Correlates of auditory decision making in prefrontal, auditory, and basal lateral amygdala cortical areas

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.jneurosci.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

1 # Correlates of auditory decision making in prefrontal, auditory, and basal lateral
2 # amygdala cortical areas

3 Julia L. Napoli[*], Corrie R. Camalier[*][†], Anna Leigh Brown, Jessica Jacobs, Mortimer M. Mishkin and Bruno
4 B. Averbeck[†]

5 [1]*Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health,*
6 *Bethesda MD, 20892-4415*

7

8 **[†]Correspondence to:**
9 Corrie R. Camalier, Ph.D.
10 Laboratory of Neuropsychology, NIMH/NIH
11 Building 49 Room 1B80
12 49 Convent Drive MSC 4415
13 Bethesda, MD 20892-4415
14 E-mail: corrie.camalier@duke.edu
15
16 Bruno B. Averbeck, Ph.D.
17 Laboratory of Neuropsychology, NIMH/NIH
18 Building 49 Room 1B80
19 49 Convent Drive MSC 4415
20 Bethesda, MD 20892-4415
21 E-mail: bruno.averbeck@nih.gov
22
23
24 *These authors contributed equally.
25

30

31 Abstract

32   Spatial selective listening and auditory choice underlie important processes including attending
33 to a speaker at a cocktail party and knowing how (or if) to respond. To examine task encoding and
34 relative timing of potential neural substrates underlying these behaviors, we developed a spatial
35 selective detection paradigm for monkeys, and recorded activity in primary auditory cortex (AC),
36 dorsolateral prefrontal cortex (dlPFC) and the basolateral amygdala (BLA). A comparison of neural
37 responses among these three areas showed that, as expected, AC encoded the side of the cue and
38 target characteristics before dlPFC and BLA. Interestingly, AC also encoded the monkey's choice before
39 dlPFC and around the time of BLA. Generally, BLA showed weak responses to all task features except
40 the choice. Decoding analyses suggested that errors followed from a failure to encode the target
41 stimulus in both AC and dlPFC, but again, these differences arose earlier in AC. The similarities between
42 AC and dlPFC responses were abolished during passive sensory stimulation with identical trial
43 conditions, suggesting that the robust sensory encoding in dlPFC is contextually gated. Thus, counter to
44 a strictly PFC-driven decision process, in this spatial selective listening task, AC neural activity represents
45 the sensory and decision information before dlPFC. Unlike in the visual domain, in this auditory task, the
46 BLA does not appear to be robustly involved in selective spatial processing.
47
48

49 Significance Statement

50   We examined neural correlates of an auditory spatial selective listening task by
51 recording single neuron activity in behaving monkeys from the amygdala, dorsal-lateral
52 prefrontal cortex, and auditory cortex. We found that auditory cortex coded spatial cues and
53 choice-related activity before dorsal-lateral prefrontal cortex or the amygdala. Auditory
54 cortex also had robust delay period activity. Therefore, we found that auditory cortex could
55 support the neural computations that underlie the behavioral processes in the task.
56
57
58
59
60
61

Introduction

Spatial selective listening is critical for solving everyday problems including the classic "cocktail party problem", which requires attending to one sound source amidst a noisy background of competing sources (Cherry, 1953). Common auditory spatial selective listening paradigms used in research with humans include modified Posner paradigms in which subjects detect auditory stimuli after being cued to a spatial location (Spence and Driver, 1994; Alho et al., 1999; McDonald and Ward, 1999; Mayer et al., 2007; Mayer et al., 2009; Roberts et al., 2009; Teshiba et al., 2013) and selective listening studies (Ahveninen et al., 2013; Frey et al., 2014; Bidet-Caulet et al., 2015). Previous work in humans has shown that auditory cortex (AC) plays an important role in spatial selective listening tasks, through interactions with prefrontal (Alho et al., 1999) and parietal (Deng et al., 2019) cortex. In addition to a role for these structures, previous studies in the visual domain in non-human primates have shown that the BLA contributes to spatial selective attention (Peck and Salzman, 2014; Costa et al., 2019).

There are only a few studies comparing multiple areas in auditory processes, especially in non-human primates, so we lack clear evidence on the relative contributions and timing of information between areas. Auditory processing is characterized by speed, especially relative to the visual system. In non-human primates, primary auditory cortex (AC) has response latencies of ~20 ms (Camalier et al., 2012), compared to ~ 40 ms for primary visual cortex (Schmolesky et al., 1998). AC is, however, further removed from the peripheral sensory receptors than primary visual cortex. This speed is consistent with a hypothesized role for the auditory system in rapid spatial alerting or orienting. However, the processing depth of primary AC has led some authors to suggest that it can also process cognitive factors such as choice, normally attributed to higher-order sensory areas (Naatanen et al., 2001; Nelken, 2004). Certainly, AC has been shown to reflect aspects of auditory decision making beyond sensory processing (Niwa et al., 2012; Tsunada et al., 2016; Christison-Lagay and Cohen, 2018; Huang et al., 2019), but it is unclear if this choice information is coming from PFC or another area (Lee et al., 2009; Plakke et al., 2015). A recent decision-making study in ferrets suggested that sensory information was encoded first in primary AC, but category information and the decision was encoded first in ferret dlFC, which is a premotor area potentially analogous to primate PFC (Yin et al., 2020). This would be consistent with auditory working memory data in non-human primates which suggests that a categorical "match" decision may emerge earlier in ventral PFC than AC (Bigelow et al., 2014). At present, the relative role of AC and dlPFC, especially in spatial decision making, is unclear. Aside from the cortical sensory and prefrontal pathways, a BLA pathway for spatially selective processing and decision making is hypothesized to be fairly fast in the visual domain (Peck and Salzman, 2014; Costa et al., 2019), but whether the BLA is involved in auditory decision making in non-human primates is unknown.

To address these outstanding questions, here we describe an experiment in which we used a spatial selective detection paradigm for monkeys, grounded in spatially cued listening tasks used in the human studies discussed above. To investigate potential neural correlates of this task, we recorded single unit activity in primary AC (A1), dlPFC, and the BLA while the monkeys carried out the task. The dlPFC recordings were located in dorsal pre-arcuate cortex (primarily area 8A, see methods; Fig. 1B), which is the primary prefrontal target of the auditory "dorsal stream" arising from caudal belt and

102   parabelt, thought to be important for auditory spatial processing (Bon and Lucchetti, 1994; Hackett et
103   al., 1999; Lanzilotto et al., 2013). Specifically these recordings targeted the zone between the principal
104   sulcus and dorsal arcuate sulcus, at least 1 mm away from the arcuate sulcus, primarily corresponding to area 8Ad,
105   but also potentially including the dorsal bank of 46d, caudal 8Adv, and caudal border of 8b. Recordings in the
106   amygdala targeted to the basal and lateral nuclei. Cortical auditory inputs to the amygdala caudal
107   parabelt, terminate in the larger lateral nucleus (Yukie, 2002). However, the rostral superior temporal
108   gyrus, which also indirectly receive auditory input, projects more broadly to the lateral and basal nuclei
109   (Stefanacci and Amaral, 2002). We examined the strength and latency of signals at the single cell level
110   related to the task across these areas. In AC and dlPFC a substantial fraction of neurons was selective to
111   the location of the cue and the subsequent target. We found that AC preceded both dlPFC and BLA in
112   sensory discrimination and also in the decision. Classification analyses of firing rate patterns in error
113   trials indicated that errors during the task were usually the result of a failure to encode the first target
114   stimulus in AC and also in dlPFC. A comparison of responses and timing with a control "passive listening"
115   condition showed that sensory target related activity in dlPFC was almost completely abolished in the
116   passive task, suggesting task-dependent gating of information to areas beyond sensory cortex.

117   **Methods**
118   The experiments were carried out using two adult male rhesus macaques (Macaca mulatta). The
119   monkeys had access to food 24 hours a day and earned their liquid through task performance on testing
120   days. Monkeys were socially pair housed. All procedures were reviewed and approved by the NIMH
121   Animal Care and Use Committee.

122   *Experimental Setup*

123   The monkeys were operantly trained to perform a spatial selective listening paradigm. The task
124   was controlled by custom software (Tucker Davis Technologies (TDT) System 3: OpenWorkbench and
125   OpenDeveloper, TDT) which controlled multi-speaker sound delivery and acquired bar presses and eye
126   movements. Eye movements were tracked using an Arrington Viewpoint eye tracking system (Arrington
127   research) sampled at 1 kHz. Monkeys were seated in a primate chair facing a 19-Inch LCD monitor 40 cm
128   from the monkey's eyes, on which the visual fixation spot was presented. Monkeys performed the task
129   in a darkened, double-walled acoustically isolated sound booth (Industrial Acoustics Company, Bronx,
130   NY). All auditory stimuli were presented from a speaker 10 cm from the left or right of the monkey's
131   head. Juice rewards were delivered using a solenoid juice delivery system (Crist Instruments).

132   *Task Design and Stimuli*

133   The monkeys carried out a spatial selective listening task (Fig. 1), modeled after spatially cued
134   tasks used in humans. The task required oculomotor fixation throughout the duration of the trial. Both
135   spatial cues and target stimuli were auditory and the monkeys were required to respond when they
136   detected a target embedded in masking noise presented on the cued side. Listening conditions (listen
137   left/right) were blocked with two types of trials (match/foil) in each condition. At the start of each trial,
138   the monkey was prompted to press a lever and fixate a central point on the screen. After a short delay
139   (2.1-2.4 s), a 50 ms 4 kHz square wave (70 dB) cue was played from a speaker on the left or right of the

140   midline.   Frozen diotic white noise (40 dB) was then played from both the left and right speakers from

141   500 ms after the initial cue until the lever was released. Following a variable delay after noise onset

142   (500, 800 or 1300 ms) a 300 ms 1 KHz square wave target sound was played from either the left or right

143   speaker. If the target sound was on the same side as the cue, it was a match trial and the animal had to

144   release the lever within 700 ms to receive a juice reward. If the target sound was on the opposite side

145   as the cue, it was a foil trial and the monkey had to continue to hold the lever. Following a second

146   interval of 800 or 1000 ms in foil trials, a second 1 KHz match target was always played on the same side

147   as the original cue. If the animal correctly released the lever following the second target in foil trials it

148   was given a juice reward. Thus, both match and foil trials were identical in terms of reward expectation.

149   If the choice was incorrect, there was a long "timeout" period before the next trial could be initiated. As

150   in our previous work (Camalier et al., 2019) the use of square waves (which contain odd harmonics)

151   allowed for wideband stimulation that was perceptually distinct, but whose broad spectral signature

152   robustly activated large swaths of AC in a way that pure tones would not. Thus, similar to human

153   paradigms, the stimuli could be kept identical across all sessions, independent of where recordings were

154   carried out in AC, and data could be collapsed across sessions for analysis.

155   To achieve maximal effort and selective effects on neurons (as well as be able to analyze sources

156   of errors), it was important that the targets be difficult to detect. Thus, several psychometric quality

157   controls were included to ensure that the monkeys were consistently performing the task across

158   sessions. The sound level of the target for the two monkeys was individually titrated to maintain

159   performance at ~70-80% correct (exact 71.14%). Thus, the detection was difficult. The cue presentation

160   was blocked to ensure the monkeys were able to maintain high accuracy on the task, as complex

161   auditory tasks in monkey have traditionally been difficult to condition operantly (Scott and Mishkin,

162   2016; Rinne et al., 2017). Analysis of the first trial after the cue switched sides showed that animals

163   were correct 76.13% of the time, indicating the monkeys were primarily using the cue in the task. For

164   monkey 1, target tones were delivered at levels between 16-40 dB, with most tones in the 17-24 dB

165   range. For monkey 2 target tones were delivered at levels between 27-45 dB, with most tones in the 29-

166   36 dB range. Within a session, the target sound varied 0-7 dB from trial to trial to ensure that the

167   monkeys were responding to the target side and not to consistencies in (or guessing based on) sound

168   level differences between speakers that may have resulted from otherwise undetectable differences in

169   calibration between the two speakers. To further ensure accurate performance, periodic "catch trials"

170   (~10% with a 0-dB target tone) were included to ensure that the monkeys were responding to the target

171   and not timing their choices relative to the presentation of the cue or noise. To encourage motivation

172   during foil trials (which were longer duration & were thus more likely to be aborted), the "match/bar

173   release" target after a foil sound was louder (and easier) than typical target sounds for each monkey

174   (monkey 1: 27 or 30 dB; monkey 2: 35 or 40dB).

175

176   Before the task was run, a battery of passive listening and mapping stimuli were played. Within

177   this battery was a control condition of "passive listening". In this condition the monkey was presented

178   with the task stimuli with trial types and timing matched to the selective listening task. However, the

179   animals did not press or release a bar, fixate, or receive juice rewards. This task allowed us to compare

180   sensory responses between active and passive task conditions. Monkeys were cued that this was a

181 passive condition as they did not have access to the lever or juice tube, and it was done as part of a
182 passive-listening mapping battery, consistently before the start of the active task.

183 *Neurophysiological Recordings*

184 Monkeys were implanted with titanium headposts for head restraint before data collection
185 began. Custom 45 x 24 mm acrylic chambers were designed and fitted to the monkeys in a separate
186 procedure. The chamber was aligned with the long axis oriented anterior-posterior. The placement
187 allowed vertical grid access to the left dorsolateral prefrontal cortex (Fig. 1B; dorsal bank of the principal
188 sulcus extending to dorsal arcuate but at least , >1mm away from arcuate sulcus, primarily corresponding
189 to area 46/8Ad, but also potentially including dorsal bank of 46d, caudal 8Adv, and caudal border of 8b), the basal
190 and lateral portions of the amygdala (entire dorsoventral extent), and auditory cortex (primarily A1 but
191 including small portions of surrounding areas). A 1 mm grid was located inside the chamber for
192 targeting (Fig. 1B, lower-right), and all penetrations were dorsal-ventral. This dorso-ventral trajectory
193 was essential for targeting AC tonotopic reversals. The chamber was custom fit to a 3D print of each
194 monkey's skull generated by a CT scan before implantation. Recording areas were verified through a T1
195 scan of grid coverage with respect to underlying anatomical landmarks (Fig. 1B), combined with maps of
196 frequency reversals and response latencies of single neurons to determine A1 location and extent
197 (Camalier et al., 2012; Camalier et al., 2019). Recordings were mainly carried out in simultaneous AC and
198 PFC sessions, with BLA sessions occurring later in the experiment, but some data included were from
199 just one, or even three simultaneously recorded areas in a given session. We recorded the activity of
200 2,387 single neurons during the task (N = 847 (AC), N = 968 (dlPFC), and N = 572 (BLA) across monkeys 1
201 (N = 1540) and 2(N = 847)).

202 In both monkeys, we recorded using either 16 or 24 channel laminar "V-trodes" (Plexon, Inc,
203 Dallas TX; 200-300μm contact spacing, respectively). The electrodes allowed for identification of white
204 matter tracts, further allowing identification of electrode location with respect to sulci and gyri. To
205 ensure vtrodes went as straight as possible, sharpened guide tubes for the buried structures (AMY, AC)
206 were advanced 10-15 mm above the structures. This was not possible for the PFC as it is a surface
207 structure, but a guide tube was used to puncture overlying granulation tissue to permit a Vtrode to
208 advance. Electrodes were advanced through the guide tubers to their target location (NAN microdrives,
209 Nazareth, Israel) and allowed to settle for at least 1 hour before recording. Neural activity was recorded
210 either primarily simultaneously (AC and PFC) or primarily individually (BLA), although there were some
211 sessions in which all 3 areas were recorded from.

212 Multichannel spike and local field potential recordings were acquired with a 64-channel Tucker
213 Davis Technology data acquisition system. Spike signals were amplified, filtered (0.3-8kHz), and digitized
214 at ~24.4 kHz. Spikes were initially sorted online on all-channels using real-time window discrimination.
215 Digitized spike waveforms and timestamps of stimulus events were saved for sorting offline (Plexon
216 sorter V 3.3.5). Units were graded according to isolation quality (single or multiunit neurons). Single and
217 multiunit recordings were analyzed separately, but patterns were similar, so they were combined. The
218 acquisition software interfaced directly with the stimulus delivery system and both systems were
219 controlled by custom software (OpenWorkbench and OpenDeveloper, controlling a RZ2, RX8, Tucker

220   Davis Technologies (TDT) System 3, Alachua, FL). For inclusion in analysis cells had to be present for at
221   least 2 blocks and 80 trials over the session.

222   *Data analysis*

223   For the ANOVA and PSTH analysis, all trials on which monkeys released the lever in the correct
224   interval were analyzed (71.14% of all trials). Trials in which the monkey answered incorrectly (28.86% of
225   all trials), were excluded. The average number of correct trials analyzed for the ANOVA and PSTH
226   analyses were 467.05 (AC: 480.11, dlPFC: 471.72, and BLA: 439.81 trials). We performed a 2 x 2 x 5
227   ANOVA (cue x target x sound level) on the activity of single neurons. The choice is given by the
228   interaction in this ANOVA. The dependent variable was the firing rates of individual neurons. Trials in
229   which the monkeys correctly released the lever within 700ms of the target and which were not catch
230   trials (target 0dB), were analyzed. The firing rate of each cell was computed in 300 ms bins advanced in
231   25 ms increments. We separated the analysis into three different segments of time, locked to the time
232   surrounding the individual presentations of the cue, noise, and first target.

233   Next, we created a population post-stimulus time histogram (PSTH) for the firing rates of the
234   individual neurons with respect to cue condition (left/right) and trial condition (match/foil). For this
235   analysis the firing rate of each cell was computed in 1 ms bins and smoothed with a 3 bin moving
236   average. Data are plotted using 25 ms bins, but t-tests, to determine onset latencies, were computed on
237   the 1 ms bins.

238   For the decoding analyses, we separately analyzed correct and error trials. A trial was
239   considered correct if the monkey released the lever after the presentation of the appropriate target
240   within 700 ms. All other trials were deemed incorrect. The average number of error trials analyzed for
241   decoding was 123.53 (AC: 127.57, dlPFC: 136.99, and BLA: 94.77). For neural analysis, the firing rate of
242   each cell was computed in 100 ms bins and advanced in 25 ms increments. Decoding analyses were
243   performed using leave-one-out cross-validation to predict which observations belong to each trial
244   condition using the SVM classifier in Matlab. All decoding was done using pseudo-populations
245   composed of all neurons recorded from a structure across all sessions. Trials were assigned randomly
246   from the different sessions within each condition.

247   For the ANOVA analyses, we used 300 ms bins, as this provided additionally sensitivity to detect
248   significant effects in neurons with low firing rates. We followed this up with the population analysis
249   which used 1 ms bins, to optimize detection of onset latencies. Finally, we used 100 ms bins for the
250   decoding analysis because the large number of neurons used in this analysis increases the signal-to-
251   noise ratio for detecting effects, and therefore a smaller bin than was used for the ANOVA analysis
252   allows us to detect timing effects more accurately.

253   For the decoding analyses, we calculated significant differences between correct and error trials
254   using a bootstrap analysis (Efron and Tibshirani, 1998). We generated data according the null
255   hypothesis that there were no differences between correct and error trials. We did this by sampling
256   with replacement, from the combined set of correct and error trials, sets of bootstrap correct and error
257   trials. Both the null correct and error bootstrap sets contained combinations of correct and error trials.

258 We then carried out the decoding analysis using the bootstrap trials to determine the decoding accuracy
259 when correct and error trials were mixed. We did this 1000 times. We calculated the difference in
260 fraction correct between correct and error trials in each time bin, for each set of bootstrap trials. This
261 gave us 1000 differences sampled from the null distribution, between correct and error trials, in each
262 time bin. We then compared the difference in the actual data to the differences in the null distribution,
263 and computed a p-value, which was the relative rank of the true difference in the null distribution
264 samples. That is to say, if the true difference was larger than, for example, 986 samples in the null
265 distribution, it was significant with a two sided p-value of 2x(1000-986)/1000 = 0.028.

266 **Results**

267 *Task and Behavior*

268 We recorded neural activity from 2 monkeys while they carried out a spatial selective listening
269 task (Fig. 1A). At the start of each trial, the monkeys acquired central fixation (Fig. 1A), and pressed a
270 bar. After a baseline hold period, an auditory stimulus (the cue) was presented from a speaker on the
271 left or right of the monkey. After the cue there was a delay period during which white noise was played,
272 continuing until bar release. Following the delay period, a second target stimulus was presented on the
273 same (match) or opposite (foil) side as the cue, at different sound levels (Fig. 2). The monkeys were
274 trained to release the bar if the cue and target stimulus were on the same side (match trials) and
275 continue to hold if they were not on the same side (foil trials). In foil trials, following a second delay
276 after the target stimulus, a third match target was played that was always on the same side. In match
277 trials the mean response time was 374.9 ms (std = 27 ms). Monkey 1 had a slightly faster mean response
278 time of 358.3 ms (std 11.5 ms) and Monkey 2 had a mean response time of 405.1 ms (std 19.9 ms).

279 *Single cell encoding of task factors*

280 While the animals carried out the task, neural activity was recorded (Fig. 1B), from three areas:
281 auditory cortex (AC, N = 847), dorsal lateral prefrontal cortex (dlPFC, N = 968) and the basal lateral
282 amygdala (BLA, N = 572). We found neurons in all structures that responded to the presented cues (Fig.
283 3A, 3C, 3E) and the targets, or the interaction of cue and target (Fig. 3B, 3D, 3F). We assessed the
284 encoding of each task factor in single neurons across the population by carrying out ANOVA analyses on
285 correct trials, for each single neuron. With the ANOVA we examined the effects of cue location, target
286 location, target sound level, and interactions (cue x target codes decision), using spike counts in a 300
287 ms window, advanced by 25 ms (Fig. 4). During the cue period, we found that activity discriminated
288 cues rapidly in AC (Fig. 4A). In dlPFC, activity discriminated cues as well, but the effect increased slowly
289 (Fig 4E). The BLA, however, showed minimal cue discriminative activity, with the number of neurons
290 coding cue location only slightly above chance (Fig. 4I). Note that cue location trials were blocked in the
291 task, which led to small baseline, statistically significant, elevation of cue side encoding prior to cue
292 presentation. Although the cue side was blocked, performance on the first trial after the cue switched
293 sides was 76.13% and therefore the animals were attending to the cue. Although encoding peaked in
294 AC and dlPFC following the cue, elevated cue discrimination was maintained during the delay interval,

295   which was not affected by the white noise, in both AC and dlPFC. The BLA showed less delay period
296   activity.

297   When the target stimulus was presented, it was rapidly and robustly encoded in AC (Fig. 4C).
298   The dlPFC also encoded the target stimulus (Fig. 4G), although later than AC, which would be expected.
299   The BLA only weakly encoded the target stimulus and only at about the time of the choice (Fig. 4K). The
300   cue x target interaction, which defined the choice in correct trials, was encoded first in AC (Fig. 4C), after
301   which it was encoded in dlPFC (Fig. 4G). The cue x target interaction, unlike the cue and target
302   locations, was robustly encoded in the BLA (Fig. 4K). Sound level was also robustly encoded in AC (Fig.
303   4C) and less robustly in dlPFC (Fig. 4G) and BLA (Fig. 4K).

304   We also followed up this ANOVA with an additional ANOVA analysis that included both correct
305   and error trials. This allowed us to dissociate the choice from the sensory processing reflecting the cue x
306   target interaction. When we carried out this analysis, we found that the choice was more robustly
307   encoded than the cue x target interaction across all areas (Fig. 4D, 4H, 4L) and most of the cue x target
308   interaction could be accounted for with the choice variable. Overall, all variables, including the delay
309   period activity and the choice, were encoded first and most robustly by AC. The dlPFC did encode all
310   task factors, but after AC. The BLA showed only weak encoding of the cue and the target but robustly
311   encoded the choice.

312   In the next analysis, we compared encoding of the choice in fast and slow reaction time
313   trials, to see if encoding of the choice (i.e. the interaction between cue side and target side in
314   the ANOVA) differed (Fig 5). We performed a median split using the reaction times for all trials,
315   both match and non-match, within a session. For non-match trials we used the release time
316   after the second target as the RT. ANOVAs were run on each neuron twice, once on trials
317   below the median reaction time, and once on trials above the median reaction time. We found,
318   in all three areas, that the choice was encoded faster when the animals responded quickly than
319   when the animals responded slowly. Only in auditory cortex did the activity related to the
320   choice diverge before the average of the fast reaction times (Fig 5A). In both dlPFC and the BLA
321   the activity diverged just before or after the average fast reaction time. Thus, the choice
322   variable from the ANOVA depends on the timing of the motor response and is not completely
323   determined by the timing of the auditory cues.

324   The results from the ANOVA analyses show the contribution of the neurons to each task
325   factor. However, they do not illustrate whether single neurons code multiple task factors
326   through time. Therefore, we also examined whether single neurons encoded more than one
327   task factor, during each epoch (Fig. 6). It could be seen that many neurons coded more than
328   one task factor and coded cue, for example, in both the cue and delay periods.

329   To further quantify whether neurons encoded more than one variable, we also estimated the
330   fraction of neurons that encoded multiple factors using a single representative bin for each factor,
331   centered on the time at which the population encoding of each factor peaked (Table 1). Most often,
332   neurons that encoded the cue during the cue presentation continued to encode the cue during the
333   delay interval. In AC, of the neurons that encoded the cue during cue presentation, 26.70% of them also

334   encoded the cue during the delay period. Neurons in dlPFC were most selective to encoding the cue
335   during both the cue presentation and through the delay interval, with an overlap of 30.90%. In the BLA,
336   25.00% of the neurons encoded the cue during both time periods. Neurons that encoded the cue also
337   often eventually encoded the target, with an overlap of 17.05%, 10.11%, and 8.33% in AC, dlPFC, and
338   the BLA respectively. Most interestingly, while a relatively small portion of neurons encoded the cue
339   through the delay period as well as eventually encoding the target in dlPFC and the BLA, AC did this with
340   an overlap of 16.9%. Neurons in AC were most likely to continue to encode other task variables, in
341   comparison to the dlPFC and the BLA.

342   Next we examined finer time-scale encoding of several task factors. The ANOVA analysis used
343   relatively large time windows to calculate sensitive statistics on potentially low firing rate neurons.
344   These time windows, however, do not allow determination of precise onset times for task factors. To
345   characterize onset times at a finer time scale, we calculated PSTHs using 1 ms time windows, smoothed
346   with a 3-point moving average, for each neuron (Fig. 7 – plotted using 25 ms bins). We then carried out
347   t-tests (p < 0.01, uncorrected) in each bin to estimate the time at which the population in each area
348   discriminated between conditions. We found that the cue was discriminated in AC at 25 ms (Fig. 7B)
349   and in dlPFC at 65 ms (Fig. 7D) after stimulus onset. Using these small bins, the population of BLA
350   neurons did not discriminate cue side, likely due to low firing rates (Fig. 7F). The target was
351   discriminated in AC at 36 ms (Fig. 7G), in dlPFC at 169ms (Fig. 7I) and in the BLA at 185ms (Fig. 7K) after
352   tone onset. Finally, the decision was discriminated in AC at 146 ms (Fig. 7H), in dlPFC at 321 ms (Fig. 7J)
353   and in BLA at 266 ms (Fig. 7L) after target onset.

354   Next, we used a bootstrap analysis to determine whether onset latencies differed significantly
355   between areas (Fig 7). We pulled samples of 100 neurons for each brain area and computed the time at
356   which the two conditions diverged (p < 0.05, consecutive bins >= 6) in each bootstrap sample. We did
357   this 100 times to create a sample distribution. We then calculated a 95% confidence interval for the
358   discrimination times for each area. If the confidence intervals overlapped, the brain areas were not
359   deemed statistically significant. From this analysis, AC preceded both dlPFC and BLA in cue and target
360   discrimination, and AC preceded dlPFC in decision discrimination. AC, however, did not statistically
361   precede the BLA in decision discrimination.

362   *Decoding correct and error trials*

363   In the next analyses we used decoding to examine error trial activity. We were interested in
364   which processes broke down in error trials. To examine this, we used leave-one-out cross validation on
365   pseudo populations (see methods) to predict, using the neural activity, the side on which the cue was
366   presented (Fig. 8), the side on which the target was presented (Fig. 9) and the choice (Fig. 10). The
367   decoding model was first estimated using only correct trials. We then classified the error trials using the
368   decoding model estimated on correct trials, to see if neural activity in error trials represented the stimuli
369   that were presented, and the choice that was made. We found that in correct and error trials the neural
370   population in both AC and dlPFC rapidly predicted the cue location (Fig. 8A, D), and maintained
371   prediction through the delay interval (Fig. 8B, D), consistent with the single-neuron results. The BLA did
372   not discriminate clearly the cue side (Fig. 8G). There were no significant differences between correct

373   and error trials for cue encoding, and this finding was consistent through the delay interval. Therefore,
374   the cue was correctly encoded in error trials.

375   When we decoded the target side using neural activity, we found that in correct trials the target
376   location was robustly predicted by AC (Fig. 9C) and dlPFC (Fig. 9F). There was minimal prediction of the
377   target in the BLA (Fig. 9I). In error trials, however, the target was not well predicted by any of the areas
378   (Fig. 9). The correct and error trial predictions diverged (p < 0.01 bootstrap) 75 ms after target onset in
379   AC and 125 ms after target onset in dlPFC.

380   In error trials, animals either released when they should not have, or did not release when they
381   should have. When we predicted the choice, relative to what the monkeys should have done, we found
382   an accurate prediction in correct trials in all 3 areas (Fig. 10C, 10F, 10I). Furthermore, in error trials, the
383   predicted choice tended to fall below chance, which indicates that the neural activity is coding the
384   choice the monkey made in error trials, as opposed to the choice the monkey should have made.
385   However, this coding was only significantly below chance late in the choice period in AC (Fig. 10C). We
386   used a smaller time bin of 5ms in the rightmost column (Fig 10C,F,I) to more precisely determine the
387   point at which the curves diverged. Consistent with the other analyses, we found that predictions in
388   error and correct trials diverged statistically in auditory cortex (270 ms after target onset) and
389   subsequently in dlPFC and BLA (275 and 300 ms after target onset).

390   Next, we examined the position of the population neural activity relative to the discrimination
391   boundary, extracted from the decoding model. For the decoding analysis (Fig. 8-10), this quantity is
392   thresholded in each trial and time-bin, and the time-bin in that trial is classified as either, e.g. cue left or
393   cue right, depending on whether the position is positive or negative. However, the average distance to
394   the decoding boundary provides a continuous estimate of how well the population discriminated the
395   conditions vs. time (Fig. 11). In general, these analyses were consistent with the thresholded decoding
396   analysis. Cue related activity diverged in correct and error trials, reflecting the cued side, and the
397   activity in error trials matched the activity in correct trials (Fig. 11A, D, G). The breakdown in activity
398   following target presentation could also be seen (Fig. 11B, E, H). However, there was some maintained
399   coding of the target, particularly in auditory cortex (Fig. 11B), which may also be reflected in the
400   decoding accuracy in error trials (Fig. 9C). Therefore, cue encoding is intact in error trials and target
401   encoding is mostly but not completely absent. The choice encoding dynamics did reflect the fact that
402   the wrong choice tended to be predicted by population activity (Fig. 11C, F, I). However, it could be
403   seen that the activity diverged less than it did in correct trials, consistent with the lower decoding
404   performance.

405   *Neural responses in the passive task*

406   In a final series of analyses, we analyzed data from a passive task, collected in each session
407   before the main, active task data. The sensory stimulation in the passive task was identical to the
408   stimulation in the active task, except the animals did not press a bar to initiate a trial, they did not
409   release the bar to indicate their choice, and there was no juice tube so they could not be rewarded.
410   When we examined encoding of cue location, we again found robust coding in AC (Fig. 12A). All of the

411 other signals, however, were much weaker. The cue responses in dlPFC dropped from a peak near 30%
412 in the active task to about 10% in the passive task (Fig. 12D). Interestingly, there was delay activity in
413 the passive task, in AC (Fig. 12B), perhaps because the animals were highly over-trained. The delay
414 activity in dlPFC was reduced from about 20% of the population to about 10% (Fig. 12E). There was also
415 a small amount of target encoding in AC (Fig. 12C). Target encoding in dlPFC did not exceed chance (Fig.
416 12F). Encoding in the BLA only sporadically exceeded chance, perhaps due to type-I errors, or low-level
417 encoding (Fig. 12G-I).

418   We also examined onset times, using small time-bins (Fig. 13). We found differences in
419 responses in AC that depended on the side of the stimulus for the cue at 32 ms (Fig. 13B) and for the
420 target at 53 ms (Fig. 13G). However, we did not detect population level differences in responses, using
421 these small time bins, in dlPFC or BLA, which suggests that responses that reached significance in the
422 ANOVA analyses were driven by low firing rates. Overall, beyond cue encoding in AC, responses across
423 all 3 areas were reduced in the passive task, relative to the active task.

424 Discussion

425   We trained monkeys on a selective listening task, based on tasks used in humans. The task
426 required animals to detect a difficult to discriminate auditory stimulus, embedded in white noise. We
427 found that AC encoded cues, targets, and decisions, prior to either dlPFC or the BLA. In addition, AC had
428 delay activity that coded the location of the initial cue. It is not clear, however, whether the AC delay
429 activity depended on dlPFC delay activity, or even parietal activity that we did not record. Activity in
430 dlPFC closely followed activity in AC. The BLA, on the other hand, only minimally encoded cue and
431 target activity. The BLA was strongly engaged, however, at the time of choice, although the choice
432 related activity followed activity in AC. Therefore, the AC appears to support many of the functions
433 required for auditory selective listening. This is in contrast to early visual areas, which represent visual
434 features, but play a minimal role in decision making aspects of tasks (Britten et al., 1992; Zaksas and
435 Pasternak, 2006).

436   Previous work has shown that AC neurons can encode non-sensory, choice-related activity
437 (Niwa et al., 2012; Christison-Lagay and Cohen, 2018; Huang et al., 2019). The Huang et al. study found
438 that whether a choice was predictable following a cue tone, based on the task condition, affected neural
439 responses in AC to the tone. Therefore, AC encoded whether the response was determined by the first
440 cue. Our results are consistent with this and other studies (Christison-Lagay and Cohen, 2018), in that
441 we show that auditory cortex encodes the necessary response. However, in our task, the choice was not
442 determined by the first cue, so choice related activity only followed the target. Our paradigm does not
443 allow us to dissociate decision making from the motor response required to indicate the decision and
444 therefore our choice coding could be related to either, though note that it begins well before monkeys
445 man reaction of time of ~ 400 ms. We also show, that encoding in AC precedes encoding in dlPFC, and
446 we dissociated through our fully crossed experimental design, encoding of cue location, target location,
447 and the required response. Although it is possible that AC inherits response encoding from a cortical
448 area other than dlPFC, the anatomical organization of this system suggests it would have to be a nearby
449 area, for example belt or parabelt auditory cortex (Romanski and Averbeck, 2009; Kajikawa et al., 2015;

450   Tsunada et al., 2016). Given that AC is deeper into the neural processing stream than, for example,
451   primary visual cortex (Mizrahi et al., 2014), it is also possible that AC could have sufficiently
452   sophisticated mechanisms to compute the required response locally. Though AC precedes PFC in the
453   encoding of the decision in both correct and error trials, the responses across areas are also quite similar
454   within this task (~50 ms differences). This tight temporal relationship between AC and dlPFC is
455   contextually dependent. When responses during the passive condition were analyzed, the fraction of
456   responsive neurons was reduced and responses were later in all areas relative to the task-related
457   responses (and BLA was completely unresponsive, consistent with a primary role in reward guided
458   behavior). Particularly, dlPFC showed a reduction of responses to the cue and delay activity and an
459   abolishment of target related activity compared to the active task condition. This is consistent with data
460   from the same animals and areas during a passive oddball task in which dlPFC activity was later (~100
461   ms) and weaker than in AC (Camalier et al., 2019). Taken together it suggests that the strength and
462   timing of the information transfer between AC and dlPFC can be flexibly allocated and is dependent on
463   task demands. Lastly, comparison of the active and passive conditions highlights the sustained
464   nonsensory motor/reward related activity in "primary" sensory cortex (AC)(Knyazeva et al., 2020).

465   Several of the analyses show that the neural responses recorded in this task were not
466   straightforward sensory responses to the auditory stimuli. This was true across areas. For example, we
467   found that the cue x target interaction, which defines choices in correct trials, was less strongly encoded
468   than the choices, when both correct and error trials were analyzed. In addition, when we split trials into
469   those with fast and slow reaction times, we found that the neural representation of the decision was
470   coded earlier when choices were made earlier. We also saw that much of the task related neural
471   activity was reduced, although not eliminated, in the passive condition, when animals did not have to
472   respond to the sensory cues.

473   Both prefrontal (Green et al., 2011; Bidet-Caulet et al., 2015) and parietal (Michalka et al., 2016;
474   Deng et al., 2019; Deng et al., 2020) cortex have been shown to play important roles in auditory spatial
475   attention in humans. AC has also been shown to have attention selective modulation of single neurons
476   when targets and distractors are separated by frequency content (Atiani et al., 2009; Schwartz and
477   David, 2018; O'Sullivan et al., 2019). Although we found clear responses related to the cued side in
478   dlPFC, they followed AC. This was true of not only the sensory responses, but also the decision
479   response. From our data it is not, however, possible to determine whether the delay period activity,
480   which may represent sustained attention/working memory for the cue location, was sustained by AC,
481   dlPFC, or their interaction. In addition, several of the spatial attention paradigms used in the human
482   work required participants to attend or discriminate sounds in one location, while ignoring sounds on
483   the contralateral side (Deng et al., 2019). It is possible that if we had required the monkeys to carry out
484   complex perceptual discriminations at one location, while ignoring distractors at another location, we
485   would have found stronger engagement of dlPFC. We did use a white masking noise, following the cue
486   signal, to examine its effects on behavior and neural representations of the cue location. Although we
487   did see some effects of the noise onset in the decoding analysis, effects which were stronger in AC than
488   dlPFC, they were transient and resulted in increased decoding accuracy for the cued location. The
489   increased accuracy may have followed from an overall increase in neural activity, which may have

490   improved decoding performance. Also, we did not record neural activity in parietal cortex, which may
491   also play a role in the sustained delay period activity, although it would be interesting to consider
492   inferior parietal cortex in future studies.

493   We found that the BLA played little role in encoding the cue location, and responses related to
494   the choice followed responses in AC. This is inconsistent with previous reports of the BLA's involvement
495   in visual-spatial attention (Peck et al., 2013). In these tasks, the amygdala neurons encoded the valence
496   of stimuli, that were saccade targets, during delay periods (Peck and Salzman, 2014). There are several
497   differences between these tasks, and ours, however. For example, the tasks used in Peck et al. were
498   based on visual-spatial paradigms instead of an auditory-spatial paradigm, and they also required eye
499   movements to spatial locations. Although the BLA receives auditory inputs(Yukie, 2002), these inputs
500   may play a smaller role in the primate than they do in rodents (Munoz-Lopez et al., 2010). In rodents,
501   auditory cues can be associated with shock in Pavlovian fear conditioning (Romanski and LeDoux, 1992).
502   These studies have shown that the amygdala plays an important role in the associative process between
503   cues and shocks. Although, the amygdala is also involved in reward guided behavior (Costa et al., 2016;
504   Averbeck and Costa, 2017; Costa et al., 2019). We did find a small, although significant, population of
505   amygdala neurons, that encoded the auditory cue and the auditory target. They did so, however, at
506   long latencies. Therefore, the BLA appears to play a minimal role in the cognitive process of selective
507   listening under reward-constant trials in highly trained animals. It is however possible that if we had
508   primarily recorded from the lateral nucleus, which receives most of the direct auditory inputs (Yukie,
509   2002), we would have found more neurons related to aspects of our task.

510   The present study also shows a substantial dissociation of function between the BLA and dlPFC.
511   This dissociation differs from the similarity between these structures seen in reinforcement learning (RL)
512   tasks, in which both dlPFC and the BLA show substantial encoding of the identity of visual stimuli, the
513   reward values associated with those stimuli, and reward outcomes (Bartolo et al., 2019; Costa et al.,
514   2019). The primary difference between the BLA and dlPFC, in RL tasks, is that the dlPFC strongly
515   encodes the direction of eye movements required to saccade to a rewarding visual stimulus (Bartolo et
516   al., 2019), whereas the BLA encodes eye movement directions only at a low level (Costa et al., 2019).
517   Thus, in RL tasks, the BLA and dlPFC show similar responses, which are also similar to those seen in the
518   ventral striatum (Costa et al., 2019) and orbito-frontal cortex (Costa and Averbeck, 2020), with which
519   the BLA is mono-synaptically connected. The current study, however, shows that in cognitive, auditory
520   selective listening tasks, the BLA and dlPFC show different responses, until the animal makes a reward
521   guided choice.

522

523   Conclusions

524   We found that AC encoded cues, targets, and decisions, before dlPFC, in an auditory selective
525   listening task. We also found that AC had delay period activity. The BLA had minimal cue or target
526   activity, although it did encode decision activity. The decision related activity in the BLA, however,
527   followed decision related activity in AC. Overall, this suggests that AC may carry out most important

528   computations relevant to auditory selective listening. The main caveat is that it is not possible to
529   determine whether delay period activity, which likely critically underlies performance in this task, is
530   supported by AC in the absence of dlPFC or parietal cortex. Future work, for example inactivating dlPFC
531   and/or parietal cortex (Plakke et al., 2015), while recording in AC, could clarify this question.

532
533
534
535
536
537

538 **References**
539
540 Ahveninen J, Huang S, Nummenmaa A, Belliveau JW, Hung AY, Jaaskelainen IP, Rauschecker JP,
541   Rossi S, Tiitinen H, Raij T (2013) Evidence for distinct human auditory cortex regions for
542   sound location versus identity processing. Nat Commun 4:2585.
543 Alho K, Medvedev SV, Pakhomov SV, Roudas MS, Tervaniemi M, Reinikainen K, Zeffiro T,
544   Naatanen R (1999) Selective tuning of the left and right auditory cortices during spatially
545   directed attention. Brain Res Cogn Brain Res 7:335-341.
546 Atiani S, Elhilali M, David SV, Fritz JB, Shamma SA (2009) Task difficulty and performance induce
547   diverse adaptive patterns in gain and shape of primary auditory cortical receptive fields.
548   Neuron 61:467-480.
549 Averbeck BB, Costa VD (2017) Motivational neural circuits underlying reinforcement learning.
550   Nature Neuroscience 20:505-512.
551 Bartolo R, Saunders RC, Mitz A, Averbeck BB (2019) Dimensionality, information and learning in
552   prefrontal cortex. bioRxiv:823377.
553 Bidet-Caulet A, Buchanan KG, Viswanath H, Black J, Scabini D, Bonnet-Brilhault F, Knight RT
554   (2015) Impaired Facilitatory Mechanisms of Auditory Attention After Damage of the
555   Lateral Prefrontal Cortex. Cereb Cortex 25:4126-4134.
556 Bigelow J, Rossi B, Poremba A (2014) Neural correlates of short-term memory in primate
557   auditory cortex. Front Neurosci 8:250.
558 Bon L, Lucchetti C (1994) Ear and eye representation in the frontal cortex, area 8b, of the
559   macaque monkey: an electrophysiological study. Exp Brain Res 102:259-271.
560 Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a
561   comparison of neuronal and psychophysical performance. J Neurosci 12:4745-4765.
562 Camalier CR, Scarim K, Mishkin M, Averbeck BB (2019) A Comparison of Auditory Oddball
563   Responses in Dorsolateral Prefrontal Cortex, Basolateral Amygdala, and Auditory Cortex
564   of Macaque. J Cogn Neurosci:1-11.
565 Camalier CR, D'Angelo WR, Sterbing-D'Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural
566   latencies across auditory cortex of macaque support a dorsal stream supramodal timing
567   advantage in primates. Proc Natl Acad Sci U S A 109:18168-18173.
568 Cherry EC (1953) Some Experiments on the Recognition of Speech, with One and with 2 Ears. J
569   Acoust Soc Am 25:975-979.
570 Christison-Lagay KL, Cohen YE (2018) The Contribution of Primary Auditory Cortex to Auditory
571   Categorization in Behaving Monkeys. Front Neurosci 12:601.
572 Costa VD, Averbeck BB (2020) Primate orbitofrontal cortex codes information relevant for
573   managing explore-exploit tradeoffs. J Neurosci.
574 Costa VD, Mitz AR, Averbeck BB (2019) Subcortical Substrates of Explore-Exploit Decisions in
575   Primates. Neuron 103:533-545 e535.
576 Costa VD, Dal Monte O, Lucas DR, Murray EA, Averbeck BB (2016) Amygdala and Ventral
577   Striatum Make Distinct Contributions to Reinforcement Learning. Neuron 92:505-517.
578 Deng Y, Choi I, Shinn-Cunningham B (2020) Topographic specificity of alpha power during
579   auditory spatial attention. Neuroimage 207:116360.
580 Deng Y, Reinhart RM, Choi I, Shinn-Cunningham BG (2019) Causal links between parietal alpha
581   activity and spatial auditory attention. Elife 8.

582 Efron B, Tibshirani RJ (1998) An introduction to the bootstrap. New York: Chapman & Hall.

583 Frey JN, Mainy N, Lachaux JP, Muller N, Bertrand O, Weisz N (2014) Selective modulation of
584   auditory cortical alpha activity in an audiovisual spatial attention task. J Neurosci
585   34:6634-6639.

586 Green JJ, Doesburg SM, Ward LM, McDonald JJ (2011) Electrical neuroimaging of voluntary
587   audiospatial attention: evidence for a supramodal attention control network. J Neurosci
588   31:3560-3564.

589 Hackett TA, Stepniewska I, Kaas JH (1999) Prefrontal connections of the parabelt auditory
590   cortex in macaque monkeys. Brain Res 817:45-58.

591 Huang Y, Heil P, Brosch M (2019) Associations between sounds and actions in early auditory
592   cortex of nonhuman primates. Elife 8.

593 Kajikawa Y, Frey S, Ross D, Falchier A, Hackett TA, Schroeder CE (2015) Auditory properties in
594   the parabelt regions of the superior temporal gyrus in the awake macaque monkey: an
595   initial survey. J Neurosci 35:4140-4150.

596 Knyazeva S, Selezneva E, Gorkin A, Ohl FW, Brosch M (2020) Representation of Auditory Task
597   Components and of Their Relationships in Primate Auditory Cortex. Frontiers in
598   Neuroscience 14.

599 Lanzilotto M, Perciavalle V, Lucchetti C (2013) Auditory and visual systems organization in
600   Brodmann Area 8 for gaze-shift control: where we do not see, we can hear. Front Behav
601   Neurosci 7:198.

602 Lee JH, Russ BE, Orr LE, Cohen YE (2009) Prefrontal activity predicts monkeys' decisions during
603   an auditory category task. Front Integr Neurosci 3:16.

604 Mayer AR, Harrington DL, Stephen J, Adair JC, Lee RR (2007) An event-related fMRI Study of
605   exogenous facilitation and inhibition of return in the auditory modality. J Cogn Neurosci
606   19:455-467.

607 Mayer AR, Mannell MV, Ling J, Elgie R, Gasparovic C, Phillips JP, Doezema D, Yeo RA (2009)
608   Auditory orienting and inhibition of return in mild traumatic brain injury: a FMRI study.
609   Hum Brain Mapp 30:4152-4166.

610 McDonald JJ, Ward LM (1999) Spatial relevance determines facilitatory and inhibitory effects of
611   auditory covert spatial orienting. J Exp Psychol Human 25:1234-1252.

612 Michalka SW, Rosen ML, Kong L, Shinn-Cunningham BG, Somers DC (2016) Auditory Spatial
613   Coding Flexibly Recruits Anterior, but Not Posterior, Visuotopic Parietal Cortex. Cereb
614   Cortex 26:1302-1308.

615 Mizrahi A, Shalev A, Nelken I (2014) Single neuron and population coding of natural sounds in
616   auditory cortex. Curr Opin Neurobiol 24:103-110.

617 Munoz-Lopez MM, Mohedano-Moriano A, Insausti R (2010) Anatomical pathways for auditory
618   memory in primates. Front Neuroanat 4:129.

619 Naatanen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I (2001) "Primitive intelligence"
620   in the auditory cortex. Trends Neurosci 24:283-288.

621 Nelken I (2004) Processing of complex stimuli and natural scenes in the auditory cortex. Curr
622   Opin Neurobiol 14:474-480.

623 Niwa M, Johnson JS, O'Connor KN, Sutter ML (2012) Activity related to perceptual judgment
624   and action in primary auditory cortex. J Neurosci 32:3193-3210.

625   O'Sullivan J, Herrero J, Smith E, Schevon C, McKhann GM, Sheth SA, Mehta AD, Mesgarani N
626   (2019) Hierarchical Encoding of Attended Auditory Objects in Multi-talker Speech
627   Perception. Neuron 104:1195-1209 e1193.
628   Peck CJ, Salzman CD (2014) Amygdala neural activity reflects spatial attention towards stimuli
629   promising reward or threatening punishment. eLife 3.
630   Peck CJ, Lau B, Salzman CD (2013) The primate amygdala combines information about space
631   and value. Nature Neuroscience 16:340-348.
632   Plakke B, Hwang J, Romanski LM (2015) Inactivation of Primate Prefrontal Cortex Impairs
633   Auditory and Audiovisual Working Memory. J Neurosci 35:9666-9675.
634   Rinne T, Muers RS, Salo E, Slater H, Petkov CI (2017) Functional Imaging of Audio-Visual
635   Selective Attention in Monkeys and Humans: How do Lapses in Monkey Performance
636   Affect Cross-Species Correspondences? Cereb Cortex 27:3471-3484.
637   Roberts KL, Summerfield AQ, Hall DA (2009) Covert auditory spatial orienting: an evaluation of
638   the spatial relevance hypothesis. J Exp Psychol Hum Percept Perform 35:1178-1191.
639   Romanski LM, LeDoux JE (1992) Equipotentiality of thalamo-amygdala and thalamo-cortico-
640   amygdala circuits in auditory fear conditioning. J Neurosci 12:4501-4509.
641   Romanski LM, Averbeck BB (2009) The primate cortical auditory system and neural
642   representation of conspecific vocalizations. Annu Rev Neurosci 32:315-346.
643   Schmolesky MT, Wang Y, Hanes DP, Thompson KG, Leutgeb S, Schall JD, Leventhal AG (1998)
644   Signal timing across the macaque visual system. J Neurophysiol 79:3272-3278.
645   Schwartz ZP, David SV (2018) Focal Suppression of Distractor Sounds by Selective Attention in
646   Auditory Cortex. Cereb Cortex 28:323-339.
647   Scott BH, Mishkin M (2016) Auditory short-term memory in the primate auditory cortex. Brain
648   Res 1640:264-277.
649   Spence CJ, Driver J (1994) Covert Spatial Orienting in Audition - Exogenous and Endogenous
650   Mechanisms. J Exp Psychol Human 20:555-574.
651   Stefanacci L, Amaral DG (2002) Some observations on cortical inputs to the macaque monkey
652   amygdala: an anterograde tracing study. J Comp Neurol 451:301-323.
653   Teshiba TM, Ling J, Ruhl DA, Bedrick BS, Pena A, Mayer AR (2013) Evoked and Intrinsic
654   Asymmetries during Auditory Attention: Implications for the Contralateral and Neglect
655   Models of Functioning. Cerebral Cortex 23:560-569.
656   Tsunada J, Liu AS, Gold JI, Cohen YE (2016) Causal contribution of primate auditory cortex to
657   auditory perceptual decision-making. Nat Neurosci 19:135-142.
658   Yin P, Strait DL, Radtke-Schuller S, Fritz JB, Shamma SA (2020) Dynamics and Hierarchical
659   Encoding of Non-compact Acoustic Categories in Auditory and Frontal Cortex. Curr Biol
660   30:1649-1663 e1645.
661   Yukie M (2002) Connections between the amygdala and auditory cortical areas in the macaque
662   monkey. Neurosci Res 42:219-229.
663   Zaksas D, Pasternak T (2006) Directional signals in the prefrontal cortex and in area MT during a
664   working memory for visual motion task. J Neurosci 26:11726-11742.

665
666

667   **Figure 1.** Task design and recording locations. (A) Structure of the spatial selective listening task. Cue
668   conditions (listen left/right) were blocked with two types of trials (match/foil) in each condition. To
669   begin each trial, the animal must depress a lever and maintain fixation at a central point on the screen.
670   After a short delay (2.1-2.4 s), the animal heard a 4 kHz square wave cue from a speaker on the left or
671   right of its head. A continuous white noise was played 500ms after the initial cue to make target
672   detection difficult. In match trials, the animal heard a 1 kHz match target (various levels, see Methods)
673   after some stimulus onset asynchrony (SOA) (500, 800 or 1300ms) on the same side as the cue. If the
674   animal released the lever within 700ms of the match target, a fixed juice reward was delivered. In foil
675   trials, the animal heard a 1 kHz foil target on the opposite side as the cue after the same SOA. The
676   animal had to continue to hold down the lever until a 1 kHz match target was presented (after 800 or
677   1000ms) on the same side as the cue. If the animal released the bar within 700ms of the match target, a
678   fixed juice reward was delivered. The "passive listening" control condition was identical to the active
679   task except the monkey listened passively and did not press a lever, fixate, respond, or receive juice. (B)
680   Recording locations of single neurons across auditory cortex (AC), dorsolateral prefrontal cortex (dlPFC)
681   and the basal lateral amygdala (BLA). (Top) Patch of dlPFC recording area morphed to anatomical
682   landmarks. (Middle) AC grid coverage on region A1 of auditory cortex based on topography, latency and
683   frequency reversals. (Bottom) Region of interest highlighted in blue--the entire left basolateral
684   amygdala--targeted by v-trodes. In all three areas, we selectively recorded from the left hemisphere of
685   the animal. The lower right image shows a structural MRI with contrast agent (betadine gel) in chamber
686   grid holes for targeting. Recording locations and trajectories were further verified using tungsten
687   electrodes inserted through grid locations to target areas. Yellow lines in each image show an
688   approximate trajectory.
689
690   **Figure 2.** Auditory target sound level and accuracy. Trial performance compared against the target
691   sound level in decibels. Numbers below line indicate percentage of trials across sessions at that sound
692   level. Note, catch trials (0 dB) are not plotted, so percentages do not add to 1. Bars at each point
693   represent the standard error. Mean values were first calculated for each session, and then means were
694   taken across sessions in which the indicated sound level was used. The standard error of the mean was
695   calculated across sessions where the number of sessions are: Monkey 1: 16-20 dB (N = 53); 21-25 dB (N
696   = 19); 26-30 dB (N = 53), 36-40 dB (N = 1 – data not shown); Monkey 2: 26-30 dB (N = 42); 31-35 dB (N =
697   42); 36-40 dB (N = 42); 41-45 dB (N = 9).
698
699   **Figure 3.** Example neurons. Left hand panel shows rasters of single trials, right hand panel show p-value
700   from ANOVA, for the indicated factor. Only correct trials are shown. X-axis for p-value plots shows right
701   hand edge of 300 ms bin used for ANOVA. A. Example neuron from cue epoch in AC. B. Example AC
702   neuron showing responses to target. Plotted p-values are for target factor. C. Example dlPFC neuron
703   showing responses to cue. D. Example dlPFC neuron showing responses to target. Plotted p-values are
704   for cue x target interaction. E. Example BLA neuron showing responses to cue. F. Example BLA neuron
705   showing response to cue x target interaction.
706
707   **Figure 4.** ANOVA analysis. Recording of single neurons from caudal AC (A1, lateral belt), dlPFC and BLA
708   while monkeys are performing a spatial selective listening task. A 2 x 2 factor ANOVA (Cue side x Target
709   side, $p < 0.05$) using 300ms bins sliding at 25ms. Bin endpoint was used to align time on the x-axis, i.e.
710   300 ms is a bin from 0 to 300 ms. Bars above each plot represent the bins in which a statistically
711   significant fraction of neurons encode each factor by color ($p < 0.01$; binomial test). (A, E, I) During
712   presentation of the cue, neurons respond differentially to the cue location. (B, F, J) Post-cue, a
713   substantial fraction of neurons is selective to cue side, during the delay period, in both AC and dlPFC. (C,
714   G, K) Post-target presentation, a substantial portion of neurons in all three areas of interest are selective

19

715   to the choice. (D, H, L) Post-target presentation analysis including error trials shows choice encoding
716   over and above cue x target interaction. Thus, choices are not a direct reflection of sensory input.
717
718   **Figure 5.** Median reaction time split ANOVA analysis. Recording of single neurons from AC, dlPFC and
719   BLA while monkeys were performing the task. The results are from a 2 x 2 ANOVA (Cue side x Target
720   side, $p < 0.05$) using 300 ms bins sliding at 25 ms. Only the interaction term (Response) is plotted on the
721   graph. Bin endpoint was used to align time on the x-axis. The trials for all neurons were split into fast
722   and slow reaction time by the median reaction time within a session and separate ANOVAs were run for
723   each set of trials. Paired t-tests ($p < 0.01$, consecutive bins >= 3) were computed to determine
724   significance between fraction of significant neurons assessed in each reaction time split. Bars above
725   each plot represent the bins in which a statistically significant difference was seen between response to
726   trials with fast reaction times vs. slow reaction times.
727
728   **Figure 6.** Contribution of individual neuron selectivity to the populational representation. 2 x 2 ANOVA
729   ($p < 0.05$) using 300 ms bins sliding at 25 ms conducted on each individual neuron. Neurons are plotted
730   along the y-axis and the time is on the x-axis. Bin endpoint was used to align time on the x-axis. Grey
731   bars represent the times in which a neuron was significant for that task factor. Blue column displays
732   encoding of the cue during the cue presentation, grey column shows the cue encoding during the delay
733   period, red column shows the target encoding during target presentation and yellow column shows the
734   response encoding during the target period.
735
736   **Figure 7.** Post-stimulus time histograms (PSTHs). Mean normalized firing rates of neurons plotted using
737   non-overlapping 25ms bins smoothed with a 3-point moving average. Only neurons significant for the
738   corresponding factor (i.e. cue, target or cue x target) were included in this analysis. Bin midpoint was
739   used to align time on the x-axis. Analysis was conducted to assess precise timing of changes in neuronal
740   firing rates in AC, dlPFC and BLA. Paired t-tests were performed on all bins to determine significant
741   difference in firing rates. Bootstrapping analysis was performed to directly compare timing differences
742   in different brain areas. Vertical bars with stars indicate non-overlapping 95% confidence intervals for
743   discrimination times between areas. (A, C, E) compares conditions that are identical in cue side but vary
744   in target side, as a measure of sensory identification. (B, D, F) compares conditions that are identical in
745   target side but vary in the cue location (left or right). (G, I, K) Conditions are matched for cue side but
746   vary in target side. (H, J, L) Conditions shown have opposite cue sides but matched target side.
747
748   **Figure 8.** Classification analysis to cue location factor comparing correct and error trials. Analysis was
749   performed using 100 ms bins, sliding at 25 ms. Bin endpoint was used to align time on the x-axis.
750   Analysis performed using leave-one-out cross-validation to predict which observations belong to each
751   cue condition. Bootstrap test performed with 1000 pseudorandom samples. No significant difference in
752   classification rates was found between correct and error trials in any brain region during any time bin.
753
754   **Figure 9.** Classification analysis to target factor by correct or error. Analysis was performed using 100 ms
755   bins, sliding at 25 ms. Bin endpoint was used to align time on the x-axis. Grey shaded areas represent
756   timepoints where correct and error classification rates differ ($p < 0.01$, bootstrap). (C) In AC, a significant
757   difference is seen in classification rates during the target epoch that begins after 75ms and ends after
758   400ms. (F) In dlPFC, the difference in classification rates begins slightly later and ends slightly earlier,
759   starting at 125ms post-target and ending at 375ms post-target.
760
761   **Figure 10**. Classification analysis to choice. Grey bar indicates timepoints where correct and error
762   classification rates differ, red bar indicates timepoints where error trials were significantly below chance
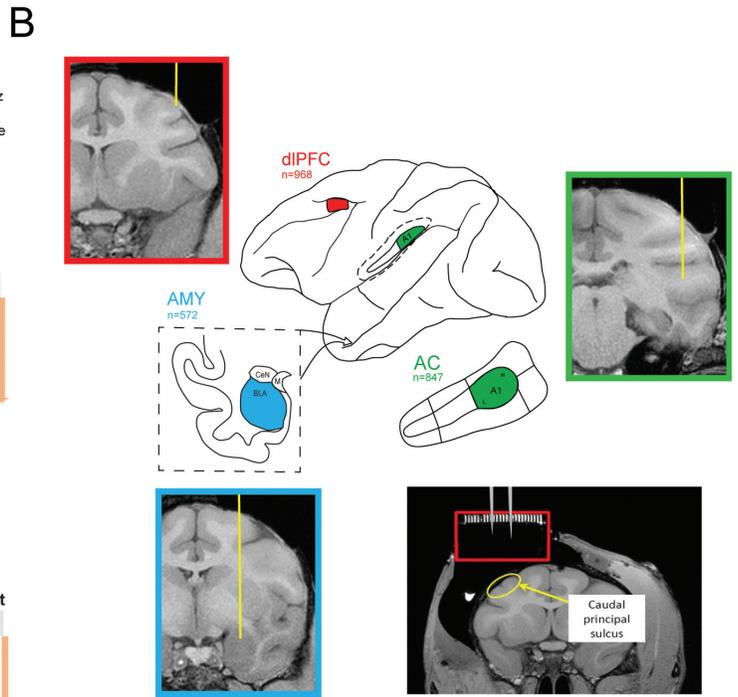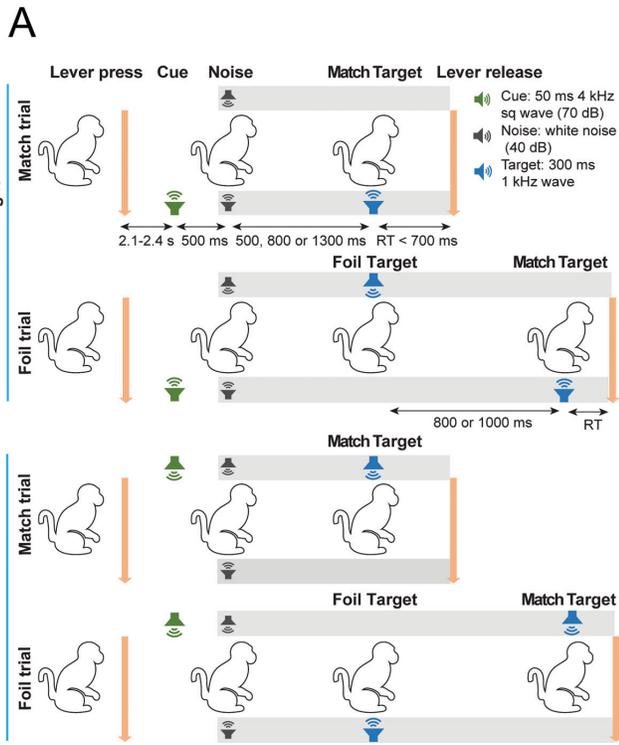
763 (0.5). Bin endpoint was used to align time on the x-axis. The reaction time for detect trials is shown as a
764 dotted line, with the standard deviation shaded, in panels C, F and I. (C) Differences between correct
765 and error trials in AC were from (270ms to 500ms), (F) In dlPFC from (275ms to 500ms) and (I) In BLA
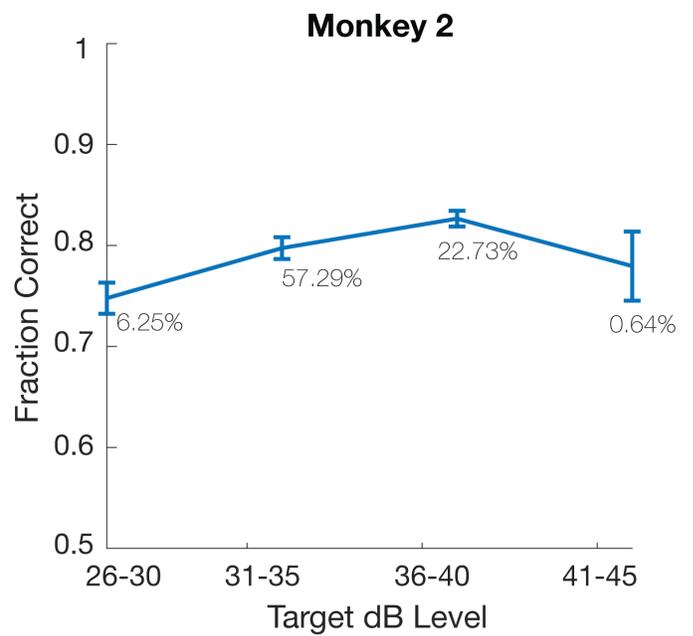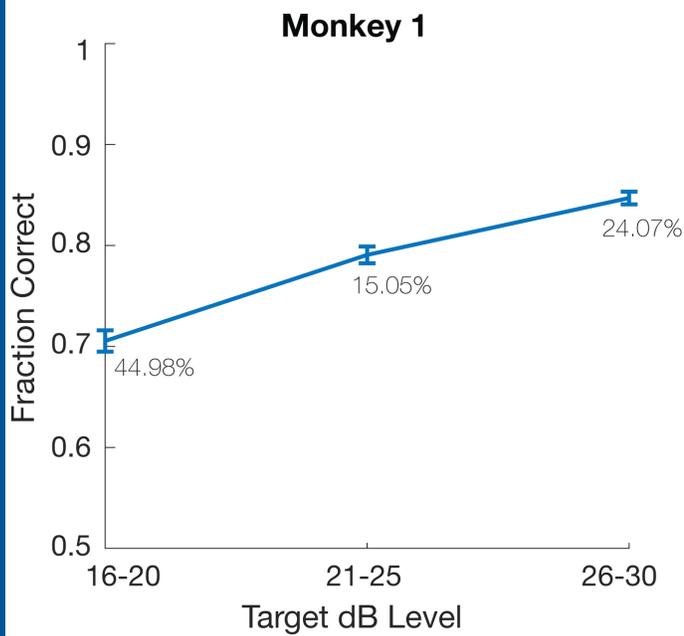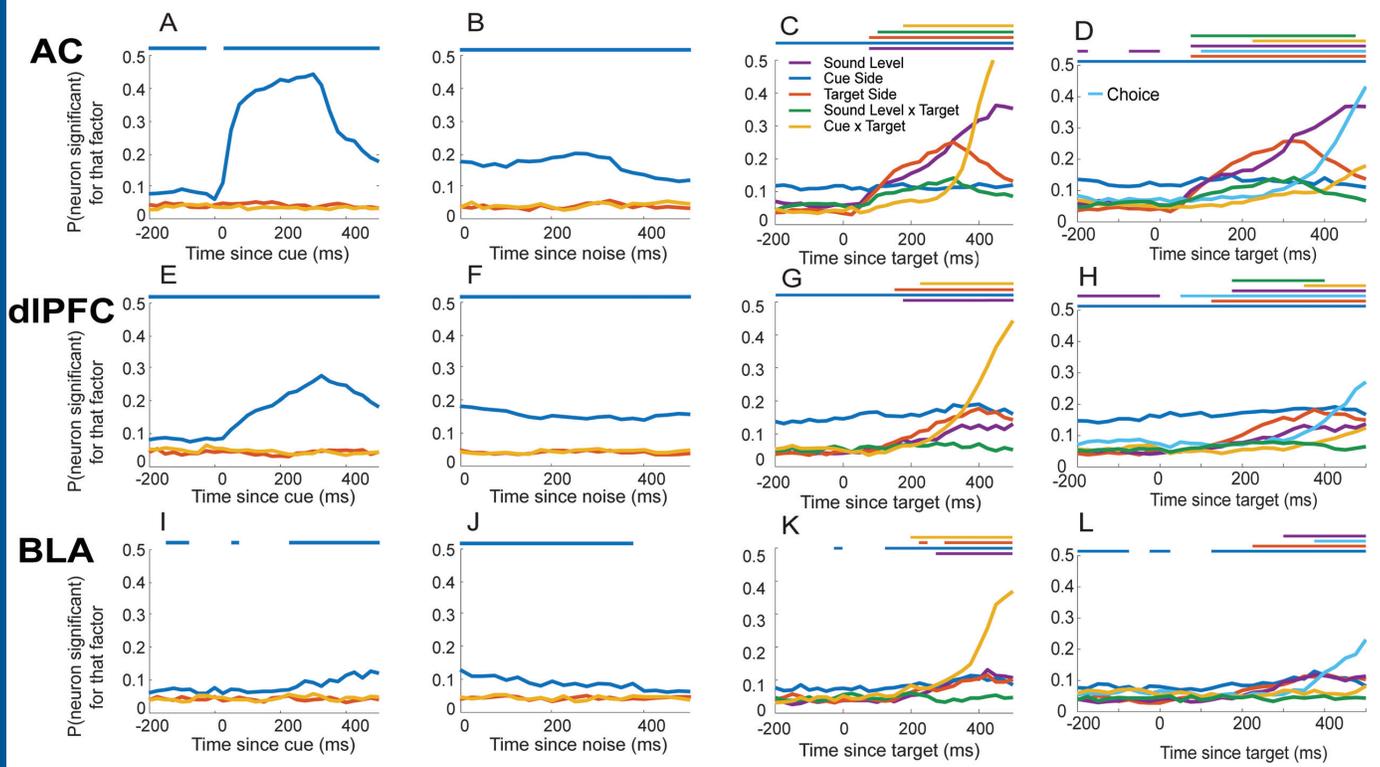766 from (300ms to 500ms).
767
768 **Figure 11**. Distance to classification boundary derived from support vector machine classifier. Analyses
769 were conducted using 200ms bins sliding at 25ms. Bin endpoint was used to align to the x-axis. Error
770 trials were defined as trials in which the animal responded incorrectly or responded outside of the
771 allotted reaction time window. Correct trials are presented in shades of blue and error trials in shades of
772 red. The reaction time for detect trials is shown as a dotted line, with the standard deviation shaded, in
773 panels C, F and I. (A, D, G) Conditions were separated by whether a trial was cued on the left or right
774 side of the animal. (B, E, H) Conditions were separated by whether the first target was presented on the
775 left or right side of the animal. (C, F, I) Conditions were separated by whether there was a response or
776 no response made by the animal, i.e. if it was a detect or foil trial, respectively. For error trials,
777 conditions were separated by whether there should have been a response or no response, regardless of
778 what the animal chose to do.
779
780 **Figure 12.** Passive task data. ANOVA analysis of data from passive condition, which was identical to the
781 task, except the monkey was simply required to sit passively and listen to the task structure's sounds.
782 Bars above each plot represent the bins in which a statistically significant fraction of neurons encode
783 each factor by color ($p < 0.01$). (A, D, G) During presentation of the cue, neurons respond differentially
784 to the cue location. (B, E, H) Post-cue, neurons are selective to cue side, during the delay period, in both
785 AC and dlPFC, but weakly, compared to the active task. (C, F, I) Post-target presentation, only AC
786 encodes the target, and none of the areas encode the response.
787
788 **Figure 13.** Passive data. Post-stimulus time histograms (PSTHs). Mean normalized firing rates of neurons
789 plotted using non-overlapping 25ms bins smoothed with a 3-point moving average. Bin midpoint was
790 used to align time on the x-axis. Analysis was conducted to assess precise timing of changes in neuronal
791 firing rates in AC, dlPFC and BLA. Paired t-tests were performed on all bins to determine significant
792 difference in firing rates. (A, C, E) compares conditions that are identical in cue side but vary in target
793 side, as a measure of sensory identification. (B, D, F) compares conditions that are identical in target side
794 but vary in the cue location (left or right). (G, I, K) Conditions are matched for cue side but vary in target
795 side. (H, J, L) Conditions shown have opposite cue sides but matched target side.
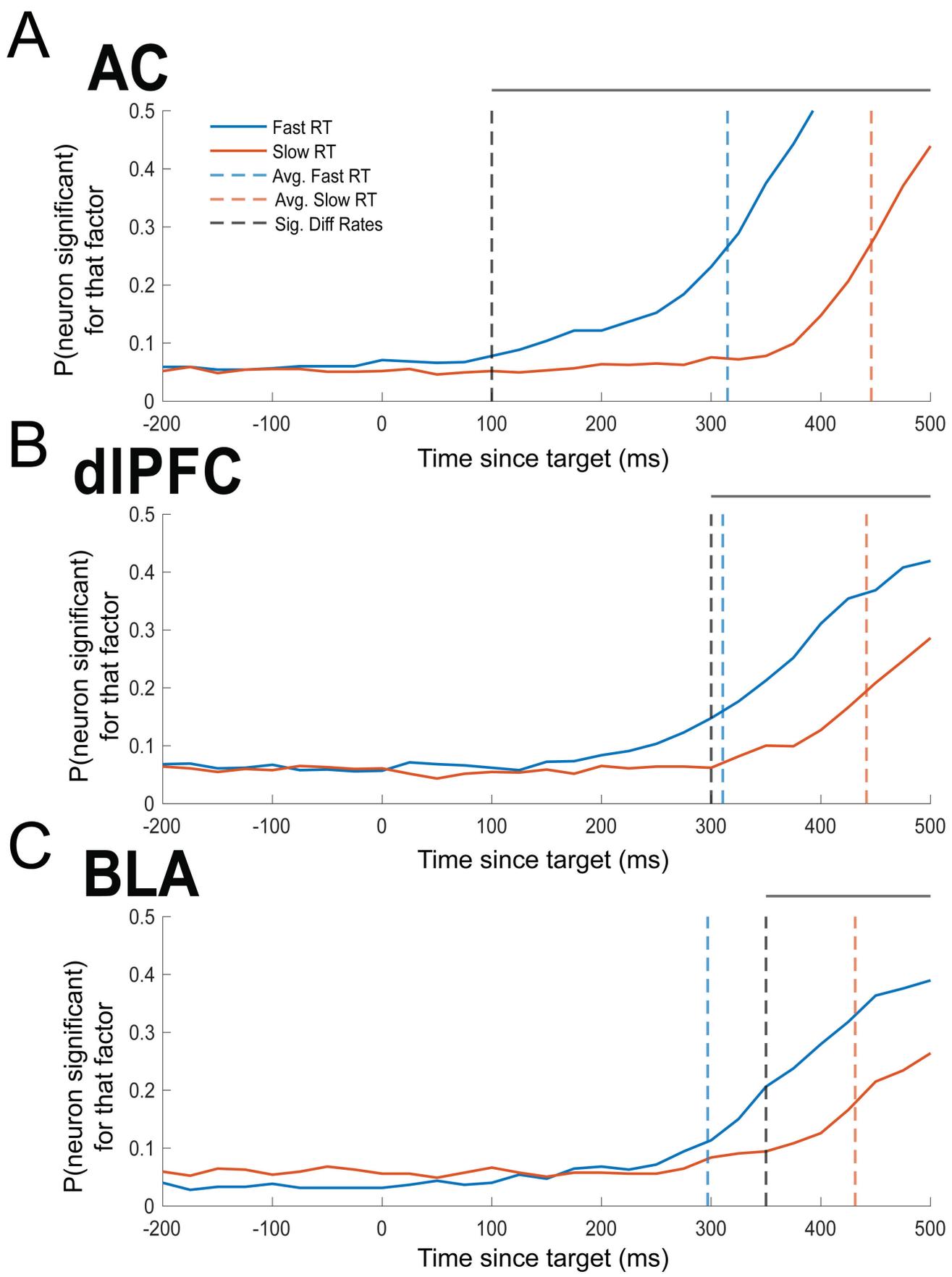
Monkey 1

Monkey 2

Cue Epoch

Target Epoch

A **AC**

B

Attention Ipsi,
Target Ipsi (detect)

Attention Ipsi,
Target Contra (foil)

Attention Contra,
Target Contra (detect)

Attention Contra,
Target Ipsi (foil)

C **dlPFC**

D

E **BLA**

F

A AC

B dlPFC

C BLA

A **AC**

Cue selectivity   Cue selectivity   Target selectivity   Response selectivity

Neuron

Time since cue (ms)   Time since noise (ms)   Time since target (ms)   Time since target (ms)

B **dlPFC**

Cue selectivity   Cue selectivity   Target selectivity   Response selectivity

Neuron

Time since cue (ms)   Time since noise (ms)   Time since target (ms)   Time since target (ms)

C **BLA**

Cue selectivity   Cue selectivity   Target selectivity   Response selectivity

Neuron

Time since cue (ms)   Time since noise (ms)   Time since target (ms)   Time since target (ms)

Cue Epoch

Target Epoch

AC

dlPFC

BLA

Legend (panel A):
- Attention Contra, Target Contra (detect)
- Attention Contra, Target Ipsi (foil)
- Attention Ipsi, Target Contra (foil)

AC

A

B

C

dlPFC

D

E

F

BLA

G

H

I

Correct Trials
Error Trials

AC

A

B

C

dlPFC

D

E

F

BLA

G

H

I

AC

A — Time since cue (ms). P(correct classification of response). Correct Trials, Error Trials, Rates Differ.

B — Time since noise (ms)

C — Time since target (ms). 270 ms

dlPFC

D — Time since cue (ms)

E — Time since noise (ms)

F — Time since target (ms). 275 ms

BLA

G — Time since cue (ms)

H — Time since noise (ms)

I — Time since target (ms). 300 ms

Cue Epoch

Target Epoch

AC

A

Mean normalized firing rate

Cue Contra, Target Contra (detect)
Cue Contra, Target Ipsi (foil)

Time since cue (ms)

B

32 ms

Cue Contra, Target Contra (detect)
Cue Ipsi, Target Contra (foil)

Time since cue (ms)

G

53 ms

Time since target (ms)

H

Time since target (ms)

dlPFC

C

Mean normalized firing rate

Time since cue (ms)

D

Time since cue (ms)

I

Time since target (ms)

J

Time since target (ms)

BLA

E

Mean normalized firing rate

Time since cue (ms)

F

Time since cue (ms)

K

Time since target (ms)

L

Time since target (ms)

also encoded	Encoding cue			Encoding delay		Encoding target
	Delay	*Target*	*Response*	*Target*	*Response*	*Response*
AC	26.70	17.05	5.40	16.9	8.81	5.83
dlPFC	30.90	10.11	4.49	9.80	3.92	5.80
BLA	25.00	8.33	2.08	5.77	7.69	8.00

Note: all results displayed are in percentages

Table 1. Neuron selectivity to task variables by brain area. 2 x 2 ANOVA ($p < 0.05$) was performed using one bin per time period. A single time bin was chosen at the peak of the ANOVA curves (Fig. 3). Cue presentation included 150 ms – 450 ms post-cue, the delay period included 100 ms – 400 ms post-noise, target presentation included 50 ms – 400 ms post-target presentation and choice period included 200 – 500 ms post-target presentation.