1   **Cascade of neural processing orchestrates cognitive control in human frontal cortex**

2

3   Hanlin Tang[1,2], Hsiang-Yu Yu[3,4], Chien-Chen Chou[3,4], Nathan E. Crone[5], Joseph R.

4   Madsen[6], William S. Anderson[7], Gabriel Kreiman[1,2,8,*]

5

6   [1] Program in Biophysics, Harvard University, Boston, MA, 02115, USA

7   [2] Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School,

8   Boston, MA, 02115, USA

9   [3] Department of Neurology, Taipei Veterans General Hospital, Taipei, Taiwan

10   [4] National Yang-Ming University, Taipei, Taiwan

11   [5] Department of Neurology, Johns Hopkins School of Medicine, Baltimore, MD, 21287,

12   USA

13   [6] Department of Neurosurgery, Boston Children's Hospital, Harvard Medical School,

14   Boston, MA, 02115, USA

15   [7] Department of Neurosurgery, Johns Hopkins School of Medicine, Baltimore, MD,

16   21287, USA

17   [8] Center for Brain Science, Harvard University, Boston, MA, 02115, USA;

18

19   * Corresponding author: gabriel.kreiman@childrens.harvard.edu

20   **Abstract**

21   Rapid and flexible interpretation of conflicting sensory inputs in the context of current

22   goals is a critical component of cognitive control that is orchestrated by frontal cortex.

23   The relative roles of distinct subregions within frontal cortex are poorly understood. To

24   examine the dynamics underlying cognitive control across frontal regions, we took

25   advantage of the spatiotemporal resolution of intracranial recordings in epilepsy patients

26   while subjects resolved color-word conflict. We observed differential activity preceding

27   the behavioral responses to conflict trials throughout frontal cortex; this activity was

28   correlated with behavioral reaction times. These signals emerged first in anterior

29   cingulate cortex (ACC) before dorsolateral prefrontal cortex (dlPFC), followed by medial

30   frontal cortex (mFC) and then by orbitofrontal cortex (OFC). These results disassociate

31   the frontal subregions based on their dynamics, and suggest a temporal hierarchy for

32   cognitive control in human cortex.

Introduction

Flexible control of cognitive processes is fundamental to daily activities, including the execution of goal-directed tasks according to stimulus inputs and context dependencies. An important case of cognitive control arises when input stimuli elicit conflicting responses and subjects must select the task-relevant response despite competition from an often stronger but task-irrelevant response (Miller, 2000; Miller and Cohen, 2001). A canonical example of this type of conflict is the Stroop task: subjects are asked to name the font color of a word where the semantic meaning conflicts with the color signal (e.g. the word "red" shown in green versus red). Such incongruent inputs lead to longer reaction times, attributed to weaker signals (color processing) that must be emphasized over the automatic processing of word information (Stroop, 1935). The Stroop task is frequently used in cognitive neuroscience and clinical psychology and forms the foundation for theories of cognitive control.

Neurophysiological, neuroimaging, and lesion studies have ascribed a critical role in cognitive control to networks within frontal cortex (Miller, 2000; Miller and Cohen, 2001), yet the neural circuit dynamics and mechanisms responsible for orchestrating control processes remain poorly understood. Lesion studies (Cohen and Servan-Schreiber, 1992; Perrett, 1974), human neuroimaging measurements (Egner and Hirsch, 2005a; MacDonald, 2000), and macaque single unit recordings (Johnston et al., 2007) implicate the dorsolateral prefrontal cortex (dlPFC) in providing top-down signals to bias processing in favor of the task-relevant stimuli (Botvinick et al., 2001; Miller and Cohen, 2001). The medial frontal cortex (mFC) also participates in cognitive control, possibly in a conflict monitoring capacity (Botvinick et al., 2001; Ridderinkhof et al., 2004; Rushworth et al., 2004). Recordings and lesions studies in the macaque anterior cingulate cortex (ACC) (Ito et al., 2003; Nakamura et al., 2005) suggest that ACC neurons are principally involved in monitoring for errors and making between-trial adjustments (Brown and Braver, 2005; Ito et al., 2003; Johnston et al., 2007; Rothé et al., 2011)—an idea that has received support by a recent study in the human ACC (Sheth et al., 2012). Recent work has also demonstrated that the supplementary motor area and the medial frontal cortex play an important role in monitoring for errors (Bonini et al., 2014). An alternative and influential theoretical framework posits that the ACC monitors for

64   potential conflicts and subsequently directs the dlPFC to engage control processes

65   (Botvinick et al., 2001; Shenhav et al., 2013). Several human neuroimaging studies are

66   consistent with this notion (Botvinick et al., 1999; Kerns, 2006; Kerns et al., 2004;

67   MacDonald, 2000) but the relative contributions of dlPFC, mFC, and ACC to cognitive

68   control remain a matter of debate (Aarts et al., 2008; Cole et al., 2009; Fellows and

69   Farah, 2005; Mansouri et al., 2007; Milham and Banich, 2005; Milham et al., 2003;

70   Rushworth et al., 2004).

71   Previously, some neuroimaging studies have suggested that these frontal cortex

72   regions can be differentiated based on the presence or absence of conflict signals

73   (MacDonald, 2000). The challenge in dissociating the relative roles of these regions

74   during Stroop-like tasks is that increased task difficulty recruits a host of executive

75   functions (attention, decision-making, uncertainty, cognitive control). These functions are

76   associated with neural activity spanning tens to hundreds of milliseconds that and the

77   underlying dynamics are difficult to untangle with the low temporal resolution of existing

78   neuroimaging techniques (Shenhav et al., 2013). Human single neuron studies provide

79   millisecond resolution but have focused on individual regions (Sheth et al., 2012). We

80   took advantage of the high spatiotemporal resolution of intracranial recordings in human

81   epilepsy patients and the ability to record simultaneously from multiple regions to

82   directly investigate the dynamics of conflict responses during cognitive control. We

83   hypothesized that subregions of frontal cortex could be differentiated based on the

84   temporal profile of their conflict responses. We recorded intracranial field potentials from

85   1,397 electrodes in 15 subjects while they performed the Stroop task or a variation in

86   which they were asked to read the word instead of focusing on its color.

87   We observed conflict-selective activity throughout several regions in frontal

88   cortex: ACC, mFC, dlPFC, and also orbitofrontal cortex. Several analyses link these

89   signals to cognitive control. Neural responses were increased for incongruent compared

90   to congruent trials, and these signals correlated with behavioral reaction time, depended

91   on the task, and exhibited adaptation over trials. We compared pairs of simultaneously

92   recorded electrodes to disassociate these different regions based on the timing of these

93   conflict responses rather than their presence or absence. Conflict responses emerged first

94 in the ACC and subsequently emerged in dlPFC and mFC and finally in OFC. These

95 observations propose a plausible flow of signals underlying cognitive control.

96

Results

98   We recorded field potentials from 15 epilepsy patients implanted with intracranial

99 electrodes in frontal cortex as they performed the Stroop task (**Figure 1, Supplementary**

100 **File 1**). After 500 ms of a fixation cross, subjects were presented with one of three words

101 (Red, Blue, Green), which were colored either red, blue, or green. We refer to congruent

102 trials (C) where the font color matched the word (60% of the trials) compared to

103 incongruent trials (I) where the font color conflicted with the word (40% of the trials).

104 Within each trial type, the word-color combinations were counter-balanced and randomly

105 interleaved. The stimuli were presented for 2 seconds (in two subjects, for 3 seconds).

106 Subjects were asked to respond verbally and either name the color (Stroop task), or read

107 the word (Reading task) in separate blocks. Performance during congruent trials was

108 essentially at ceiling (**Figure 1-figure supplement 1**).

109   An ANOVA conducted on subjects' performance with stimulus type (congruent

110 or incongruent) and task (Stroop or Reading) as repeated measures revealed a significant

111 interaction between stimulus type and task (F = 22.9, $P < 0.001$). For the Stroop task,

112 subjects made more errors during incongruent trials (average error rate: 5±3%, $P < 0.001$

113 paired t-test), as demonstrated in previous studies (Bugg et al., 2008; Egner and Hirsch,

114 2005b; Kerns et al., 2004). There was no difference in the number of error trials during

115 the Reading task ($P = 0.76$, paired t-test). Subsequent analyses focused on correct trials

116 only unless otherwise stated. Subjects' reaction times also had a significant interaction

117 between stimulus type and task (F = 65.2, $P < 10^{-5}$, ANOVA). Consistent with previous

118 observations (Stroop, 1935), subjects' response times during the Stroop task were

119 delayed for incongruent trials compared to congruent trials (**Figure 1B,** average delay:

120 215±93 ms, $P < 0.001$, paired t-test, see also **Figure 1 – figure supplement 1C** for

121 individual subject data). The reaction time delays were shorter in the Reading task

122 (**Figure 1C**, average delay: 22±31 ms, $P = 0.02$, paired t-test). Trial history also has a

123 strong effect on reaction time (known as Gratton effect in the literature (Gratton et al.,

124 1992)). A repeated measures ANOVA revealed an interaction between previous and

125   current trial type (F = 19.5, P < 0.001). Incongruent trials that were preceded by a

126   congruent trial (cI trials) elicited slower reaction times compared to incongruent trials

127   that were preceded by an incongruent trial (iI trials) (**Figure 1D**, average reaction time

128   difference: 34±14 ms, P = 0.03, paired t-test). A similar Gratton effect was found for iC

129   versus cC trials (**Figure 1D**, average reaction time difference: 72±136ms, P < 0.001,

130   paired t-test).

131   We recorded intracranial field potentials from 1,397 electrodes (average

132   93±31 electrodes per subject) while subjects performed the Stroop and Reading tasks.

133   The number of electrodes per subject and the location of these electrodes were strictly

134   dictated by clinical needs. Therefore, there was a wide distribution of electrode

135   locations, as is typical in this type of recordings (Liu et al., 2009). We excluded

136   electrodes in epileptogenic regions. We focused on the neural signals in the gamma

137   band (70-120 Hz) given their prominence in sensory, motor and cognitive phenomena

138   (Crone et al., 1998a; Liu et al., 2009; Oehrn et al., 2014); results for other frequency

139   bands are shown in **Figure 2-figure supplement 2** and **Figure 4-figure supplement**

140   **1** and **2**. Presentation of the visual stimuli evoked rapid and color/word selective

141   neural responses in visual cortical areas within 200 ms of stimulus onset, as expected

142   from previous studies (e.g. (Liu et al., 2009)). Other electrodes were activated for

143   different motor (verbal) outputs (e.g. (Bouchard et al., 2013; Crone et al., 1998a)).

144

145   **Conflict responses in frontal cortex**

146   We focused on 469 electrodes located in areas within frontal lobe which have

147   been previously implicated in executive function: medial frontal cortex (mFC, n=111),

148   orbitofrontal cortex (OFC, n=156), dorsolateral prefrontal cortex (dlPFC, n=168) and the

149   anterior cingulate cortex (ACC, n=34). We applied a non-parametric analysis of variance

150   (ANOVA) to measure whether and when the physiological responses differed between

151   congruent and incongruent trials. An electrode was considered conflict-selective if the F-

152   statistic was greater than a significance threshold computed by a permutation test with P

153   = 0.001 for 50 consecutive milliseconds (**Methods**). The latency was defined as the first

154   time of this threshold-crossing.

155   **Figure 2** shows an example electrode from the left Anterior Cingulate Cortex that

156   responded differentially between congruent and incongruent trials during the Stroop task.

157   These signals were better aligned to the speech onset than to the stimulus onset, as shown

158   in the response-aligned view (compare **Figure 2A-C** with **Figure 2D-F**). During the

159   Stroop task, the response-aligned signals were significantly stronger for the incongruent

160   (brown) trials compared to the congruent (black) trials (**Figure 2D**, $P < 10^{-5}$, ANOVA),

161   and were invariant to the particular word/color combinations (**Figure 2G**). Incongruent

162   trials could be discriminated from congruent trials at a latency of 669±31 ms

163   (mean±s.e.m.) before the onset of the response (**Figure 2D**). This conflict response was

164   also specific to the Stroop task; there was a significant interaction between congruency

165   and task (F = 13.5, $P = 0.007$, ANOVA**).** The same stimuli did not elicit differential

166   activity during the Reading task (**Figure 2F**). We assessed the correlation between the

167   neural signal strength and behavioral reaction times in single trials. The maximal gamma

168   power during each incongruent trial (using the average gamma power yielded similar

169   results) was positively correlated with the behavioral reaction times (**Figure 2H**, □ =

170   0.25, $P = 0.02$).

171   Any differences between congruent and incongruent trials in the stimulus-aligned

172   analyses can be confounded by the reaction time differences; therefore, we focus

173   subsequent analyses on the response-aligned signals. More example electrodes are shown

174   in **Figure 2-figure supplement 1** (dlPFC) and **Figure 2-figure supplement 2** (OFC).

175   Using the aforementioned criteria, we identified n=51 conflict selective frontal

176   cortex electrodes during the Stroop task, with contributions from 13 subjects

177   (**Supplementary Files 2** and **3**). These electrodes were distributed throughout different

178   subregions within frontal cortex (**Figure 3A**). To evaluate whether random variation in

179   the signals could give rise to apparent conflict-selective electrodes, we randomly shuffled

180   the congruent/incongruent trial labels 10,000 times and applied the same statistical

181   criteria (**Methods**). Across our population, we found n=4.4±0.03 false positive electrodes

182   (mean±s.e.m., out of 469 electrodes), which corresponds to a false discovery rate (FDR)

183   of q = 0.01, which is significantly less than our observation of n=51 electrodes. The

184   number of conflict-selective electrodes within each subregion was significantly greater

185   than expected by chance (**Figure 3B**, $P < 0.01$, all regions). We repeated the analyses

186   during the Reading task. In contrast with the Stroop task, we only observed n=3 conflict-

187   selective frontal cortex electrodes during the Reading task (out of 469 electrodes), a

188   number that is within the false positive rate.

189   To account for within-subject and across-subject variation, we used a multilevel

190   model (Aarts et al., 2014) to conduct a group analysis of the physiological responses,

191   with electrodes nested within subjects (**Methods**). Across the population, we observed a

192   significant interaction between the factors congruency and task on the gamma power

193   (χ^2=9.2, P = 0.002). Consistent with the single electrode examples, gamma power was

194   greater for incongruent compared to congruent trials, but only during the Stroop task

195   (**Figure 4A**, Stroop: $P < 10^{-3}$, Reading: P = 0.56). We computed the average response in

196   each region (**Figure 4B**). Each electrode's response was normalized by dividing the

197   power during incongruent trials by the power in congruent trials (dividing the brown

198   curve by the black curve in **Figure 2**), computing the logarithm and finally pooling

199   within each region. The pooled responses in the OFC are visually less compelling

200   (**Figure 4B**, bottom right subplot) due to the heterogeneity in the latency of the individual

201   electrodes but the responses in the OFC were as vigorous as the ones in other areas (e.g.

202   **Figure 2-figure supplement 2**). Similar conclusions were reached when plotting the

203   pooled responses aligned to stimulus onset (**Figure 4-figure supplement 3**).

204

205   **Behavioral relevance of physiological responses**

206   Several lines of evidence demonstrate a link between the neural signals described

207   in the previous section and cognitive control: the neural signals correlated with reaction

208   times, showed behavioral adaptation, and demonstrated error monitoring.

209   As shown in previous studies, there was a wide distribution of behavioral reaction

210   times (**Figure 1B**). Consistent with the example electrode in **Figure 2**, behavioral

211   reaction times across the population correlated with the strength of the physiological

212   signals, even after controlling for trial history (**Figure 4C**, $P < 10^{-5}$, sign-rank test).

213   The strength of these neural signals also revealed a neural correlate of the

214   behavioral Gratton effect documented in **Figure 1D**: gamma power was greater in cI

215   compared to iI trials (**Figure 4D**). Using the aforementioned multilevel model, we found

216   a significant interaction between trial history (cI or iI) and task (χ^2=4.4, P = 0.03). This

217   Gratton effect was stronger in the Stroop task ($P < 0.001$) than in the Reading task ($P =$

218   0.72). These differences were not observed for cC versus iC trials, where the interaction

219   was not significant ($\chi^2=1.9$, $P = 0.17$) (**Figure 4E**). This analysis was performed after

220   removing stimulus repetition trials. The Gratton effect was present in all four frontal

221   regions and there were no statistically significant differences in the strength of the effect

222   across regions (F = 0.25, $P=0.86$, ANOVA). To control for reaction time effects on these

223   comparisons, we ran an analysis of covariance (ANCOVA) to test for a main effect of

224   trial history on the gamma power with the behavioral reaction time as a covariate

225   (**Methods**). The neural Gratton effect during the Stroop task persisted under these

226   controlled conditions ($P = 0.0002$, multilevel model). We also explicitly ruled out

227   reaction time differences by subsampling to match the reaction time distribution between

228   conditions, with similar results ($P = 0.01$, multilevel model). Together, these results

229   suggest that the neural signals described here code for an internally perceived level of

230   conflict that exhibits conflict adaptation and correlates with the across-trial variability in

231   reaction times.

232

233   **Conflict responses in other frequency bands**

234   The results presented above focus on the neural signals filtered within the gamma

235   frequency band (70-120 Hz). We also examined the responses elicited in the broadband

236   signals (1 to 100 Hz) as well as in the theta, (4 to 8 Hz), beta, (9 to 30 Hz), and low

237   gamma (30-70 Hz) bands. No conflict selective responses were observed in the

238   broadband signals or low gamma band. We found conflict-selective responses both in the

239   theta and beta bands (see example in **Figure 2-figure supplement 2a-f**). Across theta and

240   beta frequency bands, we also observed a significant interaction between Congruency and

241   Task (theta: $P < 10^{-5}$, beta: $P < 10^{-4}$, multilevel model). Consistent with the results

242   reported in the gamma frequency band, conflict responses in the theta and beta bands

243   were more prominent during the Stroop task compared to the Reading task (**Figure 4-**

244   **figure supplement 1**). In contrast to the results in the gamma band, power in the theta

245   and beta bands *decreased* during incongruent trials. Furthermore, power in the theta and

246   beta frequency bands was not correlated with reaction times (theta: $P = 0.43$, beta: $P =$

247   0.09, sign-rank test).

248   In addition to separately examining the responses in different frequency bands, an

249   important aspect of encoding of cognitive information is the relationship between signals

250   across frequencies. In particular, several studies have demonstrated that the amplitude of

251   the gamma band is coupled to the phase of slower oscillations in the theta band (Canolty

252   et al., 2006; Oehrn et al., 2014; Tort et al., 2008). We therefore examined the degree of

253   cross-frequency coupling between the signals in the gamma and theta bands (**Figure 4-**

254   **figure supplement 2**). Consistent with previous studies, we found that 50% of the

255   electrodes demonstrated significant theta-gamma coupling. However, the strength of this

256   coupling was not different between congruent and incongruent trials across the

257   population of conflict-selective electrodes ($P = 0.52$, sign-rank test).

258

259   **Error monitoring signals**

260   The conflict responses reported above are based on correct trials only. Yet, error

261   monitoring has also been ascribed to frontal cortical circuits (Bonini et al., 2014; Shenhav

262   et al., 2013; Yeung et al., 2004). To investigate whether the same electrodes responding

263   to conflict are also involved in successful error monitoring, we analyzed the neural

264   signals during self-corrected trials. In these trials, subjects initially made an erroneous

265   response and rapidly corrected themselves with the right answer. Given the high

266   performance level of all subjects, the number of such trials is low. However, these trials

267   are particularly interesting because we can be certain that there was successful error

268   detection (as opposed to error trials without any self-correction). An example self-

269   corrected trial from the ACC electrode shown previously is illustrated in **Figure 5A**. The

270   subject initially made an incorrect response (green), which was rapidly followed with the

271   correct response (red). Increased gamma power was observed after onset of the erroneous

272   response. In contrast, the following corrected behavioral response exhibited no such post-

273   response signal. Additionally, these error-monitoring signals were not observed in correct

274   incongruent trials (**Figure 2D**), and were consistent across the n =11 self-corrected trials

275   for this subject (**Figure 5B**, $P = 0.001$, signed rank test). Another example electrode is

276   shown in **Figure 5C-D**. There were only two subjects contributing n = 7 conflict-

277   signaling electrodes that had a sufficient number of self-correction trials (greater than five

278   trials) for this analysis. For each electrode, we compared the difference in neural signals

279   during the one-second post-response window between the initial error and the following

280   self-correction. Of those n=7 electrodes, n = 5 electrodes showed evidence of error

281   monitoring (**Figure 5E**, $P < 0.05$, sign-rank test). Although the number of electrodes and

282   trials in this analysis is small, these results provide a direct correlate of error monitoring

283   signals. Furthermore, these results highlight that the same electrodes that respond to

284   conflict leading up to the behavioral response can also show post-response error

285   monitoring.

286

287   **Regional differences in conflict response latencies**

288   We observed conflict-selective responses in the anterior cingulate cortex, medial

289   frontal cortex, dorsolateral prefrontal cortex and orbitofrontal cortex. To examine the

290   dynamics of cognitive control orchestrating the transformation of conflicting visual

291   signals to motor outputs, we compared, across those four regional groups, the latencies

292   relative to behavioral response onset at which the congruent and incongruent trials could

293   be discriminated. Comparing latencies across regions is difficult, especially across

294   subjects with varying reaction times. For a controlled and direct comparison, we

295   restricted the analysis to compute the latency differences between pairs of simultaneously

296   recorded electrodes. This within-subject pairwise analysis had increased power to

297   examine the relative dynamics between frontal lobe areas (**Figure 6**). The relative

298   latencies were significantly different across the regions ($P = 0.01$, permutation test, post-

299   hoc testing was controlled for multiple comparisons using the Benjamin-Hochberg

300   procedure, **Methods**). Conflict responses in the ACC preceded those in all the other

301   frontal lobe regions, followed 207±40 ms later by dorsolateral prefrontal cortex and

302   388±83 ms later by medial frontal cortex. Signals in orbitofrontal cortex emerged 319±78

303   ms after dlPFC. This entire processing cascade took approximately 500 ms. For

304   comparison, subjects' behavioral reaction times to incongruent trials were 1,105±49 ms.

305   The latency difference between ACC and dlPFC is based on 6 electrode pairs: one ACC

306   electrode and six simultaneously recorded dlPFC electrodes. There was only one pair of

307   simultaneous recordings between ACC and OFC and we do not report this value in

308   **Figure 6**. The other region comparisons have contributions from multiple electrodes in

309   multiple subjects (Supplementary File 3). These results suggest a temporal hierarchy of

310   cognitive control mechanisms culminating in speech onset.

311

312

313   **Discussion**

314   We used intracranial field potentials to measure the dynamics of conflict

315   responses across frontal cortex leading up to the behavioral response in the Stroop task.

316   Previous physiological and functional neuroimaging studies have documented the

317   involvement of multiple of these frontal cortex areas in the Stroop or similar tasks

318   (Botvinick et al., 1999; Kolling et al., 2012; MacDonald, 2000; Niendam et al., 2012;

319   Oehrn et al., 2014; Sheth et al., 2012). The intracranial field potential recordings reported

320   here show conflict-selective signals in ACC (e.g. **Figure 2**), dlPFC (e.g. **Figure 2-figure**

321   **supplement 1**), mFC (e.g. **Figure 4B**) and OFC (e.g. **Figure 2-figure supplement 2**).

322   The mFC and dlPFC have been previously implicated in cognitive control, and these

323   structures are extensively connected to the rest of frontal cortex areas (Ridderinkhof et

324   al., 2004). The role of the OFC in cognitive control during Stroop-like tasks has not been

325   reported previously, possibly because of technical challenges in neuroimaging near this

326   area (Weiskopf et al., 2006).

327   We presented several lines of evidence that demonstrate that these conflict-

328   selective physiological signals are relevant for behavior during the Stroop task. Longer

329   behavioral reaction times were correlated with greater gamma power on a trial-by-trial

330   basis during the Stroop task but not during the Reading task, even after accounting for

331   trial history and for differences between congruent and incongruent stimuli (**Figure 2H,**

332   **4C**). The same identical stimuli can elicit a range of behavioral reaction times and this

333   internal degree of conflict can be captured, at least partly, by the strength of gamma

334   power in frontal cortex in each trial.

335   The neural correlates of behavioral adaptation (Gratton effect (Gratton et al.,

336   1992)) were observed in the ACC, consistent with prior studies based on human single

337   neuron recordings (Sheth et al., 2012), neuroimaging (Botvinick et al., 1999; Kerns,

338   2006) and also in accordance with the behavioral effects of ACC resection (Sheth et al.,

339   2012). Conflict responses throughout the other frontal cortex regions also demonstrated

340 the neural Gratton effect, suggesting a more distributed network involved in across-trial

341 adaptation than previously hypothesized. The physiological responses in these areas were

342 stronger in cI trials (incongruent trials that were preceded by congruent trials) than iI

343 trials (**Figure 4D**). While the increased activity in cI trials compared to iI trials is

344 consistent with neuroimaging studies (Botvinick et al., 1999), single neuron recordings in

345 a different Stroop-like task report the opposite relationship (iI > cI) (Sheth et al., 2012).

346 These differences point to potentially interesting distinctions between the activity of

347 individual neurons and coarser population measures that warrant further investigation.

348   Another discrepancy between neuroimaging studies and single unit recordings is

349 the presence of conflict responses and error signals. Single unit recording in macaque

350 ACC typically find error monitoring signals but not conflict-selective responses (Cole et

351 al., 2009; Emeric et al., 2010; Ito et al., 2003; Taylor et al., 2006), see however (Ebitz and

352 Platt, 2015), whereas human neuroimaging studies observe both types of signals in ACC.

353 There has been significant debate concerning whether action monitoring and conflict

354 detection represent distinct processes (Carter et al., 1998; Carter et al., 2000; Nee et al.,

355 2011; Swick and Turken, 2002). Because both processes may co-occur on the same trials,

356 high temporal resolution is required to disassociate the two computations. A recent

357 human intracranial study has found error signals in supplementary motor area and medial

358 frontal cortex (Bonini et al., 2014), and a human single unit study reported conflict

359 signals in ACC (Sheth et al., 2012). The current work demonstrates the coexistence of

360 both error signals and conflict signals. The analysis of the few self-correction trials in our

361 data suggests that the same areas responsible for pre-behavioral conflict signals can also

362 produce post-behavioral response error-monitoring signals (**Figure 5**). In addition, the

363 relative timing of the conflict and error signals surrounding the neural responses confirms

364 computational predictions based on a connectionist architecture to explain the

365 mechanisms of conflict (Yeung et al., 2004) and scalp EEG studies (Hughes and Yeung,

366 2011). These results are consistent with computational models suggesting that these

367 signals may represent a general error-likelihood prediction, of which conflict and error

368 detection are special cases (Brown and Braver, 2005).

369   It has been suggested that ACC and supplementary eye field neurons in macaque

370 monkeys respond to specific stimulus and/or behavioral combinations but are not directly

371 modulated by conflict (Cole et al., 2009; Nakamura et al., 2005). At the level of the

372 intracranial field potentials reported here, the modulation of conflict trials observed in the

373 four frontal cortex regions could not be ascribed to specific stimulus or behavioral

374 responses (e.g. **Figure 2G**) and were also task dependent (compare **Figure 2A** versus

375 **2C**). In these patients, we did not have access to single neuron responses and we

376 therefore cannot rule out the possibility that individual neurons show distinct patterns of

377 responses that are averaged out at the field potential level.

378   Besides the high gamma band, we also observed conflict responses in the beta and

379 theta bands, but not the low gamma band (e.g. **Figure 4-figure supplement 1**). Previous

380 work has suggested differential roles for distinct oscillatory components of the local field

381 potential (Cavanagh and Frank, 2014; Kahana et al., 2001; Ullsperger et al., 2014; von

382 Stein and Sarnthein, 2000). There were clear differences in the type of information

383 conveyed by distinct frequencies components. Lack of significant correlations with

384 reaction time in the theta and beta bands suggests that the gamma band better captures the

385 behavior. Additionally, conflict responses were characterized by increased power in the

386 gamma band, but decreased power in the theta and beta bands (**Figure 4-figure**

387 **supplement 1**). Previous scalp EEG recordings (Cavanagh and Frank, 2014; Ullsperger

388 et al., 2014; van Driel et al., 2015) have demonstrated that conflict and/or error trials

389 elicit increased theta power, suggesting potentially interesting differences in how theta is

390 captured across spatial scales. We also observed a decrease in beta power, which is

391 consistent with previous studies that correlate frontal cortex activation with

392 desynchronization in the beta band and increased synchronization in the gamma bands

393 (Crone et al., 1998a; Crone et al., 1998b). Differences across tasks, recording methods,

394 and targeted regions should be interpreted with caution. The roles of different oscillatory

395 components in neocortex are not clearly understood. One possibility is that lower

396 frequency bands reflect the summed dendritic input of the nearby neural population

397 (Logothetis et al., 2001; Mitzdorf, 1987) and can act as channels for communication

398 (Cavanagh and Frank, 2014), whereas higher frequency bands represent the population

399 spiking rate (Buzsaki et al., 2012; Ray and Maunsell, 2011). Along these lines, we

400 speculate that the theta desynchronization we observe could reflect a reduction of inputs,

401 leading to inhibition of the prepotent but erroneous response.

402   While we observed conflict responses throughout frontal cortex, the

403   spatiotemporal resolution of our intracranial recordings allowed us to separate regions by

404   the latency at which conflict-selective responses emerge with respect to speech onset. By

405   comparing pairs of simultaneously recorded electrodes, we found that conflict responses

406   in the ACC lead the dlPFC by ~200 ms. Medial frontal cortex is anatomically close and

407   extensively connected to the ACC, and the two regions are often grouped together

408   (Cavanagh et al., 2009; Ridderinkhof et al., 2004). Yet, conflict responses in the mFC

409   trail the ACC by hundreds of milliseconds, suggesting an important distinction between

410   the two regions (Rushworth et al., 2004). The relative latency measurements place the

411   OFC at the bottom of this cascade. The hierarchical cascade of processes described here

412   is consistent with predictions from mechanistic models of cognitive control (e.g. see

413   Figure 2 in Shenhav et al, Neuron 2013). In particular, stimulus related signals are

414   evident along the ventral visual stream early on and feed onto frontal cortex, where we

415   find that ACC activity precedes activity in other frontal regions, followed by dlPFC, and

416   finally mFC, and OFC.

417   Since the local field potential pools over many neurons, latency measures can be

418   influenced by a variety of factors, such as the proportion of neurons selective for conflict

419   and their laminar organization. Yet, at least in the ACC, the temporal profile of conflict

420   responses we observed is similar to responses from human single unit recordings (Sheth

421   et al., 2012). The relatively long delays between regions are also particularly intriguing.

422   There are monosynaptic connections that link these four regions within frontal cortex and

423   yet, it takes 100-200 ms to detect the relative activation between these areas (**Figure 6**).

424   Daily decisions require integration of different goals, contexts, input signals, and

425   the consequences of the resulting actions. The current study provides initial steps to

426   elucidate not only which brain areas participate in cognitive control on a trial-by-trial

427   basis but also their relative interactions and differential roles. The relative latency

428   measurements and correlations between neural activity and reaction time provide a

429   framework to constrain theories of cognitive control, and propose a plausible flow of

430   conflict responses through frontal cortex.

431

Materials and methods

Subjects. Subjects were 15 patients (10 male, Ages 10-50, **Supplementary File 1**) with pharmacologically intractable epilepsy treated at Children's Hospital Boston (CHB), Johns Hopkins Medical Institution (JHMI), Brigham and Women's Hospital (BWH), or Taipei Veterans General Hospital (TVGH). These subjects were implanted with intracranial electrodes in frontal cortex for clinical purposes. Five other subjects participated in this task but they were excluded from the analyses because they did not have any electrodes in frontal cortex. All studies were approved by each hospital's institutional review boards and were carried out with the subjects' informed consent.

Recordings. Subjects were implanted with 2mm diameter intracranial subdural electrodes (Ad-Tech, Racine, WI, USA) that were arranged into grids or strips with 1 cm separation. Electrode locations were determined by clinical considerations. There were 1,397 electrodes (15 subjects). Sampling rates ranged from 256 Hz to 1000 Hz depending on the equipment at each institution: CHB (XLTEK, Oakville, ON, Canada), BWH (Bio-Logic, Knoxville, TN, USA), JHMI (Nihon Kohden, Tokyo, Japan), and TVGH (Natus, San Carlos, CA). All the data were collected during periods without any seizure events or following any seizures.

Task procedures. A schematic of the task is shown in **Figure 1**. After 500 ms of fixation, subjects were presented with a word stimulus for 2 seconds. The stimulus presentation was 3 seconds in two subjects. Stimuli were one of three words (Red, Blue, Green) presented in the subjects' primary language (CHB, BWH, JHMI: English; TVGH: Mandarin) either in red, blue, or green font color. Stimuli subtended approximately 5 degrees of visual angle and were centered on the screen. Trials were either congruent (C), where the font color matched the word, or incongruent (I), where the font color conflicted with the word. The order of congruent and incongruent trials was randomized. Approximately 40% of the trials were incongruent trials. Within congruent trials and within incongruent trials all color-word combinations were counter balanced and

462 randomly interleaved. Subjects were asked to either name the color (Stroop task) or read

463 the word (Reading task) within the time limit imposed by the stimulus presentation time.

464   Each block contained 18 trials, and the two tasks were completed in separate

465 blocks. Most subjects completed 18 blocks of the Stroop task and 9 blocks of the Reading

466 task (**Supplementary File 1**). Audio was recorded using a microphone at 8192 Hz

467 sampling rate. No correct/incorrect feedback was provided.

468

469 **Electrode Localization.** Electrodes were localized by co-registering the preoperative

470 magnetic resonance imaging (MRI) with the postoperative computer tomography (CT)

471 (Destrieux et al., 2010; Liu et al., 2009). In 4 subjects without a postoperative CT,

472 electrodes were localized using intraoperative photographs and preoperative MRI. For

473 each subject, the brain surface was reconstructed from the MRI and then assigned to one

474 of 75 regions by Freesurfer. Depth electrodes were assigned to either a subcortical

475 structure or to gyri/sulci.

476   We focused on those electrodes in visual cortex and in four frontal cortex regions

477 (ACC: anterior and middle-anterior cingulate gyrus, mFC: superior frontal gyrus, dlPFC:

478 middle frontal gyrus, and OFC: orbitofrontal gyrus).

479

480 **Behavioral analyses.** To determine the behavioral reaction time for each trial, the short-

481 time energy was computed from the audio recordings. For an audio signal *x(t)*, the short-

482 time energy *E(t)* is defined as:

$$E(t) = \sum_{m=0}^{m=T} [x(m)w(t-m)]^2,$$

483 where *T* is the length of the recording and *w(t)* is a 300-point Hamming window (~40

484 ms). Speech onset was defined as the first time when the energy crossed a threshold set as

485 1 standard deviation above the baseline. Only trials where the subject gave a single verbal

486 response and the speech onset could be identified were considered correct trials.

487

488 **Preprocessing.** Unless otherwise noted, analyses in this manuscript used correct trials

489 only. Electrodes with significant spectral noise were excluded from analysis (n=25 out of

490 1,397 total electrodes). For each electrode, a notch filter was applied at 60 Hz, and the

491   common average reference computed from all channels was subtracted. Power in the

492   theta (4-8 Hz), beta (9-30 Hz), and high-gamma band (70-120 Hz) was extracted using a

493   moving window multi-taper Fourier transform (Chronux toolbox (Mitra and Bokil,

494   2008)) with a time-bandwidth product of five and seven tapers. The window size was 200

495   ms with 10 ms increments. In several figures, the gamma power was z-scored for display

496   purposes (see figure legends).

497

498   **Analyses of neural response selectivity**. To determine whether and when an electrode

499   responded selectively to conflict, we used a sliding F-statistic procedure (Liu et al.,

500   2009). Electrodes with differential responses between congruent and incongruent trials

501   were selected by computing the F-statistic, for each time bin, comparing the neural

502   responses between congruent and incongruent trials. Electrodes were denoted as 'conflict

503   selective' if (1) the F-statistic exceeded a significance threshold for 50 consecutive

504   milliseconds, and (2) the average neural response exceeded one standard deviation above

505   the baseline period at least once during the trial. A null distribution generated by

506   randomly permuting the labels was used to set the significance threshold with $P=0.001$.

507   The latency at which congruent and incongruent stimuli could be discriminated was

508   defined as the first time of this threshold crossing. For the response-aligned view, only

509   electrodes where the latency preceded the response were included in subsequent analysis.

510   This selection process was independently performed for each electrode in both stimulus-

511   aligned and response-aligned analyses, and separately for the Stroop and Reading task.

512   We used a permutation test with 10,000 shuffles to obtain a false discovery rate

513   for our selection process. The congruent/incongruent trial labels were randomized 10,000

514   times and we measured the average number of electrodes across our population that

515   passed the selection procedure.

516

517   **Single electrode analyses**. For the selected electrodes obtained with the procedure

518   described in the previous section, we performed a number of within-electrode analyses.

519   We measured single-trial correlations with behavioral reaction times, assessed the

520   significance of interactions and simple/main effects, and controlled for confounds in

521   measuring the neural Gratton effect.

522

523   *Single-trial analysis*. For single trial comparisons across conditions, signal power for
524   each trial was computed for both response-aligned and stimulus-aligned analyses. For
525   stimulus-aligned data, the signal power was defined as the maximal power from stimulus
526   onset to 1 second after stimulus onset. For response-aligned analyses, the signal power
527   was defined as the maximal power from one second before the response to the response
528   onset. Analyses using the average power within the same window yielded similar results.
529   Single-trial response latency was defined as the time of maximal activation relative to
530   stimulus onset.

531

532   *Interaction Effects*. For conflict-selective electrodes, we measured the significance of task
533   dependence by performing, at each time bin, an ANOVA on the gamma power with the
534   factors Congruency and Task (Nieuwenhuis et al., 2011). The peak F-statistic of the
535   interaction term over the pre-response window was compared against a null distribution
536   generated by randomly shuffling the trial labels. Simple effects were tested using this
537   same approach.

538

539   *Neural Gratton Effect*. We evaluated the neural signal difference between trials with
540   different histories (e.g. cI versus iI), while removing trials with stimulus repetitions.
541   Given that (1) reaction times are different for the cI versus iI trials (**Figure 1**) and (2)
542   gamma power is significantly correlated with reaction time in incongruent trials (**Figure
543   4**), we would expect differences in gamma power in cI versus iI trials. To control for this
544   potential confounding effect in our measurements of trial history dependence, we applied
545   two methods. First, for each electrode, we performed an ANCOVA on the gamma power
546   with trial history (cI or iI, for example) as the group and reaction time as a covariate. We
547   computed the regression line, extracted the RT-adjusted gamma power from the y-
548   intercept and used this value in the group analysis. Second, we performed a matched
549   reaction time analysis, where the distribution of reaction times was equalized by
550   subsampling the trials in a histogram-matching procedure with 200ms bins. This resulted
551   in using only ~50% of the trials. The same analysis was then applied to this reaction time
552   matched dataset.

553

554　**Group Analysis**. To account for both within-subject and across-subject variance,

555　statistical testing of the electrophysiological data was conducted with multilevel models

556　(Aarts et al., 2014; Goldstein, 2011) (also known as random effect models). Random

557　factors included electrodes nested within subjects. Significance of interactions and/or

558　main effects was assessed with a likelihood ratio test against a null model excluding that

559　particular term.

560　　　　For comparison of latency across regions, we restricted our analyses to

561　simultaneous measurements made within each subject. We computed the latency

562　difference for each pair of simultaneously recorded electrodes from different regions. The

563　F-statistic of this latency difference across the groups was compared against a null

564　distribution generated by shuffling, within each subject, the region labels (n = 10,000

565　shuffles). Post hoc testing used the Benjamin-Hochberg procedure to control for multiple

566　comparisons.

567

568　**Cross-Frequency Coupling**. To measure cross-frequency coupling between the theta

569　and gamma frequency bands, we used the Modulation Index (MI) defined previously

570　(Tort et al., 2008). Activity in the theta (4-8 Hz) and high gamma (70-120 Hz) bands was

571　obtained with a zero-phase least-squares finite impulse response (FIR) filter.

572　Instantaneous phase and amplitude was extracted with the Hilbert Transform. For the

573　Stroop and Reading Task separately, the MI was computed as the Kullback-Leiber

574　distance between the phase-amplitude histogram and a uniform distribution. For

575　comparison between tasks, the number of trials was equalized. This MI was compared

576　against a surrogate distribution generated by randomly lagging the time series across

577　1,000 repetitions. Similar results were obtained with the measure defined in Canolty et al.

578　(Canolty et al., 2006). Results were also similar when a surrogate distribution was created

579　by randomly pairing low-frequency phase with high-frequency power from different

580　trials.

581　　　　To compare the strength of cross-frequency coupling between congruent and

582　incongruent conditions, we computed the difference in MI between the two conditions

583 while equalizing the trial count. This difference was compared against a null distribution

584 generated by randomly shuffling the congruent and incongruent labels.

585

586 **Acknowledgments**

592

593 **Author Contributions**

594 H.T. and G.K. designed the experiment; H.T., H.Y., C.C., and N.E.C. performed the

595 experiments; J.R.M. and W.S.A. performed the neurosurgical procedures; H.T. and G.K.

596 analyzed the data; H.T. and G.K. wrote the manuscript incorporating feedback from all

597 other authors. All authors commented and approved the manuscript.

598

599   **Figure Legends**

600

601   **Figure 1. Experimental task and behavioral performance**

602   (**A**) Subjects were presented with one of three words (Red, Blue or Green); each word

603   was randomly colored red, blue, or green. Trials were incongruent (I) when the word

604   and color did not match, and were congruent (C) otherwise. The word-color

605   combinations were counter-balanced and randomly interleaved. Subjects performed

606   the Stroop task (name the color), and the Reading task (read the word) in separate

607   blocks.

608   (**B**) Distribution of z-scored behavioral reaction times (speech onset) across all

609   subjects (n=15) for congruent (black) or incongruent (brown) trials during the Stroop

610   task. Bin size = 0.2. Dashed lines indicate average reaction times.

611   (**C**) Distribution of z-scored reaction times during the Reading task.

612   (**D**) Z-scored reaction time across subjects for different trial histories during the

613   Stroop Task (cI: incongruent trial preceded by congruent trial; iI: incongruent trial

614   preceded by incongruent trial; iC: congruent trial preceded by incongruent trial; cC:

615   congruent trial preceded by congruent trial). Error bars indicate s.e.m.

616

617   **Figure 1-figure supplement 1. Behavioral data for each subject**

618   (**A-B**) Percent correct for each subject for the Stroop task (**A**) or Reading task (**B**)

619   during congruent (black) or incongruent (brown) trials. Subjects made more errors for

620   incongruent trials compared to congruent trials during the Stroop task ($P < 0.001$,

621   signed-rank test). One subject (Subject 6) did not participate in the Reading task.

622   (**C-D**) Average behavioral reaction time (speech onset) for each subject for the Stroop

623   task (**C**) or Reading task (**D**). Error bars indicate s.e.m. Subjects had delayed

624   responses for incongruent trials compared to congruent trials during the Stroop task

625   ($P < 0.001$, signed-rank test).

626

627   **Figure 2. Example electrode in left Anterior Cingulate Cortex**

628   (**A**) Average gamma power signals aligned to the stimulus onset from an electrode

629   during the Stroop task, for congruent (black) or incongruent (brown) stimuli. For

630　display purposes only, we z-scored the gamma power by subtracting the average and

631　dividing by the standard deviation of power during the baseline period (500 ms prior

632　to stimulus onset). Shaded areas indicate s.e.m. The total number of trials for each

633　condition is indicated in the upper right.

634　(**B**) Single-trial data for congruent (left) and incongruent (right) trials.

635　Each row is a trial, and the color indicates the z-scored gamma power (color scale on

636　upper right). Trials are sorted by behavioral response time (black line).

637　(**C**) Same as (**A**), but showing data from the Reading task.

638　(**D-F**) Same as in **A-C**, but aligning the data to behavioral response time. Gamma

639　power was better aligned to the behavioral response, and was stronger for incongruent

640　compared to congruent trials. The dashed line indicates the response-aligned latency,

641　defined as the first time point at which incongruent and congruent trials can be

642　discriminated.

643　(**G**) Signals elicited by each of the 9 possible stimulus combinations.

644　(**H**) There was a correlation between the maximal z-scored gamma power and behavioral

645　reaction times during incongruent trials (Pearson correlation coefficient $= 0.25$, $P = 0.02$,

646　permutation test). Each point in this plot represents a single trial.

647

648　**Figure 2-figure supplement 1. Example conflict-selective electrode in the right**

649　**dorsolateral Prefrontal Cortex**

650　Here we show a different conflict selective electrode, located in the dlPFC (format as

651　in **Figure 2**).

652

653　**Figure 2-figure supplement 2. Example conflict-selective electrode in the**

654　**Orbitofrontal Cortex comparing responses in the Theta and Gamma Bands**

655　(**A-F**) Responses in the theta power frequency band, z-scored. Same format as **Figure**

656　**2-figure supplement 1**.

657　(**G-L**) Responses in the gamma power frequency band, z-scored. Same format as

658　**Figure 2- figure supplement 1.**

659

660　**Figure 3. Electrode locations**

661   (**A**) Location of conflict-selective electrodes (black/gray) shown on a reference brain,

662   with each region colored (**Methods**). Electrodes from the right hemisphere were

663   mapped to the left hemisphere for display purposes. For more detail, see

664   **Supplementary File 2**.

665   (**B**) Percent of total electrodes in each region that were selective for conflict. Chance

666   levels were computed using a permutation test (black line). The number of observed

667   electrodes was significantly above chance for all regions ($P < 0.01$, permutation test,

668   **Methods**).

669

670   **Figure 4. Gamma power in frontal cortex correlates with behavior**

671   (**A**) Distribution of gamma power log-ratio (Incongruent/Congruent) for the Stroop

672   task (blue) and Reading task (green). Bin size = 0.05. Gamma power showed a

673   significant interaction between Congruency and Task ($P = 0.002$, multilevel model,

674   **Methods**). Power was larger for incongruent versus congruent trials during the Stroop

675   task ($P < 0.001$, n=51 frontal cortex electrodes) but not during the Reading task

676   (green, $P = 0.56$). The statistical analyses directly compare the gamma power, we

677   show the log-ratios here for display purposes only.

678   (**B**) Normalized gamma power log-ratio averaged across electrodes from each of the

679   four different frontal cortex regions during the Stroop task. We divided the power

680   during incongruent trials by the power during congruent trials, then computed the log

681   and finally averaged across electrodes. Data are aligned to the behavioral response

682   onset (t=0).

683   (**C**) Distribution of Pearson correlation coefficients between the maximal gamma

684   power and behavioral reaction time during incongruent trials for n=51 frontal cortex

685   electrodes. These correlations were significantly positive ($P < 10^{-5}$, sign-rank test).

686   Bin size = 0.1.

687   (**D**) For incongruent trials, there was a significant interaction between trial history and

688   task ($P=0.03$, multilevel model). Gamma power was larger for incongruent trials

689   preceded by congruent trials (cI) compared to incongruent trials preceded by

690   incongruent trials (iI), particularly during the Stroop task (blue, $P = 0.001$), compared

691   to the Reading task (green, $P = 0.72$). Data beyond the range of the x-axis are shown

692   in the first or last bins.

693   (**E**) For congruent trials, there was no interaction between trial history and task ($P =$

694   0.17, multilevel model). Gamma power was similar in congruent trials preceded by

695   incongruent trials (iC) compared to congruent trials preceded by congruent trials (cC)

696   during the Stroop task (blue, $P = 0.16$) and during the Reading task (green, $P = 0.19$).

697

698   **Figure 4-figure supplement 1. Theta and Beta band population results**

699   (**A**) Distribution of theta power log-ratio (Incongruent/Congruent) for the Stroop task

700   (blue) and Reading task (green). Bin size = 0.05. P values in black denote interaction

701   statistics whereas P values in blue and green denote the statistics for the Stroop and

702   Reading tasks respectively. As discussed in **Figure 4**, the average log-ratios are

703   presented here for display purposes only and the statistical tests are based on the raw

704   power values.

705   (**B**) Distribution of the gamma power log-ratio between incongruent trials preceded

706   by congruent trials (cI) compared to incongruent trials preceded by incongruent trials

707   (iI).

708   (**C**) Distribution of the gamma power log-ratio between congruent trials preceded by

709   incongruent trials (iC) compared to congruent trials preceded by congruent trials (cC).

710   (**D-F**) Same as (A-C), but for power in the beta band.

711

712   **Figure 4-figure supplement 2. Cross-frequency coupling analyses**

713   For the anterior cingulate cortex electrode in **Figure 2**:

714   (**A**) Phase-amplitude distribution during the Stroop task for the anterior cingulate

715   example electrode shown in **Figure 2** (see **Methods** for calculation of cross-

716   frequency coupling).

717   (**B**) The observed Modulation Index (MI, black arrow) is significantly greater than the

718   surrogate distribution generated by adding a lag between the phase and amplitude

719   measurements, demonstrating that the amplitude of the gamma band is strongly

720   coupled to the phase of the theta band.

721 (**C**) During the Stroop task, the difference in Modulation Index between congruent

722 and incongruent trials (black arrow) was not significantly different from 0 ($P = 0.61$).

723 The null distribution (gray bars) was generated by randomly permuting the congruent

724 and incongruent labels.

725 Across the population of electrodes:

726 (**D**) The percent of total electrodes in each region (Frontal cortex or non-frontal

727 cortex) that had significant phase-amplitude coupling. Shown on the right is the

728 percentage of the n=51 conflict selective electrodes that showed significant coupling.

729 (**E**) The MI of congruent compared to incongruent trials for all Frontal cortex

730 electrodes (gray dots) and the subset that were conflict-selective in the gamma band

731 (blue dots). For both groups, there was no significant difference in the MI between

732 congruent and incongruent trials (Frontal Cortex, $P = 0.45$; Conflict-selective, $P =$

733 0.52; signed-rank test). For this comparison, the number of congruent and

734 incongruent trials was equalized before computing the MI.

735

736 **Figure 4-figure supplement 3. Stimulus-aligned population averages**

737 Same as in **Figure 4B**, but data are aligned to the stimulus response onset (t=0).

738

739 **Figure 5. Responses during self-corrected error trials**

740 (**A**) An example self-correction trial from the ACC electrode in **Figure 2** when the

741 word Green colored in red was presented. The single trial gamma power is shown on

742 top, with the speech waveform below. The dashed lines indicate the onset of the

743 initially incorrect response ("green") and the following corrected response in bold ("no

744 – red"). Note the increased gamma power after an error response.

745 (**B**) Average gamma power aligned to the onset of the initial error response (blue) and

746 the onset of the corrected response (black) for n=11 self-correction trials. Shaded areas

747 indicate s.e.m. The post-response power was significantly greater after the error ($P =$

748 0.001, signed-rank test).

749 (**C-D**) Same as (**A-B**) for another example electrode in the dorsolateral prefrontal

750 cortex. The post-response power was significantly greater after the error response ($P =$

751 0.002, signed-rank test).

752   **(E)** Across the n=7 electrodes with n=10 or greater self-correction trials, the z-scored

753   gamma power during the initial error response was larger than during the corrected

754   response. Electrodes with significant differences ($P < 0.05$, signed-rank test) are

755   colored black. Letters mark the examples in (**A**) and (**C**).

756

757   **Figure 6. Latency Comparisons across regions**

758   Latency differences between different regions computed from all pairs of simultaneously

759   recorded electrodes. n_p denotes the number of electrode pairs. Because we only consider

760   simultaneously recorded electrodes here, not all the electrodes modulated by conflict can

761   be paired with any other electrode. Supplementary File 3 shows the number of electrodes

762   modulated by conflict in each area and subject. There was only one electrode pair

763   between ACC and OFC and therefore we do not show the latency difference between

764   these two regions here. Significant latency differences ($P < 0.05$, permutation test,

765   **Methods**) are shown in black, and non-significant differences in gray. ACC leads both

766   mFC ($P = 0.001$) and dlPFC ($P = 0.02$), with OFC following dlPFC ($P = 0.009$).

767

768   **Supplementary Files Legends**

769

770   **Supplementary File 1**. Information about each of the 15 subjects that participated in this

771   study

772

773   **Supplementary File 2**. Information about each of the 51 electrodes that were modulated

774   by conflict.

775

776   **Supplementary File 3**. Distribution of the 51 electrodes that were modulated by conflict

777   across the four frontal cortex areas and the 15 subjects in our study.

778

779   **Supplementary File 4**. Statistical reporting table describing the statistical tests used

780   throughout the text, including the corresponding n, the number of degrees of freedom, p

781   values and effect sizes.

782

783   **Source data**

784   We provide source data and code to reproduce the work reported here.

785   figure1_sourceData.mat: Figure 1-source data 1

786   figure2_sourceData.mat: Figure 2-source data 1

787   figure3+4_sourceData.mat: Figures 3 and 4-source data 1

788   figure5_sourceData.mat: Figure 5-source data 1

789   figure6_sourceData.mat: Figure 6-source data 1

790   plotFigure2A.m   Source code 1: Code to plot Figure 2A (with minor

791   modifications, this code can also be used to plot the other example figures in the text)

792   plotFigure4.m   Source code 2:Code to plot Figure 4

793   shadedErrorBar.m   Source code 3:Code used to plot error bars in the

794   plotFigure2A.m

795 **References**

796

797 Aarts, E., Roelofs, A., and van Turennout, M. (2008). Anticipatory activity in anterior
798 cingulate cortex can be independent of conflict and error likelihood. J Neurosci 28,
799 4671-4678.
800 Aarts, E., Verhage, M., Veenvliet, J.V., Dolan, C.V., and van der Sluis, S. (2014). A
801 solution to dependency: using multilevel analysis to accommodate nested data. Nat
802 Neurosci 17, 491-496.
803 Bonini, F., Burle, B., Liegeois-Chauvel, C., Regis, J., Chauvel, P., and Vidal, F. (2014).
804 Action monitoring and medial frontal cortex: leading role of supplementary motor
805 area. Science 343, 888-891.
806 Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., and Cohen, J.D. (1999). Conflict
807 monitoring versus selection-for-action in anterior cingulate cortex. Nature 402, 179-
808 181.
809 Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S., and Cohen, J.D. (2001). Conflict
810 monitoring and cognitive control. Psychological review 108, 624.
811 Bouchard, K.E., Mesgarani, N., Johnson, K., and Chang, E.F. (2013). Functional
812 organization of human sensorimotor cortex for speech articulation. Nature 495,
813 327-332.
814 Brown, J.W., and Braver, T.S. (2005). Learned predictions of error likelihood in the
815 anterior cingulate cortex. Science 307, 1118-1121.
816 Bugg, J.M., Jacoby, L.L., and Toth, J.P. (2008). Multiple levels of control in the Stroop
817 task. Memory & cognition 36, 1484-1494.
818 Buzsaki, G., Anastassiou, C.A., and Koch, C. (2012). The origin of extracellular fields
819 and currents--EEG, ECoG, LFP and spikes. Nature reviews Neuroscience 13, 407-
820 420.
821 Canolty, R.T., Edwards, E., Dalal, S.S., Soltani, M., Nagarajan, S.S., Kirsch, H.E., Berger,
822 M.S., Barbaro, N.M., and Knight, R.T. (2006). High gamma power is phase-locked to
823 theta oscillations in human neocortex. Science 313, 1626-1628.
824 Carter, C.S., Braver, T.S., Barch, D.M., Botvinick, M.M., Noll, D., and Cohen, J.D. (1998).
825 Anterior cingulate cortex, error detection, and the online monitoring of
826 performance. Science 280, 747-749.
827 Carter, C.S., Macdonald, A.M., Botvinick, M., Ross, L.L., Stenger, V.A., Noll, D., and
828 Cohen, J.D. (2000). Parsing executive processes: strategic vs. evaluative functions of
829 the anterior cingulate cortex. Proceedings of the National Academy of Sciences of
830 the United States of America 97, 1944-1948.
831 Cavanagh, J.F., Cohen, M.X., and Allen, J.J. (2009). Prelude to and resolution of an
832 error: EEG phase synchrony reveals cognitive control dynamics during action
833 monitoring. J Neurosci 29, 98-105.
834 Cavanagh, J.F., and Frank, M.J. (2014). Frontal theta as a mechanism for cognitive
835 control. Trends Cogn Sci 18, 414-421.
836 Cohen, J.D., and Servan-Schreiber, D. (1992). Context, cortex, and dopamine: a
837 connectionist approach to behavior and biology in schizophrenia. Psychological
838 review 99, 45.

839   Cole, M.W., Yeung, N., Freiwald, W.A., and Botvinick, M. (2009). Cingulate cortex:
840   diverging data from humans and monkeys. Trends Neurosci 32, 566-574.
841   Crone, N.E., Miglioretti, D.L., Gordon, B., and Lesser, R.P. (1998a). Functional
842   mapping of human sensorimotor cortex with electrocorticographic spectral analysis.
843   II. Event-related synchronization in the gamma band. Brain 121 (Pt 12), 2301-2315.
844   Crone, N.E., Miglioretti, D.L., Gordon, B., Sieracki, J.M., Wilson, M.T., Uematsu, S., and
845   Lesser, R.P. (1998b). Functional mapping of human sensorimotor cortex with
846   electrocorticographic spectral analysis. I. Alpha and beta event-related
847   desynchronization. Brain 121 (Pt 12), 2271-2299.
848   Destrieux, C., Fischl, B., Dale, A., and Halgren, E. (2010). Automatic parcellation of
849   human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage
850   53, 1-15.
851   Ebitz, R.B., and Platt, M.L. (2015). Neuronal activity in primate dorsal anterior
852   cingulate cortex signals task conflict and predicts adjustments in pupil-linked
853   arousal. Neuron 85, 628-640.
854   Egner, T., and Hirsch, J. (2005a). Cognitive control mechanisms resolve conflict
855   through cortical amplification of task-relevant information. Nat Neurosci 8, 1784-
856   1790.
857   Egner, T., and Hirsch, J. (2005b). The neural correlates and functional integration of
858   cognitive control in a Stroop task. Neuroimage 24, 539-547.
859   Emeric, E.E., Leslie, M., Pouget, P., and Schall, J.D. (2010). Performance monitoring
860   local field potentials in the medial frontal cortex of primates: supplementary eye
861   field. J Neurophysiol 104, 1523-1537.
862   Fellows, L.K., and Farah, M.J. (2005). Is anterior cingulate cortex necessary for
863   cognitive control? Brain 128, 788-796.
864   Goldstein, H. (2011). Multilevel statistical models, 4th edn (Chichester, West Sussex:
865   Wiley).
866   Gratton, G., Coles, M.G.H., and Donchin, E. (1992). Optimizing the use of information:
867   Strategic control of activation of responses. Journal of Experimental Psychology:
868   General 121, 480.
869   Hughes, G., and Yeung, N. (2011). Dissociable correlates of response conflict and
870   error awareness in error-related brain activity. Neuropsychologia 49, 405-415.
871   Ito, S., Stuphorn, V., Brown, J.W., and Schall, J.D. (2003). Performance monitoring by
872   the anterior cingulate cortex during saccade countermanding. Science 302, 120-122.
873   Johnston, K., Levin, H.M., Koval, M.J., and Everling, S. (2007). Top-down control-
874   signal dynamics in anterior cingulate and prefrontal cortex neurons following task
875   switching. Neuron 53, 453-462.
876   Kahana, M.J., Seelig, D., and Madsen, J.R. (2001). Theta returns. Current opinion in
877   neurobiology 11, 739-744.
878   Kerns, J.G. (2006). Anterior cingulate and prefrontal cortex activity in an FMRI study
879   of trial-to-trial adjustments on the Simon task. Neuroimage 33, 399-405.
880   Kerns, J.G., Cohen, J.D., MacDonald, A.W., Cho, R.Y., Stenger, V.A., and Carter, C.S.
881   (2004). Anterior cingulate conflict monitoring and adjustments in control. Science
882   303, 1023-1026.
883   Kolling, N., Behrens, T.E., Mars, R.B., and Rushworth, M.F. (2012). Neural
884   mechanisms of foraging. Science 336, 95-98.

885   Liu, H., Agam, Y., Madsen, J.R., and Kreiman, G. (2009). Timing, timing, timing: fast
886   decoding of object information from intracranial field potentials in human visual
887   cortex. Neuron 62, 281-290.
888   Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001).
889   Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150-
890   157.
891   MacDonald, A.W. (2000). Dissociating the Role of the Dorsolateral Prefrontal and
892   Anterior Cingulate Cortex in Cognitive Control. Science 288, 1835-1838.
893   Mansouri, F.A., Buckley, M.J., and Tanaka, K. (2007). Mnemonic function of the
894   dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. Science
895   318, 987-990.
896   Milham, M.P., and Banich, M.T. (2005). Anterior cingulate cortex: an fMRI analysis of
897   conflict specificity and functional differentiation. Human brain mapping 25, 328-
898   335.
899   Milham, M.P., Banich, M.T., Claus, E.D., and Cohen, N.J. (2003). Practice-related
900   effects demonstrate complementary roles of anterior cingulate and prefrontal
901   cortices in attentional control. Neuroimage 18, 483-493.
902   Miller, E.K. (2000). The prefontral cortex and cognitive control. Nature reviews
903   neuroscience 1, 59-65.
904   Miller, E.K., and Cohen, J.D. (2001). An integrative theory of prefrontal cortex
905   function. Annu Rev Neurosci 24, 167-202.
906   Mitra, P., and Bokil, H. (2008). Observed brain dynamics (Oxford ; New York: Oxford
907   University Press).
908   Mitzdorf, U. (1987). Properties of the evoked potential generators: current source-
909   density analysis of visually evoked potentials in the cat cortex. The International
910   journal of neuroscience 33, 33-59.
911   Nakamura, K., Roesch, M.R., and Olson, C.R. (2005). Neuronal activity in macaque
912   SEF and ACC during performance of tasks involving conflict. J Neurophysiol 93, 884-
913   908.
914   Nee, D.E., Kastner, S., and Brown, J.W. (2011). Functional heterogeneity of conflict,
915   error, task-switching, and unexpectedness effects within medial prefrontal cortex.
916   Neuroimage 54, 528-540.
917   Niendam, T.A., Laird, A.R., Ray, K.L., Dean, Y.M., Glahn, D.C., and Carter, C.S. (2012).
918   Meta-analytic evidence for a superordinate cognitive control network subserving
919   diverse executive functions. Cognitive, affective & behavioral neuroscience 12, 241-
920   268.
921   Nieuwenhuis, S., Forstmann, B.U., and Wagenmakers, E.J. (2011). Erroneous
922   analyses of interactions in neuroscience: a problem of significance. Nat Neurosci 14,
923   1105-1107.
924   Oehrn, C.R., Hanslmayr, S., Fell, J., Deuker, L., Kremers, N.A., Do Lam, A.T., Elger, C.E.,
925   and Axmacher, N. (2014). Neural communication patterns underlying conflict
926   detection, resolution, and adaptation. J Neurosci 34, 10438-10452.
927   Perrett, E. (1974). The left frontal lobe of man and the suppression of habitual
928   responses in verbal categorical behavior. Neuropsychologia 12, 323-330.
929   Ray, S., and Maunsell, J.H. (2011). Different origins of gamma rhythm and high-
930   gamma activity in macaque visual cortex. PLoS biology 9, e1000610.

931   Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., and Nieuwenhuis, S. (2004). The role
932   of the medial frontal cortex in cognitive control. Science 306, 443-447.
933   Rothé, M., Quilodran, R., Sallet, J., and Procyk, E. (2011). Coordination of high gamma
934   activity in anterior cingulate and lateral prefrontal cortical areas during adaptation.
935   J Neurosci 31, 11110-11117.
936   Rushworth, M.F., Walton, M.E., Kennerley, S.W., and Bannerman, D.M. (2004). Action
937   sets and decisions in the medial frontal cortex. Trends Cogn Sci 8, 410-417.
938   Shenhav, A., Botvinick, M.M., and Cohen, J.D. (2013). The expected value of control:
939   an integrative theory of anterior cingulate cortex function. Neuron 79, 217-240.
940   Sheth, S.A., Mian, M.K., Patel, S.R., Asaad, W.F., Williams, Z.M., Dougherty, D.D., Bush,
941   G., and Eskandar, E.N. (2012). Human dorsal anterior cingulate cortex neurons
942   mediate ongoing behavioural adaptation. Nature 488, 218-221.
943   Stroop, J.R. (1935). Studies of interference in serial verbal reactions. Journal of
944   Experimental Psychology 18, 643-662.
945   Swick, D., and Turken, A.U. (2002). Dissociation between conflict detection and error
946   monitoring in the human anterior cingulate cortex. Proceedings of the National
947   Academy of Sciences of the United States of America 99, 16354-16359.
948   Taylor, S.F., Martis, B., Fitzgerald, K.D., Welsh, R.C., Abelson, J.L., Liberzon, I., Himle,
949   J.A., and Gehring, W.J. (2006). Medial frontal cortex activity and loss-related
950   responses to errors. J Neurosci 26, 4063-4070.
951   Tort, A.B., Kramer, M.A., Thorn, C., Gibson, D.J., Kubota, Y., Graybiel, A.M., and Kopell,
952   N.J. (2008). Dynamic cross-frequency couplings of local field potential oscillations in
953   rat striatum and hippocampus during performance of a T-maze task. Proceedings of
954   the National Academy of Sciences of the United States of America 105, 20517-
955   20522.
956   Ullsperger, M., Danielmeier, C., and Jocham, G. (2014). Neurophysiology of
957   performance monitoring and adaptive behavior. Physiological reviews 94, 35-79.
958   van Driel, J., Swart, J.C., Egner, T., Ridderinkhof, K.R., and Cohen, M.X. (2015). (No)
959   time for control: Frontal theta dynamics reveal the cost of temporally guided conflict
960   anticipation. Cognitive, affective & behavioral neuroscience.
961   von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of
962   cortical integration: from local gamma to long range alpha/theta synchronization.
963   International journal of psychophysiology : official journal of the International
964   Organization of Psychophysiology 38, 301-313.
965   Weiskopf, N., Hutton, C., Josephs, O., and Deichmann, R. (2006). Optimal EPI
966   parameters for reduction of susceptibility-induced BOLD sensitivity losses: a whole-
967   brain analysis at 3 T and 1.5 T. Neuroimage 33, 493-504.
968   Yeung, N., Botvinick, M.M., and Cohen, J.D. (2004). The neural basis of error
969   detection: conflict monitoring and the error-related negativity. Psychol Rev 111,
970   931-959.
971

Figure 1

a

500 ms 2 s 500 ms 2 s

b

Stroop

c

Reading

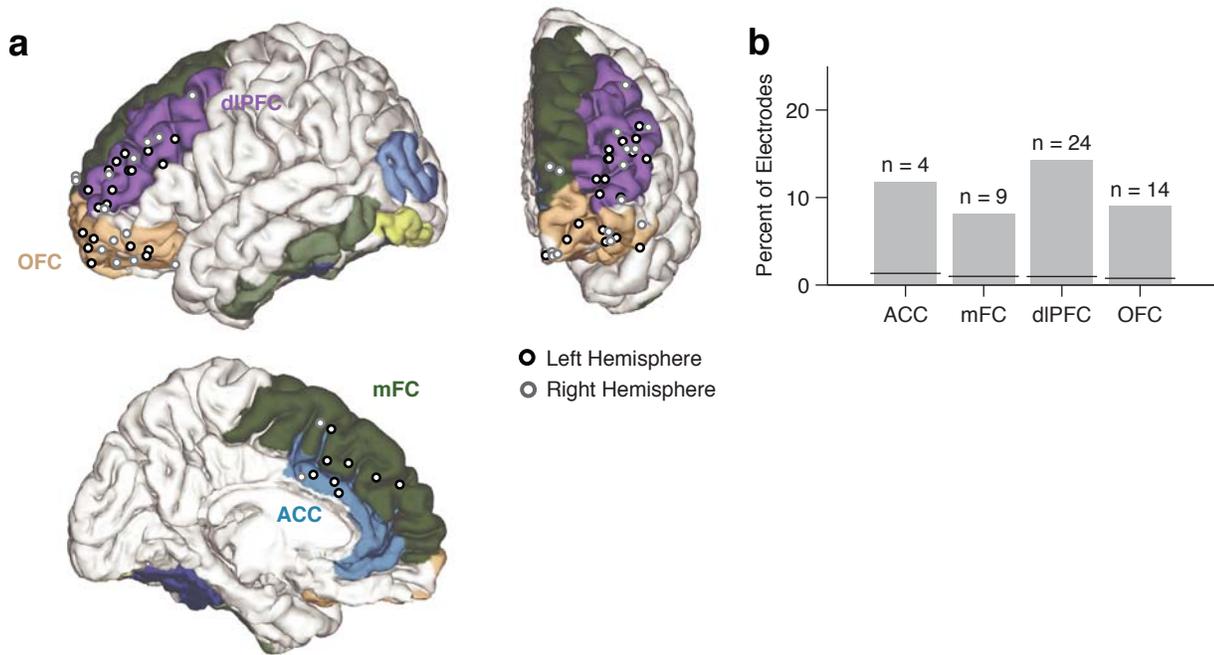**d**
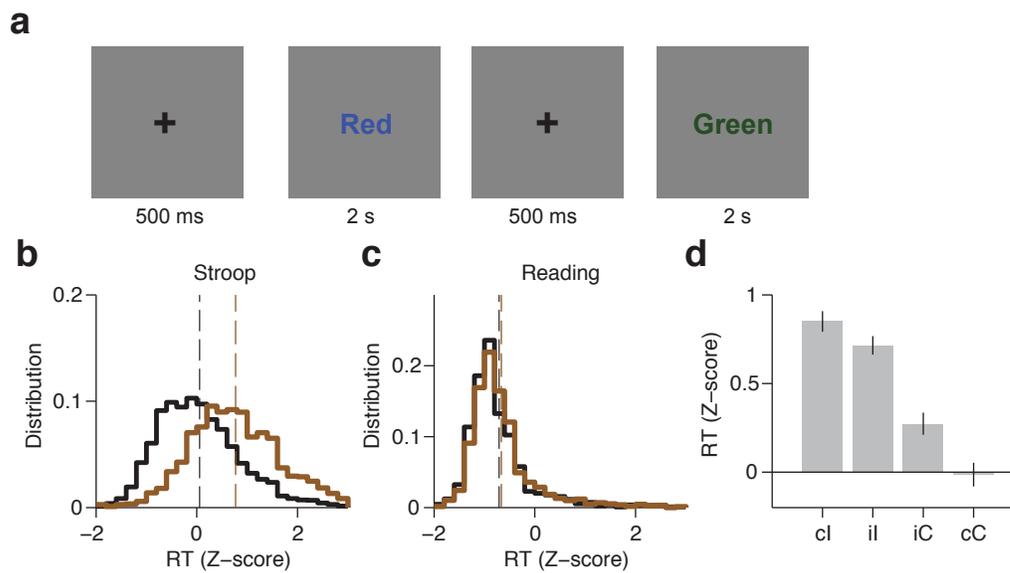
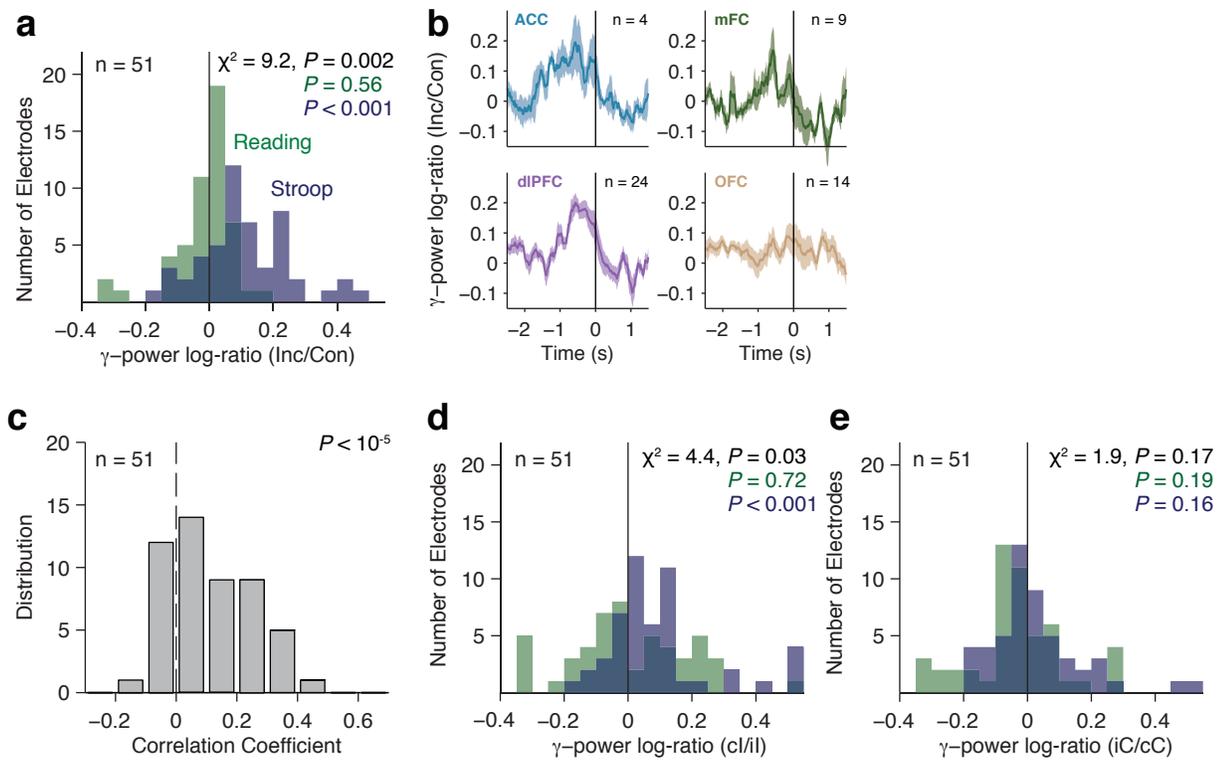Tangetal_fig1.pdf
Created in: Adobe Illustrator CS6 (Macintosh)

Figure 2

Tangetal_fig2.pdf
Created in: Adobe Illustrator CS6 (Macintosh)

Figure 3

a

dlPFC

OFC

mFC

ACC

○ Left Hemisphere
◎ Right Hemisphere

b

Percent of Electrodes

n = 4
n = 9
n = 24
n = 14

20

10

0

ACC   mFC   dlPFC   OFC

Tangetal_fig3.pdf
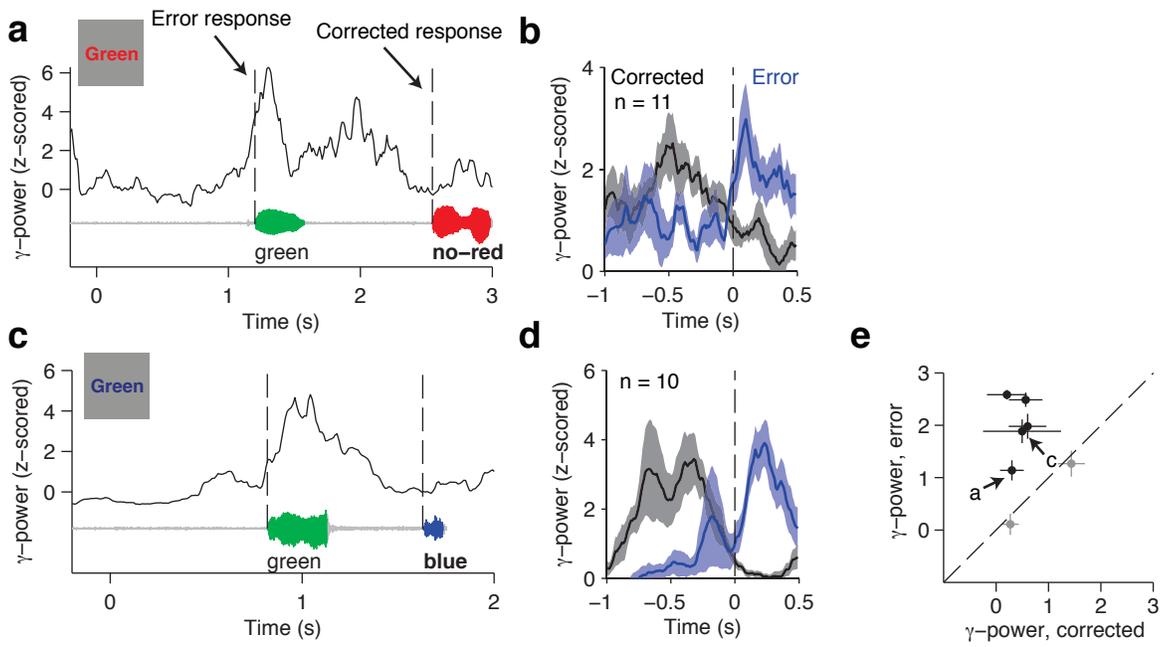Created in: Adobe Illustrator CS6 (Macintosh)

Figure 4

Tangetal_fig4.pdf
Created in: Adobe Illustrator CS6 (Macintosh)

Figure 5

Tangetal_fig5.pdf
Created in: Adobe Illustrator CS6 (Macintosh)

Figure 6

ACC

207 ms
$n_p = 6$

388 ms
$n_p = 10$

dlPFC
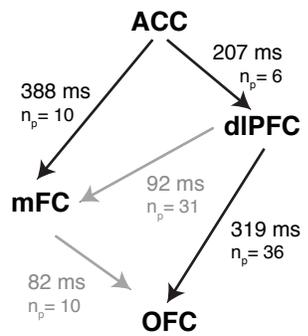
92 ms
$n_p = 31$

mFC

319 ms
$n_p = 36$

82 ms
$n_p = 10$

OFC

Tangetal_fig6.pdf
Created in: Adobe Illustrator
CS6 (Macintosh)