is maximized (Figure 3H).

Of 110 significant edge samples we detected, we tested 6189 tested candidate modulator sites. Only 31 edge samples revealed significant modulators. Of the 6189 candidate modulator sites tested, 1163 candidate modulator sites were tested from the 31 edge samples that revealed significant modulators. Overall, we revealed the 92 significant modulators. We summarized the number of tested modulator candidate sites per edge sample in Table S1.

We also performed the ROC analysis on the receiver baseline activity in Pulse Null events as a control. Pulse Event activity at the receiver may not be due to single-pulse stimulation at the sender and may simply reflect pulses of intrinsic activity. If so, the pulses of activity may exist during Pulse Null events and may also be predicted by modulators. We repeated stimAccLLR analysis on each of the 110 detected edge samples for receiver baseline activity (Pulse Null) using the detection threshold for detecting ‘Hit’ and ‘Miss’ events for receiver Pulse Event activity. We found that nearly all identified modulators (91/92) did not predict the false alarms for the receiver Pulse Null events ($P_{FDR} > 0.05$).

Modulation spectrogram of modulator baseline activity

To show how the baseline neural activity of tested modulators significantly modulated sender-receiver communication, we first sorted the events by ‘Hit’ and ‘Miss’ events at the receiver. We then estimated spectrograms of modulator baseline LFP activity (1 s epoch before Pre-tetanic pulse onset) in ‘Hit’ and ‘Miss’ events, respectively, using multitaper methods with 500-ms sliding window with ± 5 -Hz smoothing and 5-ms steps between spectral estimates up to 60 Hz. We tested the difference in LFP power between ‘Hit’ and ‘Miss’ events against a null hypothesis that there was no LFP power difference using a permutation test (10,000 permutations). To generate the null distribution for no LFP power difference, we randomly shuffled the order of combined ‘Hit’ and ‘Miss’ events, followed by computing shuffled LFP power difference between ‘Hit’ and ‘Miss’ events. For the significant regions presented in the spectrograms, we applied a cluster correction to correct for multiple comparisons at the significant level of 0.05 (Maris et al., 2007). We drew the contours to highlight the significant regions on the spectrogram. We also performed the same analysis to generate modulation spectrograms of the identified modulators using the events sorted by decoded ‘Hit’ and ‘Miss’ events at the receiver.