1   **Frequency-dependent representation of reinforcement-related information in the**

2   **human medial and lateral prefrontal cortex**

3   (Reinforcement-related information in human prefrontal cortex)

4   Elliot H. Smith[1], Garrett P. Banks[1], Charles B. Mikell[1], Syndey S. Cash[2], Shaun R. Patel[3],

5   Emad A. Eskandar[3], Sameer A. Sheth[1]

6   1. Department of Neurological Surgery, Columbia University Medical Center, New York,

7   NY, 10032

8   2. Department of Neurology, Massachusetts General Hospital, Boston, MA, 02114

9   3. Department of Neurosurgery, Massachusetts General Hospital, Boston, MA, 02114

10   Corresponding author: Sameer A. Sheth, ss4451@cumc.columbia.edu, 710 West 168th

11   Street, NI 4. New York, NY, 10032

12

13   Number of Pages: 31

14   Number of Figures: 6

15   Number of words in Abstract: 231

16   Number of words in Introduction: 505

17   Number of words in Discussion: 1405

18   Conflict of interest: The authors declare no competing financial interests

23

24   **Abstract**

25   The feedback-related negativity (FRN) is a commonly observed potential in scalp

26   electroencephalography (EEG) studies related to the valence of feedback about a

27   subject's performance. This potential classically manifests as a negative deflection in

28   medial frontocentral EEG contacts following negative feedback. Recent work has shown

29   prominence of theta power in the spectral composition of the FRN, placing it within the

30   larger class of "frontal midline theta" cognitive control signals. Although the dorsal

31   anterior cingulate cortex (dACC) is thought to be the cortical generator of the FRN,

32   conclusive data regarding its origin and propagation are lacking. Here we examine

33   intracranial electrophysiology from the human medial and lateral prefrontal cortex (PFC)

34   in order to better understand the anatomical localization and communication patterns of

35   the FRN. We show that the FRN is evident in both low- and high-frequency local field

36   potentials (LFPs) recorded on electrocorticography. The FRN is larger in medial

37   compared to lateral PFC, and coupling between theta band phase and high frequency LFP

38   power is also greater in medial PFC. Using Granger causality and conditional mutual

39   information analyses, we provide evidence that feedback-related information propagates

40   from medial to lateral PFC, and that this information transfer oscillates with theta-range

41   periodicity. These results provide evidence for the dACC as the cortical source of the

42   FRN, provide insight into the local computation of frontal midline theta, and have

43   implications for reinforcement learning models of cognitive control.

44

45   **Significance Statement:** Using intracranial electrophysiology in humans, this work

46   addresses questions about a frequently studied feedback-related electroencephalographic

47   signal, illuminating anatomical and functional properties of the representation of

48   feedback-related reinforcement during decision-making across the medial to lateral extent

49   of the human prefrontal cortex.

50

51   **Introduction**

52   Modifying future behavior based on reinforcement is essential for survival in complex

53   environments. A prominent event-related potential (ERP) observed on EEG is the

54   feedback-related negativity (FRN), which manifests as the difference between ERPs

55   evoked by positive and negative feedback (Walsh and Anderson, 2012). The FRN occurs

56   approximately 250 ms after feedback, regardless of the feedback modality (Miltner et al.,

57   1997).

58   The FRN is most commonly observed on frontocentral EEG contacts, where its

59   amplitude is highest (Miltner et al., 1997; Luu et al., 2003; Silvetti et al., 2014), but has

60   also been observed on parietal contacts (Cohen and Ranganath, 2007; Pfabigan et al.,

61   2010). EEG and MEG source localization has commonly identified the dorsomedial

62   prefrontal cortex (dmPFC), especially the dorsal anterior cingulate cortex (dACC), as the

63   source of the FRN (Miltner et al., 2003; Herrmann et al., 2004; Luu et al., 2004; Roger et

64   al., 2010; Doñamayor et al., 2011; Walsh and Anderson, 2013). fMRI studies have found

65   areas that have increased BOLD signal for positive feedback, compared with negative

66   feedback (Nieuwenhuis et al., 2005b). These studies have implicated a broad network in

67   FRN generation, including cingulate cortex, dorsolateral PFC (dlPFC) the basal ganglia,

68   and amygdalae (Holroyd et al., 2004b). One recent study that combined the temporal

69   resolution of EEG with the spatial resolution of fMRI localized the FRN to either the

3

70   dACC specifically, or to a distributed frontal "salience" network, depending on the

71   analysis technique used (Hauser et al., 2013).

72   An influential account of FRN function based on reinforcement learning concepts

73   links it to the reward system (Holroyd and Coles, 2002a; Holroyd et al., 2008).

74   Mesencephalic dopamine projections to the striatum and dmPFC reinforce advantageous

75   behaviors (Schultz, 2002). Phasic dopamine release is thought to encode reward

76   prediction error (RPE) signals, or differences between expected and actual rewards

77   (Schultz, 2013). Errors, resulting in negative feedback or reward omission, produce

78   phasic decreases in dopamine release. As dopaminergic input is inhibitory to the dACC,

79   phasic decreases disinhibit layer V dACC neurons, allowing them to depolarize and

80   produce ERPs (Holroyd and Coles, 2002b). Indeed, FRN magnitude is directly

81   proportional to RPE magnitude and the amount of behavioral adaptation in subsequent

82   trials (Cavanagh et al., 2010).

83   Recent studies have shown that the FRN, along with other frontocentral

84   negativities such as the error-related negativity (ERN) and N2, share common features.

85   Spectral analyses have demonstrated increased 4-7 Hz power, suggesting a common

86   mechanism of communication for these frontal midline theta (FMϴ) signals (Holroyd et

87   al., 2002; Chase et al., 2011; Cavanagh and Frank, 2014; Cavanagh and Shackman,

88   2014). These signals are thought to convey aspects of cognitive control and may have

89   evolved as reactions to threatening situations, engaging cognitive resources in times of

90   high demand (Cavanagh and Shackman, 2014).

91   Here we examine intracranial electrophysiology from the human dACC and

92   dlPFC to localize and characterize the neural representation of prefrontal feedback

4

93   signals. We show that the dACC is the source of the FRN, that the information contained

94   in this signal propagates to dlPFC, and that this communication occurs in the theta

95   frequency band.

96

Materials and Methods

98   *Ethics statement.* The Columbia University Medical Center and Massachusetts General

99   Hospital Institutional Review Boards approved these experiments. All patients in this

100   study provided informed consent prior to participating in this research.

101   *Subjects.* We examined intracranial electrocorticographic (ECoG) signals in 7

102   patients (2 female) undergoing monitoring for medically refractory epilepsy. Eight-

103   contact depth electrodes (AdTech Medical, Racine, WI) were implanted through the

104   mediolateral extent of prefrontal cortex (10 left hemisphere and 9 right hemisphere), thus

105   providing simultaneous LFP recordings from dACC and dlPFC. Electrode placement was

106   determined solely on clinical grounds, and participants were free to withdraw from the

107   study without influencing their clinical care.

108   *Depth Electrode localization.* Depth electrodes were localized using pre-operative

109   magnetic resonance images (MRI) and post-implant computed tomography (CT) scans. A

110   linear transform from CT space to pre-operative MRI space was designed using FMRIB

111   Software Library's (FSL) linear transformation algorithm (Jenkinson and Smith, 2001).

112   Next, a transform from patient MRI space to 2 mm Montreal Neurological Institute

113   standard space using both linear and non-linear transforms was implemented (Jenkinson

114   et al., 2005; Andersson et al., 2013). Depth electrodes' medial and lateral-most contacts

5

115   were then respectively superimposed on the medial and lateral surface of the standard

116   brain for visualization.

117   *Behavioral task.* Recordings were acquired while patients performed the Multi-

118   Source Interference Task (MSIT) (Figure 1A) (Sheth et al., 2012). The MSIT is a Stroop-

119   like cognitive interference task requiring the subject to view a stimulus consisting of

120   three numbers (composed of integers 0 through 3), and identify the one number ("target")

121   different from the other two ("distracters"). The subject has to indicate his/her choice on

122   a 3-button pad. If the target number is "1", the left button is the correct choice; if "2", the

123   middle button; and if "3", the right button. Importantly, the subject must choose the

124   correct button regardless of where the target number appears in the sequence. Reaction

125   time (RT) was defined as the time between stimulus presentation and subject response.

126   Differences in RT among conditions was tested for significance using a one way Kruskal-

127   Wallis test over interference types, and a Wilcoxon signed rank test between RTs for

128   feedback conditions.

129   The task contains two types of cognitive interference. Simon interference occurs

130   due to the presence of spatial incongruence between the position of the target number in

131   the stimulus and the position of the correct button on the button pad. For example, the

132   correct choice for both "200" and "020" is the middle button; the first cue contains spatial

133   interference (incongruence between the left position of the target number in the stimulus

134   and middle position of the correct button choice), whereas the second one does not.

135   Eriksen Flanker interference occurs due to the presence of distracters that are possible

136   button choices. For example, the correct choice for both "133" and "100" is the left

6

137 button; the first cue contains distracter interference ("3" is a possible button choice),

138 whereas the second one does not ("0" is not a possible choice).

139   Patients received either valenced or neutral feedback in blocks of 10 trials. During

140 valenced feedback trials, the target number changes color to indicate whether the patient

141 responded correctly: green for correct responses and red for incorrect responses. During

142 neutral feedback trials, the target number changes to blue regardless of the response

143 (Figure 1A).

144   *Data collection and preprocessing.* Neural data from depth electrodes were

145 sampled at 500 Hz on the clinical data acquisition system (Xltek, Natus Medical Inc., San

146 Carlos, CA). First, removal of the common mode was achieved by reconstructing the data

147 without the first principal component, based on the covariance of all examined contacts

148 (including white matter contacts) in each patient. This method removed DC offsets,

149 common artifacts, and line noise present in the raw data. Spectrograms were then

150 calculated using multi-tapered spectral analysis with a time-bandwidth product of 5 and 9

151 tapers, 200 ms windows and 10 ms step sizes, and normalized relative to mean baseline

152 spectra (500 ms prior to the appearance of the fixation point). High gamma activity was

153 derived from the mean values of spectrograms between 70 and 125 Hz. These frequency

154 limits were chosen to be above the line noise on the low end, and half of the Nyquist

155 frequency on the upper end. To generate low-frequency ERPs, data were low-pass

156 filtered at 40 Hz (150th order Fir filter) and averaged over trials.

157   In order to motivate our frequency-specific hypotheses and compare this study

158 with previous EEG studies, spectrograms and inter-trial phase coherence (ITC) were

159 calculated from Morlet wavelet decompositions on 72 scales from 1 to 125 Hz. The

7

160   group spectrogram is the average of the absolute value of each subject's spectrogram.

161   ITC was derived from averaging the absolute value of the phase of the LFP at each

162   frequency value in the wavelet decomposition over trials. Paired t-tests over trials were

163   used to assess significant differences between feedback conditions during the outcome

164   period for each frequency band in the time- frequency representations. These results were

165   corrected for the number of frequency bands using the Benjamini-Hochberg false

166   discovery rate (FDR) (Hochberg and Tamhane, 1987).

167   *FRN quantification.* The FRN for both low and high frequency signals was

168   defined as the difference in signal between correct and neutral feedback trials. We did not

169   calculate the difference between correct and incorrect feedback responses because of the

170   low error rate (see below). Previous work has shown that the FRN is dependent upon

171   expectations established by the context, such that an FRN is observed for negative

172   feedback when the range of outcomes varies from neutral to negative, but for neutral

173   feedback when the range varies from positive to neutral (Holroyd et al., 2004a;

174   Nieuwenhuis et al., 2005a; Jessup et al., 2010).

175   To account for variation in the size of the mean ERP between medial and lateral

176   sites, low frequency FRN amplitude was quantified by normalizing the maximum FRN

177   amplitude by the ERP amplitude (i.e the difference between the minimum and maximum)

178   at each contact. FRN amplitude is therefore reported relative to the size of the mean ERP

179   at each site. For high frequency signals, the FRN was quantified by examining

180   differences in high gamma power during feedback presentation. Hypothesis testing for

181   these results were carried out using Mann-Whitney U tests within patients and between

182   feedback conditions for both low frequency signals and high gamma power. In order to

8

183   address causal influences of feedback-related representations in dACC and dlPFC, we

184   calculated the latency of the FRN signal as the point of separation of the averaged high

185   gamma traces for neutral and correct feedback as long as that separation was maintained

186   for at least 500 ms. Latency was further quantified as the peak cross-correlation of

187   evoked high gamma signals between medial and lateral contacts.

188   *Phase-amplitude coupling*. To determine whether there was a functional

189   interaction between low and high frequency LFPs we examined phase-amplitude

190   coupling between a range of low frequency (2-25 Hz) phases and high-gamma power for

191   each channel separately. The high-gamma power was determined with the Hilbert

192   transform of the band-pass filtered signal between 70 and 125 Hz (4th order butterworth).

193   Frequencies for phase were filtered with a 3 Hz-wide band-pass filter centered on each

194   phase frequency (4th order butterworth). Phase values were extracted using the arctangent

195   of the real and imaginary values of the Hilbert transforms of each filtered low-frequency

196   band. High gamma amplitudes were then binned in 32 phase bins between $-\pi$ and π for

197   each low-frequency band and retained for further analysis. To determine significantly

198   modulated phase frequencies, each phase frequency's distribution of high gamma

199   amplitudes was tested for significant modulation with a Kolmogorov-Smirnov goodness-

200   of-fit test against a uniform null distribution and corrected for the number of frequency

201   bands using the FDR (Hochberg and Tamhane, 1987). The best frequency-for-phase for

202   each electrode and patient was also determined from the Kolmogorov-Smirnov goodness-

203   of-fit test, as the most significant frequency-for-phase band.

204   In order to quantify the amount of modulation in significant frequency-for-phase

205   bands, modulation indices were calculated for significantly modulated frequency-for-

9

206   phase bands. Modulation index was calculated in a similar manner to Canolty et al.

207   (Canolty et al., 2006) and Miller et al. (Miller et al., 2010; 2012). Briefly, the analytic

208   signals returned from the Hilbert transforms of the filtered ECoG data for each trial, n,

209   were separated into their amplitude and phase components a_n and ϕ_n. The modulation

210   index used here is the magnitude of the vector defined by the complex variable

211

212        $$z_n = a_n e^{i\phi_n}$$

213

214   during the feedback period. Significance was determined by comparing distributions of

215   these vectors projected onto the mean vector for each feedback condition using ANOVA,

216   and the Tukey-Kramer test for pairwise comparisons.

217   *Granger Causality.* Granger causality (GC) is a commonly used metric to infer

218   directional connectivity among brain areas (Seth, 2010); (Boatman-Reich et al., 2010);

219   (Barnett and Seth, 2014; Rodrigues and Andrade, 2014). GC measures the extent to

220   which one signal can predict another by imposing time lags on one signal and examining

221   resulting changes in the regression coefficients of an autoregressive model of both

222   (Granger, 1969; Seth and Edelman, 2007). Here within subject GC was determined using

223   the Multivariate Granger Causality Toolbox for Matlab (Barnett and Seth, 2014). We

224   examined all possible interactions between the two medial- and lateral-most contacts

225   among all electrodes in each patient, for the feedback period. We considered interactions

226   between combinations of pairs of electrodes in this frontal network (630 possible

227   connections). The Lyapunov exponents of the autoregressive model based on the raw

228   data were initially greater than one, indicating that the data did not meet the criteria of

10

229   stationarity required for GC analysis. A μ-law compressor, with $\mu = 255$, was therefore

230   applied to the z-scored broadband time series to compress the dynamic range of the

231   signals while maintaining the phase structure necessary to infer directional connectivity.

232   The multivariate autoregressive model, A, henceforth had a stable spectral radius

233   ($\rho(A) = 0.96$). Appropriate model order specification was determined using the

234   Bayesian information criterion (Bressler and Seth, 2011). A model order of 22 was

235   determined to be optimal for each of the seven patients and the mean±standard deviation

236   (over the seven patients) number of lags used was 104.34±36.71 samples.

237   Statistical significance was determined by testing the distribution generated by the

238   multivariate autoregressive model against a theoretical χ^2-distribution. The criterion for

239   significance was set at 0.05 after correction for multiple hypotheses (Hochberg and

240   Tamhane, 1987).

241   *Permutation Conditional Mutual Information.* GC analysis relies on several

242   assumptions that may not always hold in ECoG data, including stationarity, as mentioned

243   above. Conditional mutual information (CMI) was therefore examined as an additional

244   measure of information transfer between medial and lateral contacts. CMI is an

245   information theoretic measure of information transfer between cortical areas or recording

246   sites. CMI has been used to describe information transfer in cortical circuits without

247   relying on Gaussian statistics or imposing time delays on neural data (Quinn et al., 2010;

248   Salvador et al., 2010; Smith et al., 2012; 2013).

249   Here we computed the permutation, or symbolic, CMI (Li and Ouyang, 2010;

250   King et al., 2013). For permutation CMI, broadband time series were discretized into

251   groups of 6 symbols, 3 samples at a time, representing changes in the recorded voltage on

252   each contact. Probability distributions were then generated from the pairwise frequencies

253   of symbols derived from the two signals being compared (Keller and Lauffer, 2003;

254   Zunino et al., 2010). CMI was then computed as

255

$$I(F; X|Y) = \sum_{f \in F} p(f) \sum_{y \in Y} \sum_{x \in X} p(x, y|f) log \frac{p(x, y|f)}{p(x|f)p(y|f)}$$

256

257   Where F represents feedback, X is the response in one cortical area and Y is the response

258   in another cortical area. $p(f)$ is the probability a trial contains feedback; $p(x)$ and $p(y)$ are

259   the probabilities of observing responses x and y respectively.

260   We calculated CMI for epochs during and just before the feedback period for

261   medial- and lateral-most pairs of contacts on each electrode. Symbolic CMI therefore

262   tells us how much information a signal contains about feedback valence, given that we

263   know the response to feedback valences on the other end of the electrode. CMI regarding

264   feedback valence was calculated for the medial electrode given that the response in the

265   lateral electrode was known and then for the lateral electrode given that the response in

266   the medial electrode was known. This calculation was made in 250 ms windows sliding 5

267   ms in time for 4 seconds surrounding feedback presentation.

268   We also calculated CMI for the same data with randomly shuffled samples for use

269   in statistical comparisons. Permutation CMI is a relative measure of information transfer.

270   Statistically significant differences from shuffled data were therefore interpreted as

271   representing information transfer from one cortical area to the other cortical area.

272   Statistical significance was assessed within patients using a modified χ^2 test, comparing

273   χ^2 distributions generated from $2Nlog_2$ times the lateral-to-medial and medial-to-lateral

274 CMI to the χ^2 distribution from the shuffled data (Fan et al., 2000). Here N is the number

275 of trials, and the degrees of freedom for each distribution are based on the number of

276 possible responses for each stimulus presentation and the total number of presentations

277 (Ince et al., 2011).

278   To quantify the amplitude and phase of oscillations observed in the CMI signal,

279 multi-taper spectra (5 leading tapers and a time-bandwidth product of 3) of the CMI

280 signals were calculated.

281

282 **Results**

283

284 **Behavioral data**

285 Seven subjects performed the task with similar behavior as has been previously reported,

286 with RT increasing with amount of interference in the stimulus (Figures 1B and C)

287 (Horga et al., 2011; Sheth et al., 2012). There was no significant difference in RT

288 between trials with valenced vs. neutral feedback (Wilcoxon signed rank, P = 0.78, N =

289 7) (Figure 1D). Subjects had a 1.6$\pm$1.3 % error rate over an average of 289$\pm$65.3 trials

290 per subject.

291

292 **Depth electrode localization**

293 Each depth electrode had 8 contacts. In most electrodes, the two medial- and lateral-most

294 contacts were determined to be within gray matter. The middle four contacts in each

295 electrode were predominantly in white matter and therefore were not examined. Two

296 medial contacts on one electrode were also determined to be in white matter. This

13

297 electrode was excluded from further analysis. We therefore examined responses on 72

298 cortical contacts over all 7 patients. Figure 2 shows the medial and lateral most contacts

299 in each patient projected onto the medial and lateral surfaces of the MNI standard brain.

300

301 **FRN is evident in high and low frequency signals and is larger in dACC than dlPFC**

302 We found evidence for differential representations of feedback valence in both low and

303 high frequency components of the ECoG recordings (Figures 3A and 3B; Paired t-tests,

304 all $p < 0.01$). Figure 3C depicts a representative feedback-locked ERP demonstrating the

305 low-frequency FRN. This intracranial FRN was significant in the low frequency averaged

306 LFP across subjects (Kruskal-Wallis one-way ANOVA, $p << 0.01$, df = 71 contacts).

307 There was a main effect of feedback valence in 32 of 36 medial electrodes and 32 of 36

308 lateral electrodes (Tukey-Kramer, all $p = 0.0013$, N = 36 contacts each), indicating that

309 the FRN commonly observed on scalp EEG is evident in intracranial recordings. Low

310 frequency FRN was significantly greater in medial (dACC) contacts than lateral (dlPFC)

311 contacts (Mann-Whitney U test, $p = 0.0041$, N = 32 contacts) (Figure 3E).

312   Given the relationship between high-gamma ECoG signals and local neuronal

313 processing (Manning et al., 2009; Miller et al., 2009; Ray and Maunsell, 2011; Buzsáki et

314 al., 2012; Einevoll et al., 2013), we also examined the high-gamma representation of the

315 FRN in dACC and dlPFC (Figure 3D). Feedback significantly affected high-gamma

316 responses across subjects (Kruskal-Wallis one-way ANOVA, $p << 0.01$, df = 71

317 contacts). Within subjects, and across trials, there were also significant differences in

318 evoked high-gamma power between correct and neutral feedback trials in 34 of 36 medial

319 electrodes and 22 of 36 lateral contacts. As with low frequency potentials, a medial-to-

14

320   lateral decreasing gradient in FRN amplitude was observed in high-gamma responses

321   (Mann-Whitney U test, p = 0.00096, N = 22 contacts) (Figure 3F). Furthermore, latency

322   of the FRN response was shorter in medial compared to lateral contacts (two-sample t-

323   test, p = 0.02, N = 22 contacts) (Figure 3G). Cross-correlation of medial and lateral high

324   gamma signals corroborated this result by estimating the mean latency as 79 ms over

325   patients (Figure 3H). These results indicate that the FRN is likely generated medially, in

326   the dACC.

327

328   **High and low frequency LFPs are functionally related, more so in medial PFC**

329   We next examined coupling between low frequency (2 – 25 Hz) LFP phase and high

330   frequency (70 – 125 Hz) LFP amplitude. We found significant phase-amplitude coupling

331   (PAC) between theta phase and high gamma amplitude in all subjects (Kolmogorov-

332   Smirnov goodness-of-fit test, all p < 0.01, N = 50 contacts each, corrected for FDR due to

333   23 null hypotheses) (Figure 4A, B). The peak frequency-for-phase was 5.1 $\pm$0.2 Hz for

334   correct feedback trials and 4.8$\pm$0.4 Hz for neutral feedback. There was no significant

335   difference in best frequency-for-phase between medial and lateral contacts (Kruskal-

336   Wallis one-way ANOVA, p = 0.27, dF = 71 contacts). The magnitude of the theta-high

337   gamma coupling, as measured by the modulation index, was significantly greater for

338   correct feedback trials than for neutral feedback trials in medial (Tukey-Kramer, p =

339   0.0052, N = 36 contacts), but not in lateral (Tukey-Kramer, p = 0.0598, N = 36 contacts)

340   contacts (Figure 4C). In addition, the difference in theta-high gamma coupling between

341   correct and neutral trials was greater in medial contacts than lateral contacts (Tukey-

342   Kramer, p = 0.0098, N = 36 contacts) (Figure 4E). These results tie the local activity of

15

343   the dACC and PFC to the theta rhythm that can be observed on scalp EEG and show that

344   feedback information is also represented in the strength of theta-gamma coupling in

345   dACC.

346

347   **Information transfer between dACC and dlPFC**

348   The above results provide support for the hypothesis that the FRN is generated in the

349   dACC, and that modulation of theta-gamma coupling in the dACC encodes feedback-

350   related information. We next sought to determine the functional relationship between

351   feedback signals across dACC and dlPFC, hypothesizing that these signals propagate

352   from dACC to dlPFC.

353   To do so, we first used Granger causality (GC) techniques to study temporal

354   correlations between dACC and dlPFC activty. To account for the non-stationarity of

355   ECoG data, we compressed response-aligned broadband signals on all electrodes. Both

356   medial-to-lateral and lateral-to-medial interactions were significant (χ^2-Test, both p <

357   0.01, df = 6, corrected for 12 hypotheses using the FDR), although GC was

358   insignificantly greater for medial-to-lateral interactions (Figure 5A). Peak spectral

359   pairwise-conditional causalities were greatest in the low theta range (mean±std =

360   4.44±0.67 Hz) (Figure 5B). These results suggest that there are reciprocal interactions

361   between dACC and dlPFC.

362   To avoid contending with the violations of GC assumptions typical of ECoG data,

363   we further employed an analysis insensitive to these assumptions. We computed

364   information transfer between medial and lateral PFC using a conditional mutual

365   information (CMI) approach. CMI is a single-trial measure of the reduction in uncertainty

16

366   about the variable of interest obtained by observing the current trial's neural data from

367   one contact given that one knows the response on another contact. In our case, CMI is a

368   relative measure of the dependence of feedback valence representation in one cortical

369   area on the other cortical area. Lateral-to-medial information transfer for each subject

370   was not significantly greater than shuffled data after feedback presentation (χ^2 test, all p

371   > 0.01) and did not oscillate above a constant model (F-test, F = 0.288, p = 0.83) (Figure

372   6). However, CMI was significantly greater than shuffled data for medial-to-lateral

373   information transfer for each subject (χ^2 test, all p < 0.01), and oscillated (F-test, F =

374   36.50, p << 0.01) in the theta range (mean±std = 5.1±2.4 Hz) (Figure 6A,B). These

375   results support the notion of a largely unidirectional dACC-to-dlPFC transfer of

376   feedback-related information that oscillates with theta-range periodicity.

377

Discussion

379   We examined the neural representation of prefrontal feedback signals using human ECoG

380   with simultaneous medial and lateral PFC recordings. We report four main findings: 1)

381   Low and high frequency ECoG signals from human PFC are sensitive to feedback

382   valence. 2) These signals are larger and arise earlier in dACC compared to dlPFC. 3)

383   Spectral analysis demonstrated that theta-gamma coupling modulation underlies these

384   signals, and is also greater in dACC. 4) Information transfer analyses showed that

385   feedback-related information is transferred from dACC to dlPFC with theta periodicity.

386   These data thus provide direct evidence that the dACC is the cortical source of the FRN,

387   and that theta modulation serves as a mechanism for communication of feedback-related

388   information between dACC and lateral PFC.

17

389

FRN localization: Direct support for a dACC source

391 Previous EEG source localization studies have posited various sites of origin for the

392 FRN. Whereas the dACC is a common finding, others include the posterior cingulate

393 cortex and basal ganglia (reviewed in (Walsh and Anderson, 2012)) All these structures

394 receive input from midbrain dopaminergic cells, making them candidate regions for

395 producing RPE signals. Source modeling of EEG has been helpful, but the requirement of

396 solving the inverse problem from peripherally (scalp) recorded data makes identification

397 of a single source difficult.

398   Our direct human intracranial recordings demonstrated the FRN at the level of

399 individual electrode contacts in both low and high frequency ranges. Although signals

400 were seen in both dACC and lateral PFC, those from dACC were larger and had shorter

401 latencies. Human ECoG recordings have their own limitations: most importantly, that

402 electrodes can only be placed based on clinical grounds. Because the basal ganglia are

403 not a typical onset site for seizures, they are seldom placed there for epilepsy

404 investigations. On the other hand, electrodes are frequently placed into the subthalamic

405 nucleus and globus pallidus internus during deep brain stimulation procedures to treat

406 movement disorders such as Parkinson disease and dystonia(Zavala et al., 2013; 2014).

407 Investigation of feedback responses in these nuclei in humans is therefore possible, and

408 will be the subject of future work.

409   A recent combined EEG-fMRI study modeled the fMRI response informed by the

410 EEG and also identified the dACC as the FRN source (Hauser et al., 2013). Their

411 dynamic causal modeling analysis also showed that FRN signals were not arriving to

18

412   dACC from other structures. Thus taking the current study and previous studies together,

413   there seems little doubt that the dACC is the source of the FRN.

414

415   **FRN dynamics: theta modulation**

416   There is increasing evidence that the FRN and similar potentials, such as the conflict-

417   elicited N2 and the error-related negativity, share a physiological basis, namely

418   modulation of theta power (Holroyd et al., 2004a; Cavanagh and Frank, 2014). This

419   commonality has lead some to propose the existence of an interrelated group of cognitive

420   control signals known as frontal midline theta (FMΘ). FMΘ is an increase in theta-range

421   (4-7 Hz) spectral power observed in the event-related potentials recorded on frontocentral

422   EEG contacts. This group of FMΘ signals is thought to employ theta modulation as a

423   means of communicating a certain class of information with other brain regions. FMΘ

424   may be a signature of dACC activation in response to threats, either corporeal or

425   cognitive, in order to signal the need for behaviorally relevant attentional control

426   (Cavanagh and Shackman, 2014).

427   The current data show that the amplitude of high gamma activity, thought to be an

428   indicator of local neuronal processing (Miller et al., 2009; Buzsáki et al., 2012), coupled

429   to the phase of low frequency oscillations specifically in the theta band. The modulation

430   of this phase-amplitude coupling was sensitive to feedback valence, and was stronger in

431   dACC than in dlPFC. Furthermore, our information transfer analyses both showed peaks

432   in the theta range. The linear, parametric analysis, GC, indicated reciprocal interaction

433   between dACC and dlPFC that was dominated by the theta range. In the context of the

434   current study, GC has some caveats. First, that we applied a compressor to the signal in

19

435   order to achieve stationarity and second, that the duration of the responses, especially

436   those high frequency responses lasting longer than 500 milliseconds, could bias the

437   analyses. Conditional mutual information analysis, which is nonlinear and nonparametric,

438   showed that feedback information arising in dACC propagated to dlPFC, and information

439   from dlPFC did not propagate to dACC. Despite the inclusion of broadband LFP in this

440   these analyses, the spectral composition of this information transfer showed a prominent

441   peak centered at 5 Hz. All of these results support the contention that the dACC utilizes

442   theta range oscillations to communicate with other cognitive control centers.

443   Our data thus lend support to the importance of FMΘ signals in general, and to the

444   association of the FRN with this group in particular. The conception that FRN, and FMΘ,

445   originates in the dACC, yet is represented in dlPFC is also consistent with our results.

446   While some EEG studies have suggested that FMΘ arises through reciprocal activation of

447   medial and lateral prefrontal cortex (Asada et al., 1999), the intracranial

448   electrophysiology we present indicates a largely unidirectional, medial-to-lateral transfer

449   of feedback-related information. Another EEG study suggested that FMΘ is associated

450   with increased theta range synchrony among increasingly lateral scalp electrodes

451   (Cavanagh et al., 2010). However, theta power on these lateral scalp electrodes correlated

452   with neither prediction error nor behavioral adaptation (Cavanagh et al., 2010). Our

453   results indicating that dACC sends information to lateral prefrontal areas via a theta

454   rhythm could explain these EEG results.

455   Theta oscillations have also been observed in the dACC in other species.

456   Simultaneous single neuron and LFP recordings have enabled investigation into the

457   mechanism and function underlying theta modulation. Previous rat (Narayanan et al.,

20

458 2013) and monkey (Womelsdorf et al., 2010) studies have demonstrated that medial

459 PFC neuronal spiking synchronizes with theta rhythms, and the magnitude of this

460 coherence correlates with behavioral measures of cognitive control. These results argue

461 for a mechanism by which theta-band oscillations sharpen spiking output to particular

462 phase intervals within a theta cycle (Womelsdorf et al., 2010). This mechanism of

463 temporal focusing may synchronize distant regions within a cognitive control network

464 and promote more efficient spike-based communication between them.

465   Spike-phase coupling has been previously described as a means of integrating

466 brain-wide networks (Fries, 2005; Lakatos, 2005; Lakatos et al., 2009; Womelsdorf et al.,

467 2010). One salient example of low-frequency LFP phase coordinating neuronal

468 population activity in distributed networks showed that prefrontal neurons are sensitive to

469 distinct, yet diverse array of rhythms (Canolty et al., 2010). As in the current study, theta

470 band modulation has frequently been observed in PFC (Canolty et al., 2006; Voytek et

471 al., 2010; Miller et al., 2012; Zavala et al., 2014). Whether the dACC establishes this

472 rhythm, or it is entrained by another structure remains to be determined.

473   Theta rhythms are not unique to the dACC and to mechanisms of learning and

474 cognitive control. It is known in humans and rodents that action potentials commonly

475 couple to hippocampal theta rhythms, which are 4-8 Hz in the rat and 1-4 Hz in humans

476 (Jacobs, 2014), and that this coupling likely supports memory processes (Buzsáki, 2005).

477 Interestingly, action potentials couple preferentially to 4-6 Hz rhythms in the dACC in

478 humans (Jacobs et al., 2007), which is the same range in which functional differences in

479 phase amplitude coupling were observed in the current study. The medial-to-lateral CMI

480 presented here also took on rhythmicity in precisely the 4-6 Hz range. These results

21

481 support FMθ arising from an information transfer architecture, as has been proposed in

482 the rodent hippocampus (Mizuseki et al., 2009).

483

484 **Conclusions**

485

486 We have shown that feedback is associated with broadband ERPs and high gamma

487 responses in the dACC and dlPFC. We have also shown that these signals are coupled to

488 theta-band oscillations, and that information is conveyed from dACC to dlPFC entrained

489 to a theta rhythm. Our data support a novel model for the generation of FMθ associated

490 with feedback in human subjects, and pave the way for future experiments to test how

491 humans use FMθ signals to coordinate behavioral responses to feedback or other

492 indications for the need for cognitive control. Together these results suggest that

493 feedback-related information is initially processed in the dACC and quickly

494 communicated to lateral prefrontal cortex as part of a cognitive control signal (Sheth et

495 al., 2012; Shenhav et al., 2013). Lateral PFC presumably incorporates this information to

496 influence ongoing behavioral modifications relevant to the immediate context and long-

497 term goals (Miller and Cohen, 2001; Carter and van Veen, 2007)**.**

498

499 **References**

500

501 Andersson JL, Jenkinson M, Smith S (2013) Andersson: Non-linear registration, aka
502   spatial normalisa... - Google Scholar. TR07JA2 www fmrib ox ac uk analysis/ ….

503 Asada H, Fukuda Y, Tsunoda S, Yamaguchi M, Tonoike M (1999) Frontal midline theta
504   rhythms reflect alternative activation of prefrontal cortex and anterior cingulate
505   cortex in humans. Neurosci Lett 274:29–32.

506 Barnett L, Seth AK (2014) The MVGC multivariate Granger causality toolbox: a new
507   approach to Granger-causal inference. J Neurosci Methods 223:50–68.

508 Boatman-Reich D, Franaszczuk PJ, Korzeniewska A, Caffo B, Ritzl EK, Colwell S,
509   Crone NE (2010) Quantifying auditory event-related responses in multichannel
510   human intracranial recordings. Front Comput Neurosci 4:4.

511 Bressler SL, Seth AK (2011) Wiener–Granger Causality: A well established
512   methodology. NeuroImage.

513 Buzsáki G (2005) Theta rhythm of navigation: link between path integration and
514   landmark navigation, episodic and semantic memory. Hippocampus 15:827–840.

515 Buzsáki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and
516   currents--EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407–420.

517 Canolty RT, Edwards E, Dalal SS, Soltani M, Nagarajan SS, Kirsch HE, Berger MS,
518   Barbaro NM, Knight RT (2006) High gamma power is phase-locked to theta
519   oscillations in human neocortex. Science 313:1626–1628.

520 Canolty RT, Ganguly K, Kennerley SW, Cadieu CF, Koepsell K, Wallis JD, Carmena JM
521   (2010) Oscillatory phase coupling coordinates anatomically dispersed functional cell
522   assemblies. Proc Natl Acad Sci USA 107:17356–17361.

523 Carter CS, van Veen V (2007) Anterior cingulate cortex and conflict detection: an update
524   of theory and data. Cogn Affect Behav Neurosci 7:367–379.

525 Cavanagh JF, Frank MJ (2014) Frontal theta as a mechanism for cognitive control.
526   Trends Cogn Sci (Regul Ed) 18:414–421.

527 Cavanagh JF, Frank MJ, Klein TJ, Allen JJB (2010) Frontal theta links prediction errors
528   to behavioral adaptation in reinforcement learning. NeuroImage 49:3198–3209.

529 Cavanagh JF, Shackman AJ (2014) Frontal midline theta reflects anxiety and cognitive
530   control: Meta-analytic evidence. Journal of Physiology-Paris.

531 Chase HW, Swainson R, Durham L, Benham L, Cools R (2011) Feedback-related
532   Negativity Codes Prediction Error but Not Behavioral Adjustment during
533   Probabilistic Reversal Learning. J Cogn Neurosci 23:936–946.

534 Cohen MX, Ranganath C (2007) Reinforcement learning signals predict future decisions.
535   Journal of Neuroscience 27:371–378.

536 Doñamayor N, Marco-Pallarés J, Heldmann M, Schoenfeld MA, Münte TF (2011)
537   Temporal dynamics of reward processing revealed by magnetoencephalography.
538   Hum Brain Mapp 32:2228–2240.

539 Einevoll GT, Kayser C, Logothetis NK, Panzeri S (2013) Modelling and analysis of local

23

540   fieldpotentials for studying the function ofcortical circuits. Nat Rev Neurosci
541   14:770–785.

542   Fan J, Hung H-N, Wong W-H (2000) Geometric Understanding of Likelihood Ratio
543   Statistics. Journal of the American Statistical Association 95:836–841.

544   Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through
545   neuronal coherence. Trends Cogn Sci (Regul Ed) 9:474–480.

546   Granger CWJ (1969) Investigating Causal Relations by Econometric Models and Cross-
547   spectral Methods. Econometrica 37:424.

548   Hauser TU, Iannaccone R, Stämpfli P, Drechsler R, Brandeis D, Walitza S, Brem S
549   (2013) The feedback-related negativity (FRN) revisited: New insights into the
550   localization, meaning and network organization. NeuroImage 84C:159–168.

551   Herrmann MJ, Römmler J, Ehlis A-C, Heidrich A, Fallgatter AJ (2004) Source
552   localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe).
553   Cognitive Brain Research 20:294–299.

554   Hochberg Y, Tamhane AC (1987) Multiple comparison procedures. John Wiley & Sons,
555   Inc.

556   Holroyd CB, Coles MGH (2002a) The neural basis of human error processing:
557   Reinforcement learning, dopamine, and the error-related negativity. Psychol Rev
558   109:679–709.

559   Holroyd CB, Coles MGH (2002b) The neural basis of human error processing:
560   reinforcement learning, dopamine, and the error-related negativity. Psychol Rev
561   109:679–709.

562   Holroyd CB, Coles MGH, Nieuwenhuis S (2002) Medial prefrontal cortex and error
563   potentials. Science 296:1610–1authorreply1610–1.

564   Holroyd CB, Larsen JT, Cohen JD (2004a) Context dependence of the event-related brain
565   potential associated with reward and punishment. Psychophysiology 41:245–253.

566   Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom L, Mars RB, Coles MGH, Cohen JD
567   (2004b) Dorsal anterior cingulate cortex shows fMRI response to internal and
568   external error signals. Nature Publishing Group 7:497–498.

569   Holroyd CB, Pakzad-Vaezi KL, Krigolson OE (2008) The feedback correct-related
570   positivity: sensitivity of the event-related brain potential to unexpected positive
571   feedback. Psychophysiology 45:688–697.

572   Horga G, Maia TV, Wang P, Wang Z, Marsh R, Peterson BS (2011) Adaptation to
573   conflict via context-driven anticipatory signals in the dorsomedial prefrontal cortex.
574   Journal of Neuroscience 31:16208–16216.

575   Ince RAA, Mazzoni A, Bartels A, Logothetis NK, Panzeri S (2011) A novel test to
576   determine the significance of neural selectivity to single and multiple potentially
577   correlated stimulus features. J Neurosci Methods:1–17.

578   Jacobs J (2014) Hippocampal theta oscillations are slower in humans than in rodents:
579   implications for models of spatial navigation and memory. Philos Trans R Soc Lond,
580   B, Biol Sci 369:20130304.

581   Jacobs J, Kahana MJ, Ekstrom AD, Fried I (2007) Brain oscillations control timing of
582   single-neuron activity in humans. Journal of Neuroscience 27:3839–3844.

583   Jenkinson M, Pechaud M, Smith S (2005) Jenkinson: BET2: MR-based estimation of
584   brain, skull... - Google Scholar. … meeting of the organization for human brain ….

585   Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration
586   of brain images. Med Image Anal 5:143–156.

587   Jessup RK, Busemeyer JR, Brown JW (2010) Error Effects in Anterior Cingulate Cortex
588   Reverse when Error Likelihood Is High.

589   Keller K, Lauffer H (2003) Symbolic Analysis of High-Dimensional Time Series.
590   http://dxdoiorg/101142/S0218127403008168 13:2657–2668.

591   King J-R, Sitt JD, Faugeras F, Rohaut B, Karoui El I, Cohen L, Naccache L, Dehaene S
592   (2013) Information Sharing in the Brain Indexes Consciousnessin
593   Noncommunicative Patients. Curr Biol 23:1914–1919.

594   Lakatos P (2005) An Oscillatory Hierarchy Controlling Neuronal Excitability and
595   Stimulus Processing in the Auditory Cortex. J Neurophysiol 94:1904–1911.

596   Lakatos P, O'Connell MN, Barczak A, Mills A, Javitt DC, Schroeder CE (2009) The
597   leading sense: supramodal control of neurophysiological context by attention. Neuron
598   64:419–430.

599   Li X, Ouyang G (2010) Estimating coupling direction between neuronal populations with
600   permutation conditional mutual information. NeuroImage 52:497–507.

601   Luu P, Tucker DM, Derryberry D, Reed M, Poulsen C (2003) Electrophysiological
602   Responses to Errors and Feedback in the Process of Action Regulation.
603   Psychological Science 14:47–53.

604   Luu P, Tucker DM, Makeig S (2004) Frontal midline theta and the error-related
605   negativity: neurophysiological mechanisms of action regulation. Clinical
606   Neurophysiology 115:1821–1835.

607   Manning JR, Jacobs J, Fried I, Kahana MJ (2009) Broadband Shifts in Local Field
608   Potential Power Spectra Are Correlated with Single-Neuron Spiking in Humans.
609   Journal of Neuroscience 29:13613–13620.

25

610 Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. Annu
611   Rev Neurosci 24:167–202.

612 Miller KJ, Hermes D, Honey CJ, Hebb AO, Ramsey NF, Knight RT, Ojemann JG, Fetz
613   EE (2012) Human motor cortical activity is selectively phase-entrained on underlying
614   rhythms. PLoS Comput Biol 8:e1002655.

615 Miller KJ, Hermes D, Honey CJ, Sharma M, Rao RPN, Nijs den M, Fetz EE, Sejnowski
616   TJ, Hebb AO, Ojemann JG, Makeig S, Leuthardt EC (2010) Dynamic modulation of
617   local population activity by rhythm phase in human occipital cortex during a visual
618   search task. Front Hum Neurosci 4:197.

619 Miller KJ, Sorensen LB, Ojemann JG, Nijs den M (2009) Power-Law Scaling in the
620   Brain Surface Electric Potential Sporns O, ed. PLoS Comput Biol 5:e1000609.

621 Miltner WH, Braun CH, Coles MG (1997) Event-related brain potentials following
622   incorrect feedback in a time-estimation task: evidence for a "generic" neural system
623   for error detection. J Cogn Neurosci 9:788–798.

624 Miltner WHR, Lemke U, Weiss T, Holroyd C, Scheffers MK, Coles MGH (2003)
625   Implementation of error-processing in the human anterior cingulate cortex: a source
626   analysis of the magnetic equivalent of the error-related negativity. Biol Psychol
627   64:157–166.

628 Mizuseki K, Sirota A, Pastalkova E, Buzsáki G (2009) Theta Oscillations Provide
629   Temporal Windows for Local Circuit Computationin the Entorhinal-Hippocampal
630   Loop. Neuron 64:267–280.

631 Narayanan NS, Cavanagh JF, Frank MJ, Laubach M (2013) Common medial frontal
632   mechanisms of adaptive control in humans and rodents. Nat Neurosci 16:1888–1895.

633 Nieuwenhuis S, Heslenfeld DJ, Geusau von NJA, Mars RB, Holroyd CB, Yeung N
634   (2005a) Activity in human reward-sensitive brain areas is strongly context
635   dependent. NeuroImage 25:1302–1309.

636 Nieuwenhuis S, Slagter HA, Geusau von NJA, Heslenfeld DJ, Holroyd CB (2005b)
637   Knowing good from bad: differential activation of human cortical areas by positive
638   and negative outcomes. Eur J Neurosci 21:3161–3168.

639 Pfabigan DM, Alexopoulos J, Bauer H, Sailer U (2010) Manipulation of feedback
640   expectancy and valence induces negative and positive reward prediction error signals
641   manifest in event-related brain potentials. Psychophysiology 48:656–664.

642 Quinn CJ, Coleman TP, Kiyavash N, Hatsopoulos NG (2010) Estimating the directed
643   information to infer causal relationships in ensemble neural spike train recordings. J
644   Comput Neurosci 30:17–44.

645 Ray S, Maunsell JHR (2011) Different origins of gamma rhythm and high-gamma

646   activity in macaque visual cortex. PLoS Biol 9:e1000610.

647   Rodrigues J, Andrade A (2014) Lag-based effective connectivity applied to fMRI: a
648   simulation study highlighting dependence on experimental parameters and
649   formulation. NeuroImage 89:358–377.

650   Roger C, Bénar CG, Vidal F, Hasbroucq T, Burle B (2010) Rostral Cingulate Zone and
651   correct response monitoring: ICA and source localization evidences for the unicity of
652   correct- and error-negativities. NeuroImage 51:391–403.

653   Salvador R, Anguera M, Gomar JJ, Bullmore ET, Pomarol-Clotet E (2010) Conditional
654   mutual information maps as descriptors of net connectivity levels in the brain. Front
655   Neuroinform 4:115.

656   Schultz W (2002) Getting formal with dopamine and reward. Neuron 36:241–263.

657   Schultz W (2013) Updating dopamine reward signals. Curr Opin Neurobiol 23:229–238.

658   Seth AK (2010) A MATLAB toolbox for Granger causal connectivity analysis. J
659   Neurosci Methods 186:262–273.

660   Seth AK, Edelman GM (2007) Distinguishing Causal Interactions in Neural Populations.
661   Neural Comput 19:910–933.

662   Shenhav A, Botvinick MM, Cohen JD (2013) The expected value of control: an
663   integrative theory of anterior cingulate cortex function. Neuron 79:217–240.

664   Sheth SA, Mian MK, Patel SR, Asaad WF, Williams ZM, Dougherty DD, Bush G,
665   Eskandar EN (2012) Human dorsal anterior cingulate cortex neurons mediate
666   ongoing behavioural adaptation. Nature 488:218–221.

667   Silvetti M, Nuñez Castellar E, Roger C, Verguts T (2014) Reward expectation and
668   prediction error in human medial frontal cortex: An EEG study. NeuroImage 84:376–
669   382.

670   Smith E, Duede S, Hanrahan S, Davis T, House P, Greger B (2013) Seeing Is Believing:
671   Neural Representations of Visual Stimuli in Human Auditory Cortex Correlate with
672   Illusory Auditory Perceptions Zhang LI, ed. PLoS ONE 8:e73148.

673   Smith EH, Kellis SS, House PA, Greger B (2012) Information transfer along the ventral
674   auditory processing stream in the awake macaque. Conf Proc IEEE Eng Med Biol
675   Soc 2012:5178–5181.

676   Voytek B, Canolty RT, Shestyuk A, Crone NE, Parvizi J, Knight RT (2010) Shifts in
677   gamma phase-amplitude coupling frequency from theta to alpha over posterior cortex
678   during visual tasks. Front Hum Neurosci 4:191.

679   Walsh MM, Anderson JR (2012) Learning from experience: Event-related potential

680   correlates of reward processing, neural adaptation, and behavioral choice.
681   Neuroscience and Biobehavioral Reviews 36:1870–1884.

682 Walsh MM, Anderson JR (2013) Electrophysiological Responses to Feedback during the
683   Application of Abstract Rules. J Cogn Neurosci 25:1986–2002.

684 Womelsdorf T, Johnston K, Vinck M, Everling S (2010) Theta-activity in anterior
685   cingulate cortex predicts task rules and their adjustments following errors. Proc Natl
686   Acad Sci USA 107:5248–5253.

687 Zavala B, Brittain JS, Jenkinson N, Ashkan K, Foltynie T, Limousin P, Zrinzo L, Green
688   AL, Aziz T, Zaghloul K, Brown P (2013) Subthalamic Nucleus Local Field Potential
689   Activity during the Eriksen Flanker Task Reveals a Novel Role for Theta Phase
690   during Conflict Monitoring. Journal of Neuroscience 33:14758–14766.

691 Zavala BA, Tan H, Little S, Ashkan K, Hariz M, Foltynie T, Zrinzo L, Zaghloul KA,
692   Brown P (2014) Midline Frontal Cortex Low-Frequency Activity Drives Subthalamic
693   Nucleus Oscillations during Conflict. Journal of Neuroscience 34:7322–7333.

694 Zunino L, Soriano MC, Fischer I, Rosso OA (2010) Phys. Rev. E 82, 046212 (2010) -
695   Permutation-information-theory approach to unveil delay dynamics from time-series
696   analysis. Physical Review E.

697

698

699 **Figure Legends**

700

701 **Figure 1. Task description and behavioral performance. A,** MSIT task

702 description and timeline. A fixation cross appears on the screen 500 milliseconds

703 before stimuli are shown. Four labeled examples of the different types of

704 interference are shown above the stimulus period in the task timeline. Feedback

705 occurs following the subject's response to the stimulus. In valenced feedback trials,

706 the target number turns green for correct responses and red for incorrect

707 responses. In neutral feedback trials, the target number turns blue regardless of

708 whether the subject responded correctly. Examples of the four possible types of

28

709   feedback are shown for the example stimulus that contains both types of

710   interference. **B,** Distributions of RTs for each subject colored by subject. **C,** Mean RTs

711   across subjects for trials containing the four different types of interference. The

712   asterisk indicates statistical significance over conditions for the ANOVA. **D,** Mean

713   RTs across subjects for trials following correct and neutral feedback. "n.s." indicates

714   no significant difference.

715

716   **Figure 2. Depth Electrode locations in seven subjects.**

717   The locations of the medial and lateral–most contacts on each depth electrode

718   projected onto the medial and lateral surfaces of each hemisphere of the MNI

719   standard brain. Colors represent individual subjects.

720

721   **Figure 3. FRN in high and low frequency LFP. A,** Response aligned spectrograms

722   averaged over trials (mean$\pm$S.D. = 289$\pm$65.32) and patients. Frequency bands

723   during the outcome period that are significantly different between feedback

724   conditions arte outlined with gray boxes (paired t-tests, corrected for 72 hypotheses, all

725    $p<0.01$). **B,** Response aligned inter-trial phase coherence averaged over trials and

726   patients. **C,** Representative event related potentials from one subject for medial and

727   lateral electrode contacts. Green and blue lines represent correct and neutral

728   feedback conditions, respectively, and black lines represent the difference. Shading

729   represents standard errors. **D,** Representative evoked high gamma response from

730   one subject. **E,** Quantification of FRN amplitude from ERP (Mann-Whitney U test, $p =$

731   0.0041, N = 32 contacts). **F,** Quantification of high gamma power for each feedback

29

732   condition (Mann-Whitney U test, p = 0.00096, N = 22 contacts). **G,** Quantification of

733   latency differences between medial and lateral high gamma responses (two-sample

734   t-test, p = 0.02, N = 22 contacts). Asterisks indicate significant differences (p<0.05). **H,**

735   Lagged cross-correlogram between medial and lateral high gamma responses. The

736   peak indicated that mean high gamma responses in lateral contacts lagged those in

737   medial contacts by 79 ms.

738

739   **Figure 4. Phase-amplitude coupling (PAC) of low and high frequency LFPs. A,**

740   **B,** Representative PAC for a medial (**A**) and lateral (**B**) contact in one subject for

741   each feedback condition (correct and neutral). **C, D,** Plots of the real (x axis) and

742   imaginary (y axis) components of the modulation indices over trials for a

743   representative patient for medial (**C**) and lateral (**D**) contacts. Blue dots represent

744   the modulation indices for neutral trials and green dots represent the modulation

745   indices for correct trials. **E,** Mean modulation indices across patients for medial and

746   lateral contacts. Green and blue bars represent correct and neutral feedback trials,

747   respectively. Asterisks indicate significant differences (p<0.01). Modulation indices

748   were greater for correct feedback trials than neutral feedback trials in medial (Tukey-

749   Kramer, p = 0.0052, N = 36 contacts), but not in lateral (Tukey-Kramer, p = 0.0598, N =

750   36 contacts) contacts. The difference in theta-high gamma coupling between correct and

751   neutral trials was greater in medial contacts than lateral contacts (Tukey-Kramer, p =

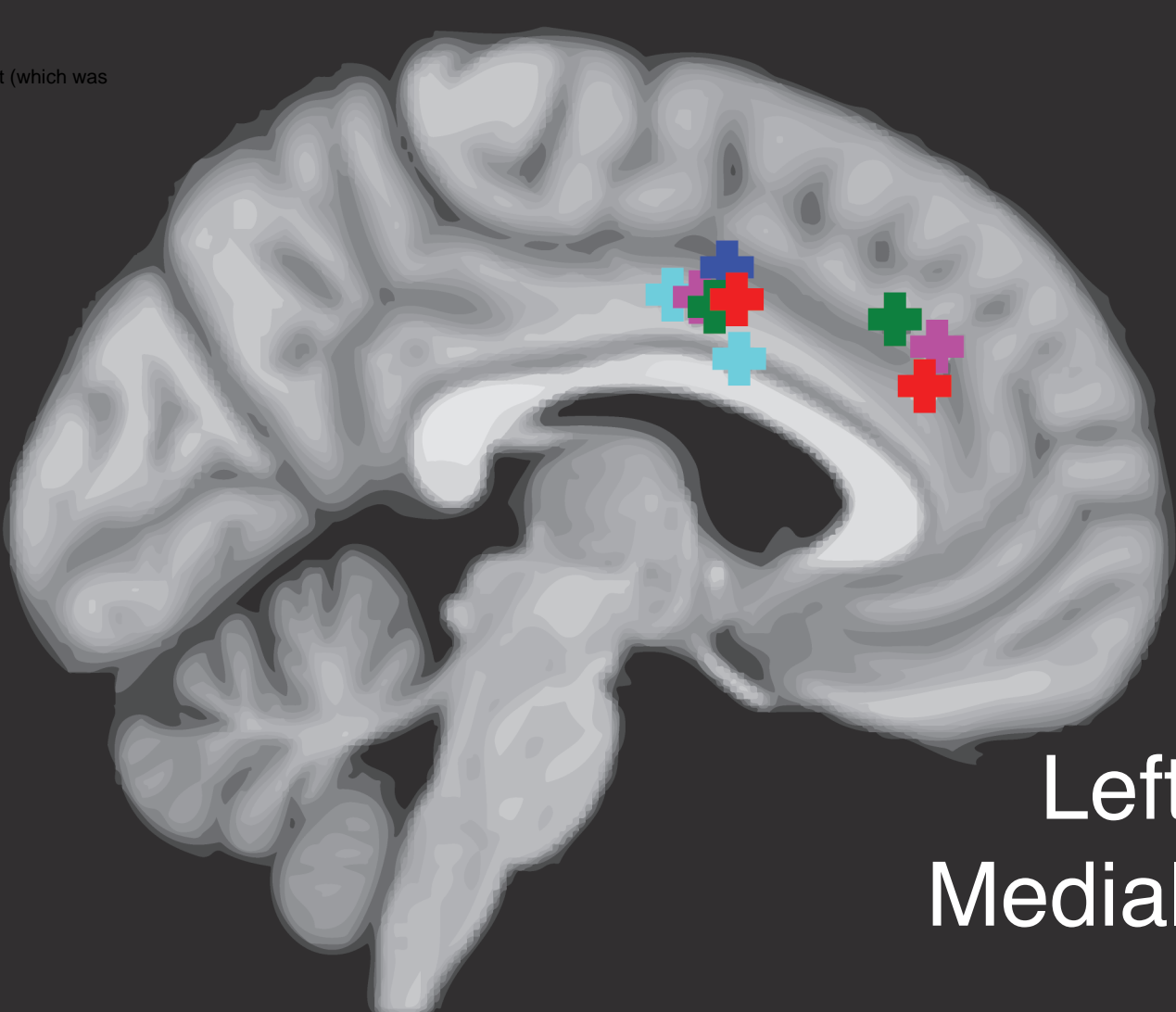752   0.0098, N = 36 contacts).

753

754 **Figure 5. Pairwise conditional Granger Causality between dACC and dlPFC. A,**

755 Pairwise GC over frequencies for the lateral-to-medial direction. Mean GC values

756 across patients are shown in dark bold orange, and patient specific data are shown

757 in light orange. Peak frequencies averaged over patients are indicated in the colored

758 text in the plot. **B,** Pairwise GC over frequencies for the medial-to-lateral direction.

759 Mean GC values across patients are shown in dark bold blue, and patient specific

760 data are shown in light blue. Peak frequencies averaged over patients are indicated

761 in the colored text in the plot.

762

763 **Figure 6. Conditional mutual information exchange between dACC and dlPFC.**

764 **A,** Mean CMI across subjects after feedback. Pink lines represent medial-to-lateral

765 information transfer, and brown lines represent lateral-to-medial transfer. Gray

766 lines represent shuffled data. Dashed and dotted lines show sinusoidal fits for

767 medial-to-lateral and lateral-to-medial CMI, respectively. Vertical line represents

768 feedback onset. **B,** Mean spectra of the CMI time series show increased theta power
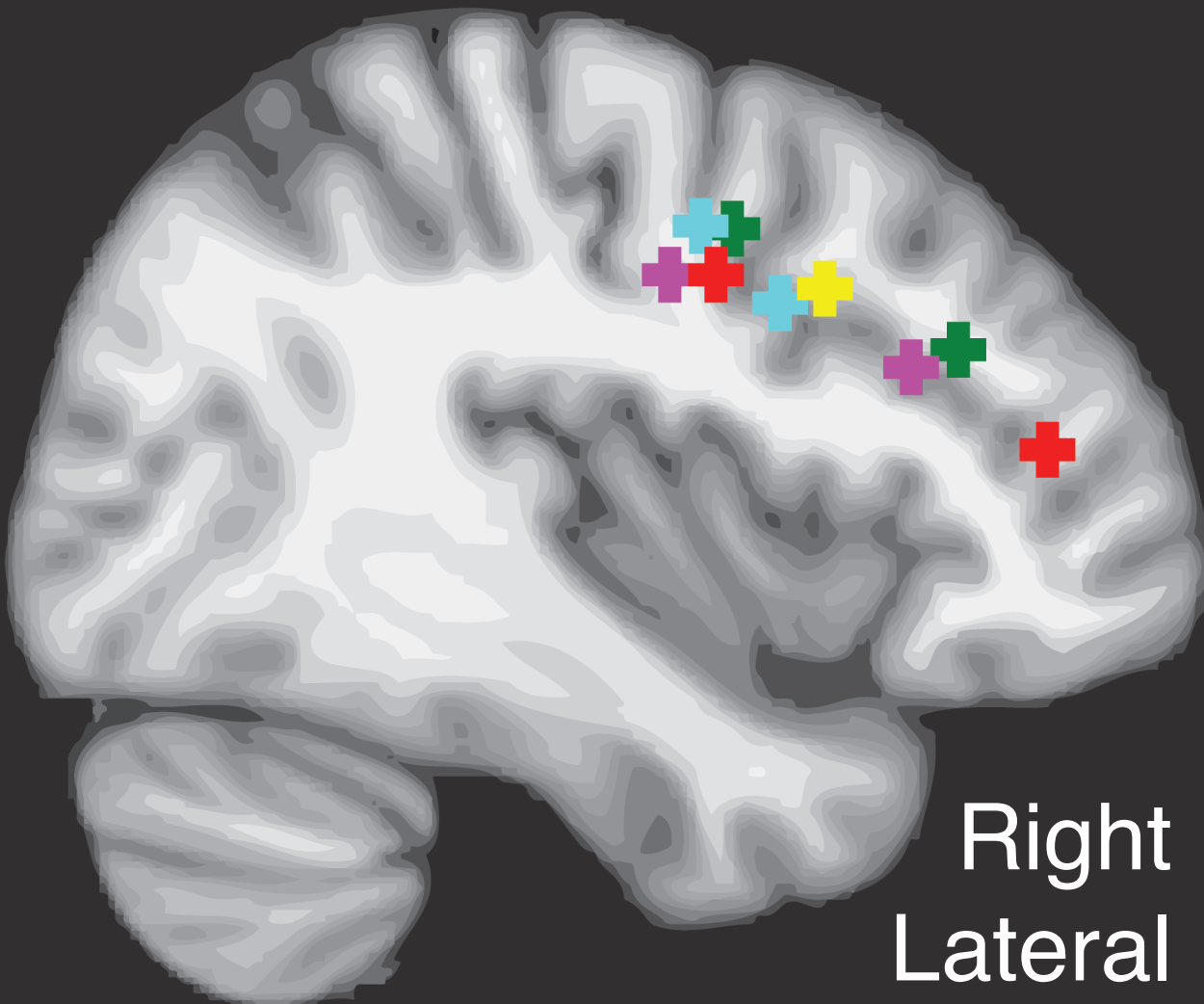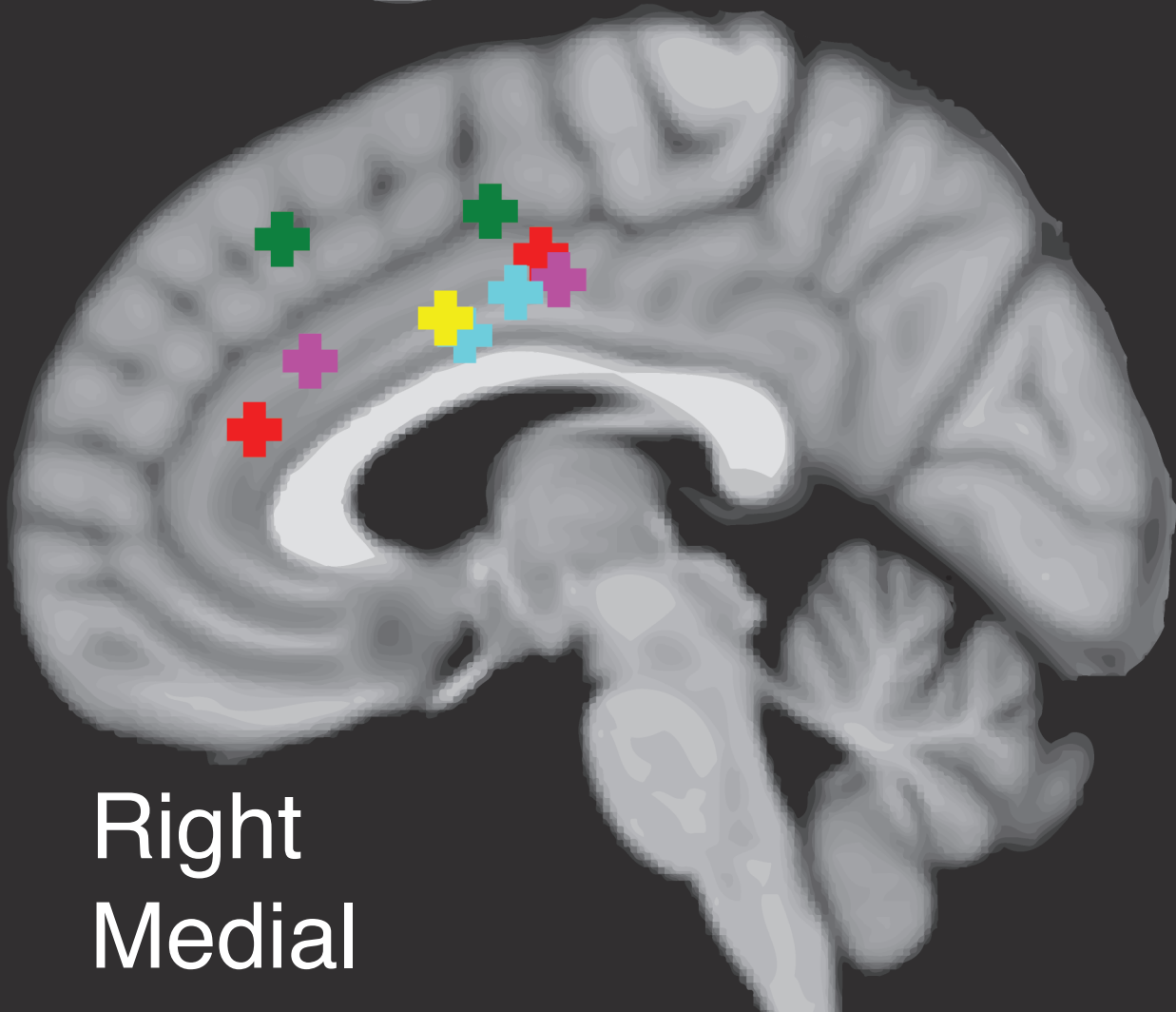
769 for medial-to-lateral CMI.
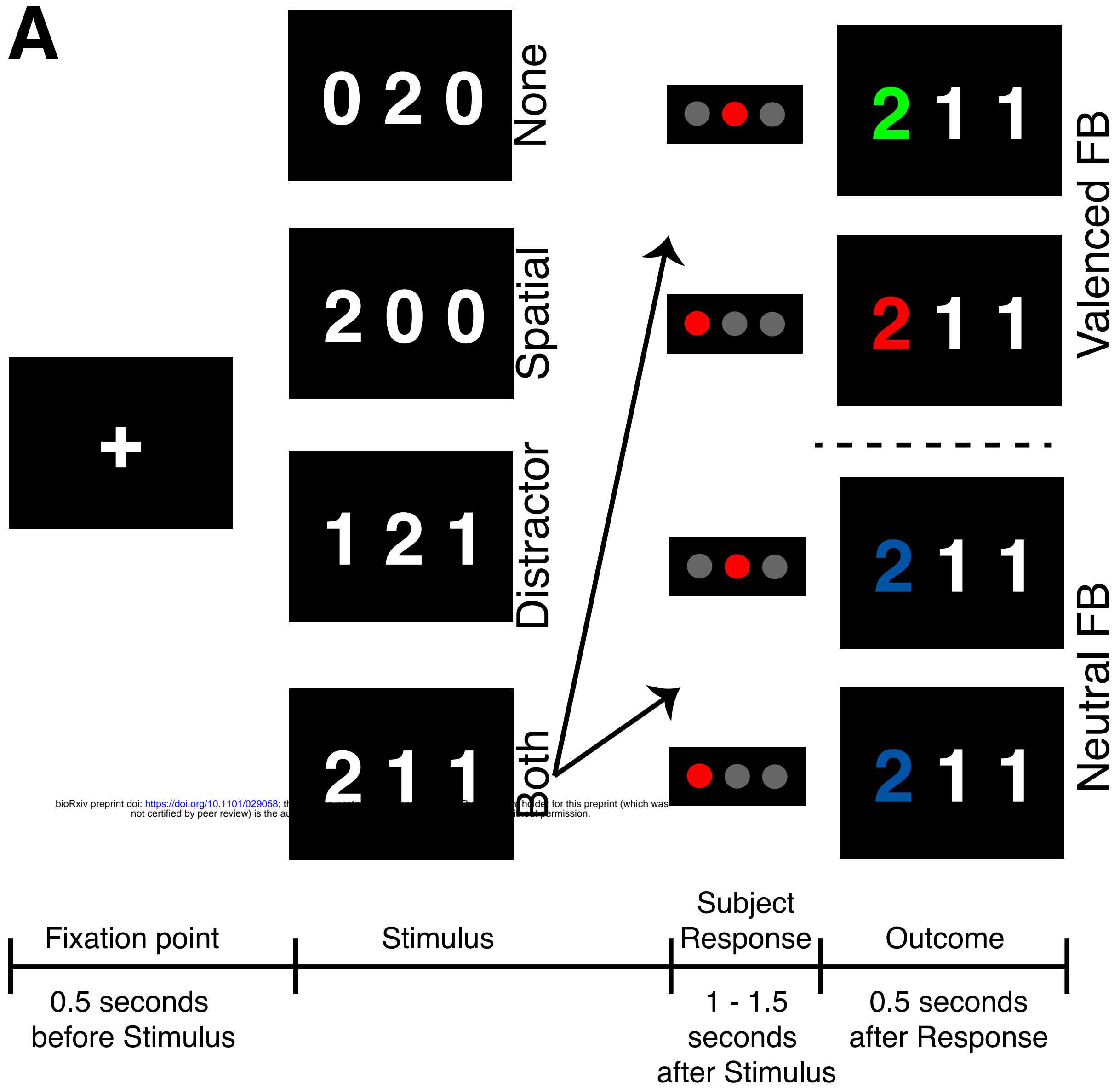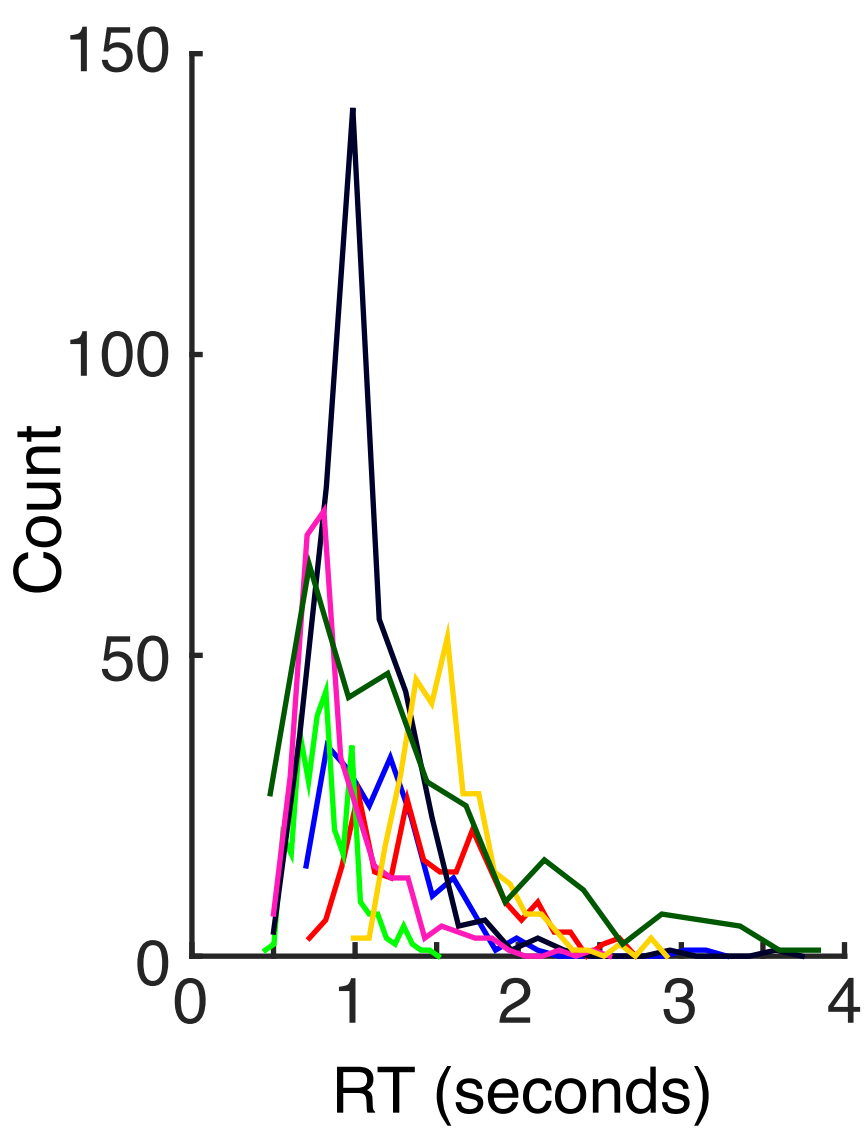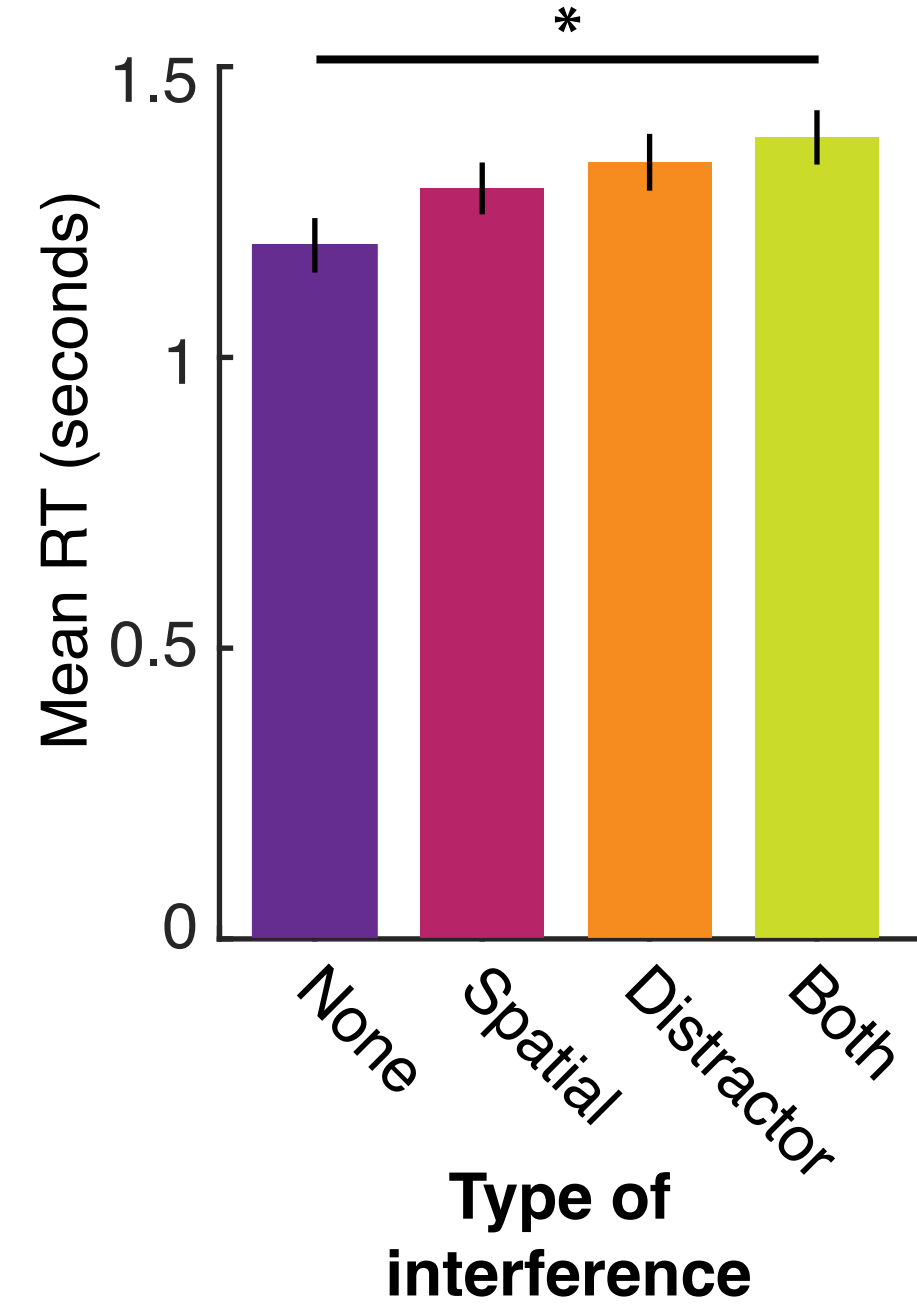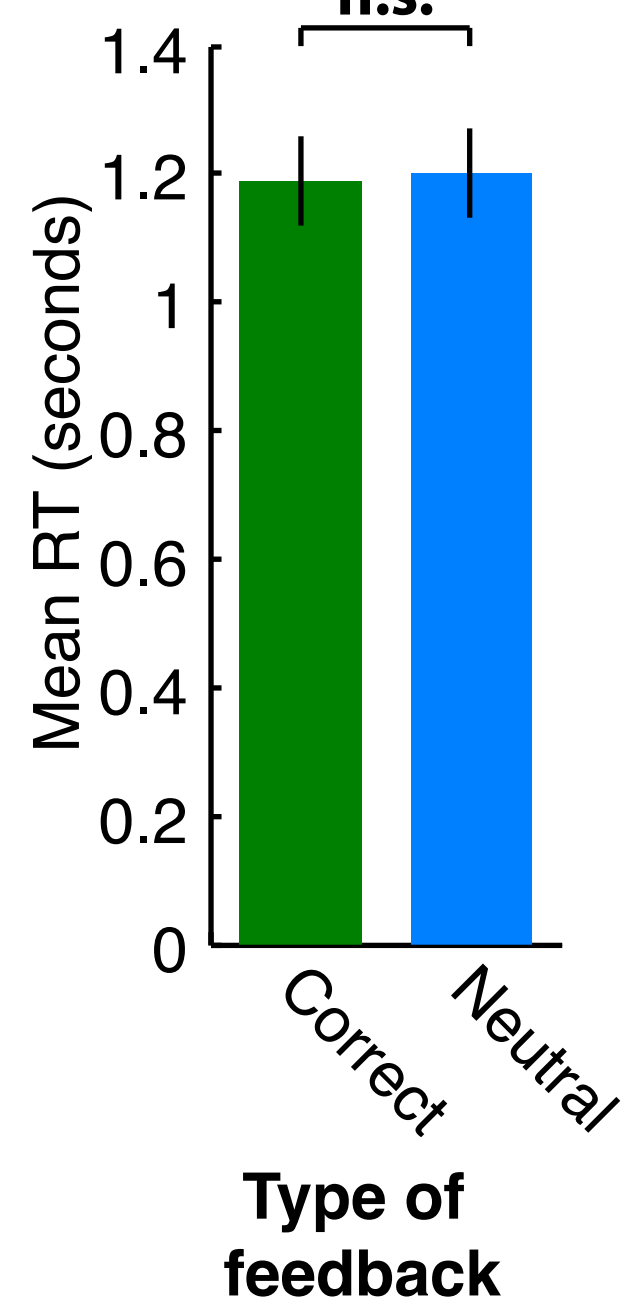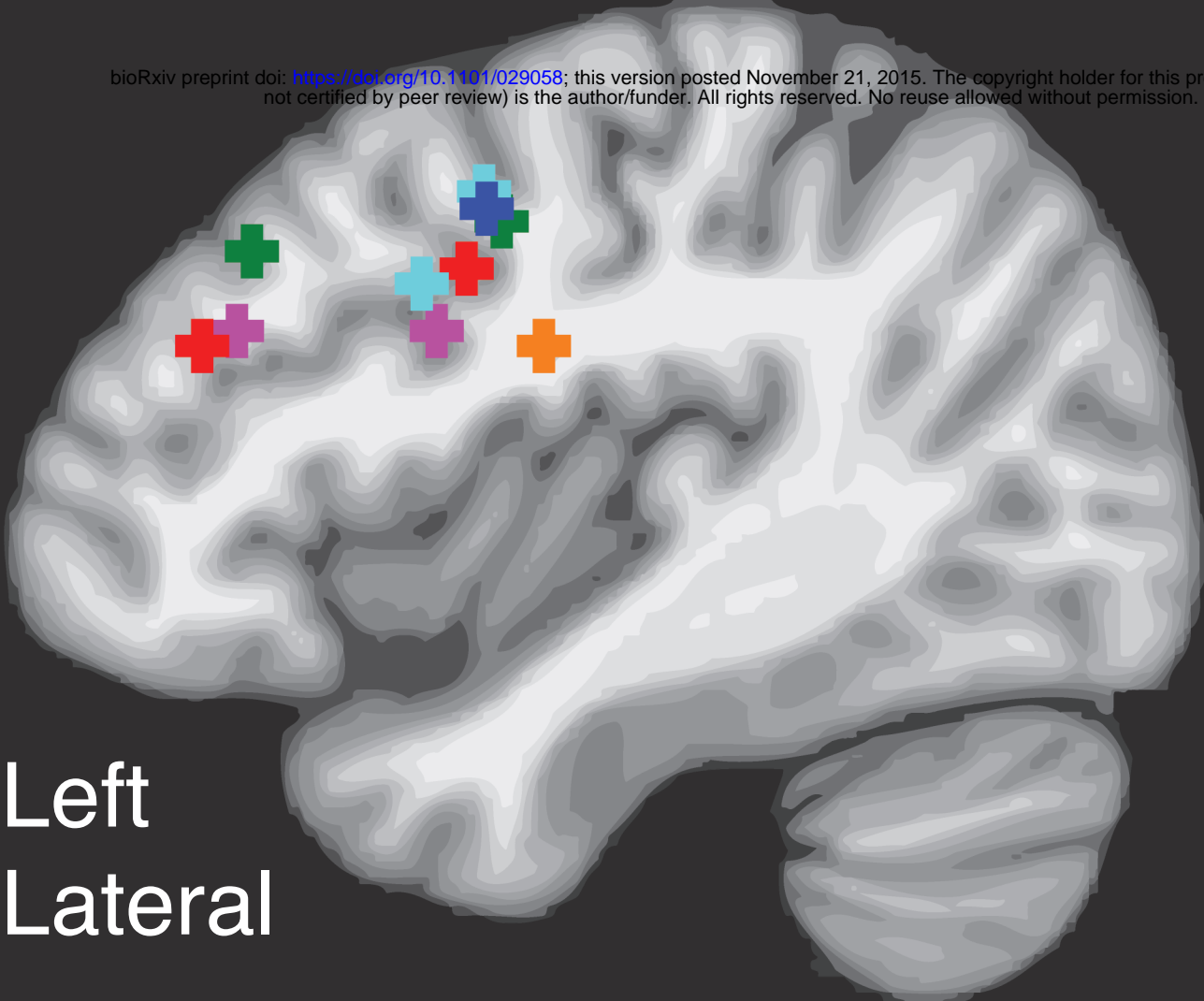
Left Lateral

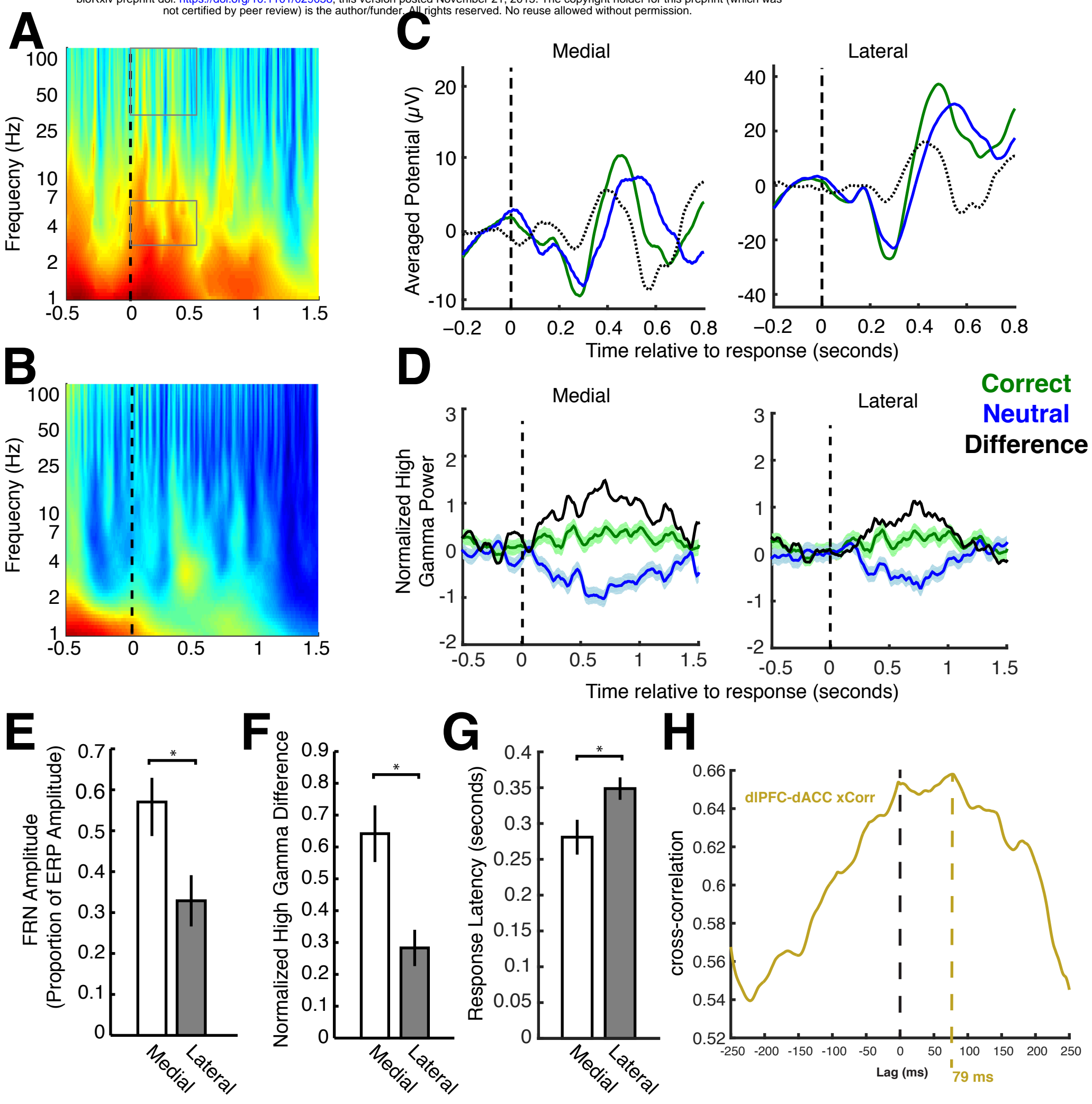Left Medial

Right Lateral
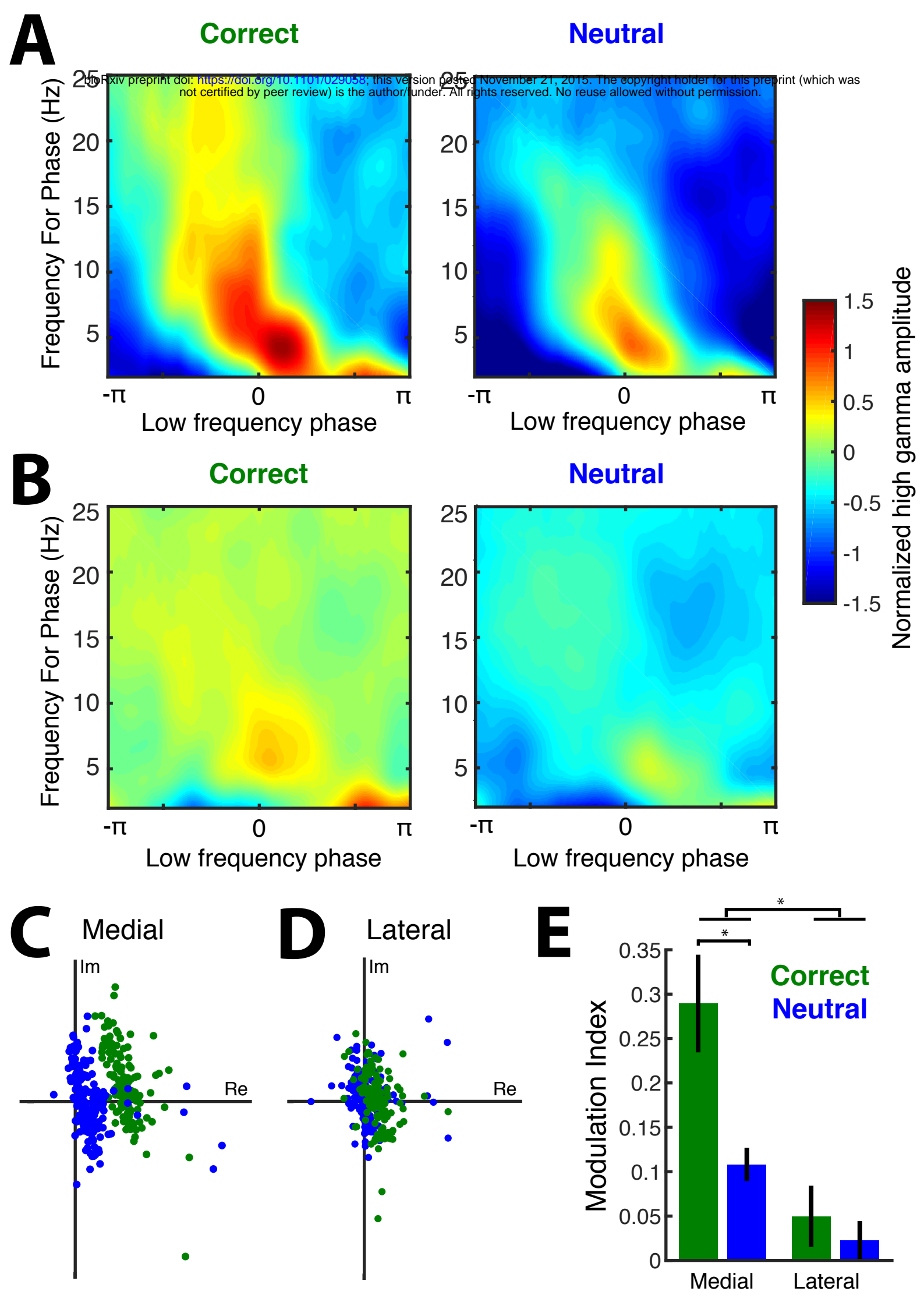
Right Medial

A dlPFC ⟶ dACC

Pairwise GC

4.91 Hz

0.3

0.2

0.1

0   50

Frequency (Hz)

B dACC ⟶ dlPFC

Pairwise GC

3.97 Hz

0.3

0.2

0.1

0   50

Frequency (Hz)

A

CMI (bits) vs Time relative to feedback (seconds)

B

Power (dB) vs Frequency (Hz)

mPFC-> lPFC
lPFC -> mPFC
shuffled data