

Eye movements during auditory attention predict individual differences in dorsal attention network activity

Rodrigo M. Braga^{1, 2, 3*}, Richard Z. Fu¹, Barry M. Seemungal¹, Richard J. Wise¹, Robert Leech¹

¹Computational, Cognitive and Clinical Neuroimaging Laboratories, Imperial College London, United Kingdom, ²Center for Brain Sciences, Harvard University, USA, ³Athinoula A. Martinos Center for Biomedical Imaging, USA

Submitted to Journal:
Frontiers in Human Neuroscience

ISSN:
1662-5161

Article type:
Original Research Article

Received on:
20 Dec 2015

Accepted on:
01 Apr 2016

Provisional PDF published on:
01 Apr 2016

Frontiers website link:
www.frontiersin.org

Citation:
Braga RM, Fu RZ, Seemungal BM, Wise RJ and Leech R(2016) Eye movements during auditory attention predict individual differences in dorsal attention network activity. *Front. Hum. Neurosci.* 10:164. doi:10.3389/fnhum.2016.00164

Copyright statement:
© 2016 Braga, Fu, Seemungal, Wise and Leech. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution and reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Provisional

1 Eye movements during auditory attention predict individual 2 differences in dorsal attention network activity

3 **Rodrigo M. Braga*^{1,2,3}, Richard Z. Fu¹, Barry M. Seemungal¹, Richard J. S. Wise¹ & Robert
4 Leech¹**

5 ¹Computational, Cognitive and Clinical Neuroimaging Laboratory, Division of Brain Sciences,
6 Imperial College London, Hammersmith Hospital Campus, Du Cane Road, London, UK.

7 ²Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA.

8 ³Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, Massachusetts, USA.

9 * **Correspondence:** Rodrigo M. Braga, Computational, Cognitive and Clinical Neuroimaging
10 Laboratory C3NL Burlington Danes Building, Hammersmith Hospital Campus, Du Cane Road,
11 Imperial College London, W12 0NN

12 rodrigo.braga@imperial.ac.uk

13 **Keywords: Auditory attention, Listening, Eye Movements, Saccades, Dorsal Attention
14 Network, Spatial Attention.**

15 Abstract

16 The neural mechanisms supporting auditory attention are not fully understood. A dorsal
17 frontoparietal network of brain regions is thought to mediate the spatial orienting of attention across
18 all sensory modalities. Key parts of this network, the frontal eye fields (FEF) and the superior parietal
19 lobes (SPL), contain retinotopic maps and elicit saccades when stimulated. This suggests that their
20 recruitment during auditory attention might reflect crossmodal oculomotor processes; however this
21 has not been confirmed experimentally. Here we investigate whether task-evoked eye movements
22 during an auditory task can predict the magnitude of activity within the dorsal frontoparietal network.
23 A spatial and non-spatial listening task was used with on-line eye-tracking and functional magnetic
24 resonance imaging. No visual stimuli or cues were used. The auditory task elicited systematic eye
25 movements, with saccade rate and gaze position predicting attentional engagement and the cued
26 sound location, respectively. Activity associated with these separate aspects of evoked eye-
27 movements dissociated between the SPL and FEF. However these observed eye movements could
28 not account for all the activation in the frontoparietal network. Our results suggest that the
29 recruitment of the SPL and FEF during attentive listening reflects, at least partly, overt crossmodal
30 oculomotor processes during non-visual attention. Further work is needed to establish whether the
31 network's remaining contribution to auditory attention is through covert crossmodal processes, or is
32 directly involved in the manipulation of auditory information.

33

34

35 **1 Introduction**

36 The visual and auditory sensory systems can be thought of as serving a single role – to gather
37 information about our surroundings so that we may adapt our behavior accordingly. When a loud
38 sound alerts us to a potentially dangerous situation, our eyes instinctively orient towards the source of
39 that sound to gain further knowledge of its identity. This example highlights that the two sensory
40 systems are intimately linked, with attention-capture in one modality often leading to recruitment of
41 the other. Despite this, we are able to control which features of a given sensory modality we wish to
42 pay attention to. This suggests some degree of modal separation in the ‘top-down’ or ‘endogenous’
43 attentional modulation of sensory information. This nuanced relationship poses a conundrum for
44 establishing the neural correlates of auditory and visual attention, if indeed they are subserved by
45 separate systems.

46 In vision, a great body of research has established that a dorsal frontoparietal network comprised of
47 the superior parietal lobe (SPL) and frontal eye fields (FEF) activated during top-down attention
48 (Kincade et al. 2005; Vossel et al. 2006, Corbetta et al. 2008). This network is sometimes called the
49 “dorsal attention network” (DAN; Corbetta et al. 2008), and is active for example during the
50 maintenance of attention to a visual field location in anticipation of a visual stimulus (Kastner et al.
51 1999). There is complementary evidence that the core nodes of this network, the SPL and FEF, do
52 have a primarily visuospatial role. The regions are known to be involved in eye movement control
53 (Büttner-Ennever and Horn 1997) and visuospatial processing (Behrmann 2004). Retinotopic maps
54 have been located in both the FEF and SPL using direct stimulation and functional neuroimaging
55 (Moore et al. 2003; Ruff et al. 2008; Saygin and Sereno 2008). The FEF and neighboring
56 supplementary eye fields (SEF) were shown to be activated during the planning of subsequent
57 saccades using electrophysiology and functional MRI (Isoda and Tanji 2003; Hu and Walker 2011).
58 Further, transcranial magnetic stimulation (TMS) of the FEF delays voluntary saccades (Muggleton
59 et al. 2011). These different lines of evidence converge on the DAN having a role in oculomotor
60 control and visual orienting.

61 In hearing, the networks subserving top-down attention are not as well understood. A frontotemporal
62 network consisting of the middle and inferior frontal gyri (IFG) and regions near the posterior
63 superior temporal sulcus (pSTS), has been proposed to mediate the orienting of attention to non-
64 spatial features of sounds, such as frequency and identity (Braga et al. 2013, Maeder et al. 2001,
65 Salmi et al. 2007, Seydell-Greenwald et al. 2013). In support of this, activity in the SPL and FEF is
66 notably absent from tasks that require auditory attention such as speech and music perception
67 (Hickok et al. 2003; Warren 2008). However, during sound localization dorsal frontoparietal activity
68 is often observed (e.g. Alho et al. 2015) even in the absence of visual stimuli (Maeder et al. 2001;
69 Shomstein and Yantis 2006; Hill and Miller 2010; Alain et al. 2001; Lewis et al. 2000; Petit et al.
70 2007). The FEF even shows preparatory activity for spatial listening in the absence of auditory or
71 visual stimuli (Lee et al. 2013). This has led to the theory that the dorsal frontoparietal network is
72 “amodal” and directly mediates attentional orienting to all sensory modalities (Posner and Petersen
73 1990; Driver and Spence 1998; Macaluso 2010).

74 It is difficult to reconcile the DAN’s role in eye movement control with its recruitment during
75 auditory orienting without recourse to a possible crossmodal cause (Driver and Spence, 1998).
76 Auditory attention is likely to involve both direct modality-specific as well as indirect cross-modal
77 processes, and the role of the DAN in this regard is not clear. On the one hand, activation of the DAN
78 could represent direct manipulation of auditory processes, such as the top-down tuning of auditory

79 spatial receptive fields in the auditory cortex (Fritz et al. 2010). Alternatively, the activation of the
80 DAN during listening may be representative of indirect processes such as visual spatial orienting or
81 task-induced eye movements. Such processes may be facilitatory for auditory attention, even if not
82 directly involved in modulating auditory information.

83 There is behavioral evidence that auditory attention elicits systematic eye movements (Paulsen and
84 Ewertzen 1966; Rolfs et al. 2005; Valsecchi and Turatto 2009; Kerzel et al. 2010; Yuval-Greenberg
85 and Deouell 2011; Zou et al. 2012). For example, an auditory stimulus leads reliably to visual
86 saccades towards the source of the sound (Zahn et al. 1978; Zambarbieri et al. 1982; Van Grootel and
87 Van Opstal 2009). Rotating a sound about a subject's head can induce nystagmus (Paulsen and
88 Ewertzen 1966). Further, the presentation of an auditory stimulus can reduce the rate of saccades
89 (Kerzel et al. 2010; Yuval-Greenberg and Deouell 2011; Zou et al. 2012; Rolfs et al. 2005). There is
90 also evidence that gaze position can affect auditory localization accuracy (Maddox et al. 2013).
91 However, neuroimaging studies rarely consider the influence of eye movements on auditory
92 attention, meaning that parts of the networks implicated in auditory attention may in fact be
93 mediating these crossmodal effects.

94 Given that the DAN, and particularly the FEF, is known to be involved in the generation of saccades,
95 it is possible that its recruitment during listening tasks reflects task-induced indirect oculomotor
96 processes. To test this hypothesis, eye movements need to be recorded in an auditory attention task in
97 the absence of visual stimuli and without any requirement for saccades or fixation. In contrast, the
98 usual approach has been to employ fixation conditions to investigate how gaze position affects
99 auditory processing (e.g. sound localization performance; Maddox et al. 2013). If DAN activity
100 during listening is associable with indirect processes such as increased eye movement control while
101 other parts of the auditory network are not, this would provide evidence that DAN activation is the
102 result of an indirect attentional route. This evidence would be particularly strong if the magnitude of
103 DAN activity represents the magnitude of indirect crossmodal influences. In the present manuscript
104 we recorded eye movements during a purely auditory attention task using functional magnetic
105 resonance imaging (fMRI) and in-scanner eye-tracking (Figure 1). Our hypothesis was that attentive
106 listening would be associated with systematic effects on eye movements, and that these effects would
107 be associated with the magnitude of activity within visuospatial regions of the DAN.

108 **2 Materials and Methods**

109 **2.1 Subjects**

110 Twenty healthy right-handed volunteers (9 female, mean age 26.2, range 21 to 36). All participants
111 reported no hearing problems and had normal or corrected vision (via contact lenses or MRI
112 compatible glasses). The study was conducted in accordance with the guidelines of Imperial College
113 Research Ethics Committee, and written consent was obtained from all volunteers before their
114 participation. Participants were screened for contraindications to MRI, and were excluded on the
115 basis of hearing difficulties and previous psychiatric or neurological disorders. Four additional
116 participants were scanned but had to be excluded due to technical issues with the eye tracking
117 equipment (3) and excessive motion (1).

118 **2.2 Auditory Task**

119 Subjects listened to 12 different looped melodies. Each melody lasted 2s, and was repeated four
120 times in each trial so that each trial lasted 8s. The melodies were generated by manually selecting

121 sequences of diatonic notes within one octave using prepackaged synthesizer sounds from the Logic
122 Pro X software (version 10.2.2). Each melody contained between 7 and 12 staccato notes (mean 10
123 notes). Six of the melodies had a low tonal center (F above C1), and six had a high tonal center (C3),
124 with no overlap in pitch between high and low pitch sequences. In each trial, two looped melodies,
125 one high and one low, were presented simultaneously, either dichotically (one sound in each ear) or
126 diotically (both sounds in both ears equally; Figure 1). The competing melodies overlapped in terms
127 of note onset and duration, but not pitch. Stimuli were presented using Sensimetrics S14 sound-
128 attenuating in-ear MR-compatible headphones. Subjects were trained outside the scanner to listen out
129 for an oddball target in the form of a “pitch change”, which was in fact a transposition of the whole
130 2s melody to a tonal center 7 semitones above the original key. This key change made the two
131 melodies incongruous and the target detectable. The task was split into 2 identical blocks. Each block
132 contained 70 trials, of which 23 contained a pitch change in the cued melody (the target), 7 contained
133 a distractor pitch change in the non-cued melody (catch trials), and 10 were silent rest trials. Targets
134 were not presented in 30/70 trials. Targets and distractors were presented in either the second (7/70),
135 third (9/70) or fourth (14/70) repeat of the 2s diatonic melody, in a pseudo-randomized order to avoid
136 long repeats and ensure an even distribution of Rest trials. Distractors and targets were never
137 presented in the same trial. The order of stimulus presentation was changed halfway through the
138 experiment (6 subjects received one order and 14 received the other) to control for order-effects. We
139 performed a confirmatory analysis with balanced groups (n=6 subjects receiving each presentation
140 order) which confirmed that the eye movement behavioral results reported were not due to order-
141 effects. Before each trial, subjects were presented with a diotic auditory spoken word (“Right”,
142 “Left”, “High” or “Low”) which cued them to listen to the melody presented in their right or left ear,
143 or that was higher or lower in pitch, respectively. The cue period lasted 2s, and the spoken cue onset
144 was at the start of those 2s (not centered within the 2s). The task period lasted 8s, and the response
145 period lasted 3s (including an auditory ‘Please respond’ cue). Each trial was followed by a period of
146 silence lasting between 1-3s. Ten silent ‘Rest’ trials were also interspersed between listening trials.
147 These were preceded by an auditory spoken cue (“Rest”), and no auditory stimuli were presented for
148 the same duration as a normal trial and response period. No “Respond” cue was presented after ‘Rest’
149 trials. Subjects were instructed to keep their eyes open throughout the listening experiment. A
150 featureless black screen was displayed during the whole experiment and no instructions to fixate
151 were given. Subjects were naïve to the purpose of the experiment, and were told that the eye tracker
152 would be used for a separate visual task that took place between the two blocks of the auditory task.

153 2.3 Saccade Distractor task

154 In between each block of the auditory task, subjects performed a visual distractor task that had 2
155 components: 1) visual fixation to a central cross (white on black background), and 2) forced saccades
156 to a white cross that appeared unpredictably on each corner of the black screen (see Eye Tracking).
157 These two tasks lasted 32 seconds each and were repeated 4 times. Four rest periods of 32 seconds
158 duration were interspersed between tasks, wherein a blank screen was presented. The distractor block
159 served as an explanation to subjects for the presence of the eye tracking equipment, for calibration of
160 the eye tracker, and to functionally localize the DAN.

161 2.4 Eye Tracking

162 Vertical and horizontal gaze displacements were recorded at 500 Hz using a MR-compatible head-
163 mounted infrared camera (Jazz-NOVO, Ober Consulting, Eye movement range - vertical: +/- 20°,
164 horizontal +/-35°; sampling frequency - 500Hz). The voluntary saccades element of the visual
165 distractor task served as a four-point calibration and was performed in between the two runs of the

Eye movements during auditory attention

166 auditory task to be close in time to both runs. In this task, white crosses were presented in each
167 corner of a black 7.5" IFIS-SA LCD screen. The screen was at a viewing distance of 13 cm. The
168 crosses subtended a horizontal angle of 60° and vertical angle of 40° from one another. Gaze
169 displacements to the left, right, upper and lower visual spaces were quantified relative to a center
170 point, which was defined as the average vertical and horizontal gaze position across the run.

171 Eye movements were analyzed using the Jazz-Manager software (saccade detection, blink removal)
172 and using in-house software based on MATLAB (normalization and gaze displacement
173 measurement). For each participant, gaze displacement along both axes was detrended to remove low
174 frequency drifts, and normalized by dividing by the standard deviation within each 17min run. The
175 saccade detection algorithm involved first a differentiation of the raw eye position signal (in degrees
176 of angle) to velocity ($^{\circ}/s$). A saccade was then determined if it satisfied all of the following criteria:
177 (i) an initial velocity of 35 $^{\circ}/s$ or greater; (ii) a minimum peak velocity of 100 $^{\circ}/s$; (iii) a maximum
178 duration of 300 ms; (iv) a minimum duration of 20 ms; (v) an inter-saccadic interval of 50 ms or
179 greater (since very short inter-saccadic intervals of <50 ms would indicate either artifact or
180 pathological eye movements such as ocular flutter or opsoclonus).

181 The eye blink detection algorithm uses the fact that during an eye blink, the eye moves primarily in
182 the vertical plane, first up and then down ("Bell's phenomenon") and hence this algorithm used only
183 the vertical eye signal. It also follows that blinks are bi-phasic with two velocity peaks per blink. The
184 eye blink detection algorithm used the following criteria: (i) a first peak minimum velocity of 200
185 $^{\circ}/s$; (ii) a second peak minimum velocity of 100 $^{\circ}/s$; (iii) a maximum inter-velocity-peak duration of
186 150 ms; (iv) a total blink duration of between 100 ms (minimum) and 500 ms (maximum); (v) an
187 amplitude of 15 $^{\circ}$ or greater; (vi) since the vertical eye position at the end of a blink is typically close
188 to the pre-blink position, a ratio of the final to initial vertical eye position for a blink should be close
189 to 1. This ratio was set to between 0.6 (minimum) and 1.4 (maximum).

190 2.5 MRI acquisition

191 MRI data were obtained using a Phillips Intera 3T MRI system with an 8-element phased array head
192 coil and sensitivity encoding. High-resolution (1 mm \times 1 mm \times 1 mm) T1-weighted whole-brain
193 structural images were obtained for each participant to allow accurate spatial registration of the
194 functional images. Functional MRI data were acquired using an echoplanar imaging (EPI) sequence.
195 Continuous data acquisition was used to collect whole-brain images in 44 axial slices with a slice
196 thickness of 3.5mm, and a repetition time (TR) of 3s (TE=45ms, FOV=220 \times 143 \times 190mm, 44 slices,
197 128 x 128 voxels per slice, voxel size: 2.18 x 2.18 x 3.25 mm). The first three dummy scans were
198 discarded. A total of 674 whole brain functional images were acquired for each subject, split into two
199 runs of 337 images. Paradigms were programmed using MATLAB and stimuli presented through an
200 IFIS-SA system (In Vivo Corporation). Eye tracking data and stimulus presentation were
201 synchronized to the initial scanner pulse. Responses were recorded through a fiber optic response box
202 (Nordicneurolab, Norway), interfaced with the stimulus presentation PC running MATLAB.

203

204 2.6 fMRI image analysis

205 Standard preprocessing was carried out using FSL (FMRIB's Software Library (Smith et al. 2004),
206 www.fmrib.ox.ac.uk/fsl). Image pre-processing involved realignment of EPI images to reduce the
207 effects of motion between scans, spatial smoothing using a 8mm full-width half-maximum Gaussian
208 kernel, pre-whitening using FILM and temporal high-pass filtering using the default cut-off
209 frequency of 100 Hz. FMRIB's Linear Image Registration Tool (FLIRT) was used to register EPI
210 functional datasets into a standard MNI space using the participant's individual high-resolution
211 anatomical images. We also ran three separate confirmatory analyses using nonlinear registration and
212 spatial smoothing at 6 and 10mm to test that our choice of registration and smoothing parameters
213 were not affecting the result. These analyses produced qualitatively similar activation patterns to
214 those reported.

215 The listening task was divided into four conditions depending on the cue and dichotic/diotic method
216 of stimulus presentation. Seven variables were entered into a general linear model with the onsets and
217 durations of the Cue, Response, Listen-Left, Listen-Right, Listen-High/Low (diotic) and Listen-
218 High/Low (dichotic) trial periods, and one additional Error variable that modeled the trials in which
219 subjects responded incorrectly (across all conditions). The model included the full duration of each
220 condition. 'Rest' trials were not modeled in the general linear model and served as the implicit
221 baseline along with the silent periods following each trial. To rule out that the observed behavioral
222 and neuroimaging results might be a result of participants closing their eyes during the trials, we ran
223 a confirmatory analysis where we removed any trials during which no saccades were detected. In this
224 separate confirmatory analysis, no-saccade trials were included in the Error variable and excluded
225 from the task condition variables. This analysis produced similar neuroimaging and eye tracking
226 results. Individual variability in the number of trials without saccades also did not correlate with any
227 of the behavioral rankings (gaze position bias, saccade inhibition or task inhibition). A synthetic
228 double-gamma hemodynamic response function was convolved with each explanatory variable and
229 its first temporal derivative was included to account for variability in the hemodynamic delay
230 function. Six motion parameters were included in the general linear model as confound regressors.
231 To conclusively rule out motion as a potential confound, we ran a confirmatory analysis using 24
232 motion regressors which produced qualitatively similar results.

233 **2.7 Group fMRI analysis**

234 Mixed effects analysis of session and group effects was carried out using FLAME (FMRIB's Local
235 Analysis of Mixed Effects(Beckmann et al. 2003)). Final statistical images were thresholded using
236 Gaussian Random Field based cluster inference with an initial cluster-forming threshold of $Z > 2.3$ and
237 a cluster significance threshold of $p < 0.05$. This resulted in statistical maps of clusters significantly
238 activated by the task. Group-mean images were produced by giving each subject equal weighting.
239 Subjects were also ranked by three behavioral variables; 1) mean task performance, 2) mean
240 difference in saccade rate between all task and 'Rest' trials, and 3) mean gaze position difference
241 between 'Listen-Left' and 'Listen-Right' conditions. These rankings were zero-meaned and used as
242 weightings for each subject to probe individual differences in each behavioral measure. These
243 weightings were entered into the same higher-level model. We also carried out a confirmatory group-
244 level analysis using each subject's (demeaned) t-values for gaze position bias and saccade inhibition
245 instead of their relative ranking. This analysis revealed qualitatively similar results to those reported.

246 **2.8 Functional connectivity analysis**

247 To test the network membership of our eye-movement-derived regions of interest, we used resting
248 state data from 20 participants data from the Human Connectome Project (Smith et al. 2013; Van

249 Essen et al. 2013). This dataset was used because of its high quality, and because resting state data
250 was not collected from the participants who took the auditory task. The data consisted of the first 20
251 subjects from the third phase public release from the Human Connectome Project. Each resting state
252 acquisition was composed of 1200 whole-brain volumes, with a TR=0.72, collected on a 3T MRI
253 scanner. We used the spatially and temporally preprocessed version of the data. Each run was
254 corrected for spatial distortions from gradient nonlinearity and from motion by registration to a
255 reference image as well as corrected for B0 distortion before being registered to a high-resolution
256 structural image and into MNI standard space. A liberal 2000s cut-off for a high-pass temporal filter
257 was applied to the data. Twenty-four motion parameters were then aggressively temporally filtered
258 out of the data, along with other non-neural structured noise identified with FIX automatic
259 independent component denoising approach (Salimi-Khorshidi et al. 2014), as described by Smith et
260 al (Smith et al. 2013). In addition, the data was downsampled into 4x4x4mm space, to reduce the
261 computational overhead. Functional connectivity was calculated using the dual regression
262 (Beckmann et al. 2009) pipeline from FSL version 5. Two regions of interest (the activation pattern
263 corresponding to individual variability in either eye gaze displacement or saccade inhibition) were
264 entered separately into the dual regression as the design matrix for a general linear model with the
265 HCP rest data as the dependent variable. This resulted in a timecourse which was then regressed
266 again with the rest data, resulting in a whole-brain spatial map of regression coefficients estimating
267 functional connectivity with the initial pattern of activation. The functional connectivity map for each
268 subject was then entered into a higher-level general linear model. Results were thresholded using a
269 family-wise error correction for multiple comparisons.

270 3 Results

271 3.1 Behavioral results:

272 3.1.1 Natural gaze position is biased towards attended sound location

273 Despite the absence of visual cues and stimuli, following spatial ('Left' or 'Right') auditory cues,
274 subjects tended to shift their gaze left or right consistent with the auditory spatial task (left/right;
275 Figure 2A). A significant difference in mean gaze position was observed between trials when
276 subjects were cued to listen to their left vs. right ear (t-test of "Left" vs "Right" cue trials, $t_{19} = -4.54$,
277 $p < 0.001$, Figure 3A). This gaze position bias was not observed for dichotic trials preceded by a
278 spectral (i.e. "High" or "Low") cue (Figure 2B). No effects on vertical gaze position were observed
279 for Right-Left or High-Low discriminations (Left-Right: $t_{19} = -0.42$ $p = 0.68$; High-Low: $t_{19} = 0.25$
280 $p = 0.81$).

281

282 3.1.2 Saccade rate decreases during auditory attention

283 We used a within-subject design to compare saccade rate between task trials and the ten silent 'Rest'
284 trials which were interspersed across each block. The auditory attention task was associated with a
285 lower saccade rate compared to rest ($t_{19} = 1.973$, $p < 0.05$, Figure 3B). This was observed across all
286 trials, regardless of cue type.

287

288 **3.1.3 Individual variability in eye movements and performance**

289 Subjects performed the task with high accuracy (average 92.1% correct responses, standard deviation
 290 8.3%). No significant differences in performance were observed between spatial (Left-Right) and
 291 spectral (High-Low) discrimination conditions (Left-Right: 93.9 %, High-Low: diotic 91.4 %,
 292 dichotic 91.0%, pairwise t-tests: all $p > 0.05$, n.s.). Subjects ignored the distractor pitch change
 293 successfully in the majority of catch trials (percentage of catch trials with false alarms: 8.2 %). In
 294 terms of eye movements, high variability in effect of task was observed across participants for both
 295 saccade rate and gaze displacement following spatial cues (Figure 3C). We assessed whether the
 296 individual differences in eye movement control predicted individual differences in performance on
 297 the auditory task. The task performance scores did not correlate with the inhibition of saccade rate
 298 (Figure 3D, $R^2 < 0.01$, n.s.) or gaze position bias ($R^2 = 0.01$, n.s.) across subjects. Gaze position bias
 299 and saccade rate inhibition across participants were also not correlated with each other (R^2 all < 0.01 ,
 300 n.s.).

301

302 **3.2 Neuroimaging results:**

303 **3.2.1 Activity in frontotemporal and DAN areas relates to auditory attention and individual**
 304 **variation in task performance**

305 Across all task conditions, activation in widespread regions was observed during attentive listening
 306 (Figure 4). This network included the SPL, FEF and middle frontal gyrus (MFG) in both
 307 hemispheres, as well as both superior temporal gyri and sulci, and regions of the cerebellum. We
 308 ranked subjects by their performance scores (% percentage correct) in the pitch-change detection task
 309 (Figure 3C), and assessed which brain regions had activity relating to this rank order during the task.
 310 Increased activation of the same widespread DAN and frontotemporal network was associated with
 311 improved performance on the task. No differences in the mean (unranked) signal were observed
 312 between specific task conditions, such as “Left”, “Right”, “High” or “Low” (diotic and dichotic)
 313 trials.

314

315 **3.2.2 Gaze position bias mediated by the SPL**

316 To determine which brain regions mediated the behavioral gaze bias evoked by the spatial task
 317 (which were observed during the ‘Left’ and ‘Right’ cue conditions; Figure 3A) whilst controlling for
 318 some of the listening requirements and auditory input, we contrasted the spatial and spectral listening
 319 conditions (‘Listen-Left’ + ‘Listen-Right’ > ‘Listen-High or Low’ dichotic). Although no differences
 320 were found in the mean signal for this contrast, this may have been a consequence of the individual
 321 variability present in gaze position bias (Figure 3C). Therefore, we ranked subjects by their
 322 difference in mean gaze displacement between ‘Listen-Left’ and ‘Listen-Right’ trials (average gaze
 323 position during all “Listen-Right” trials minus average gaze position during all “Listen-Left” trials).
 324 Subjects that showed the greatest mean gaze position difference between left and right trials were
 325 given the highest rank, and displayed higher activation of the posterior parietal and superiolateral
 326 occipital lobes bilaterally during the spatial compared to spectral listening trials (Figure 4). No
 327 differences were found for the direct contrast of ‘Listen-Left’ > ‘Listen-Right’.

328

3.2.3 Saccade inhibition effect mediated by the FEF

Substantial individual differences in the task-induced inhibition of saccades were also observed (Figure 3B). Therefore, to assess which brain regions mediated this effect, we ranked the participants by the amount of reduction in their saccade rate during the task (saccade rate during listening trials minus saccade rate during rest periods; Figure 3C). Subjects with the largest reduction in saccade rate between rest and task trials were given the highest rank. No regions of activation were observed during the task (compared to the implicit baseline) for this rank analysis. However, during the cue period immediately preceding the task subjects that showed the greatest reduction in saccade rate displayed higher activation of the FEF and MFG bilaterally, as well as some activation in the left superior temporal gyrus (Figure 4).

3.2.4 Auditory task-evoked eye movement regions overlap with and dissociate anterior and posterior components of the DAN

The visual distractor task was used to functionally localize the DAN in our dataset (Figure 5). The activation patterns obtained from the rank analyses of gaze position bias and saccade inhibition were located primarily within regions of the DAN, as evoked by the independent visual task acquired with the same participants. To provide a more detailed description of these two patterns of activation, their functional connectivity with the rest of the brain was assessed using resting state data acquired as part of the Human Connectome Project (Van Essen et al. 2013). The regions activated by variability in gaze position were functionally connected to the DAN, including SPL, FEF and supplementary eye fields near the midline. There was also extensive connectivity with the dorsal and ventral visual streams bilaterally, extending from the SPL via the occipital lobes to the fusiform gyri (Figure 5). Regions activated by variability in saccade inhibition were also functionally connected to the whole DAN, including FEF and SPL, and also the MFG, and visual streams to a lesser extent. There was also extensive connectivity with posterior superior and posterior middle temporal cortices, and anterior regions of the lateral prefrontal cortices. Both functional connectivity maps overlapped considerably with the DAN as evoked by the visual task (Figure 5).

4 Discussion

The study demonstrates that attentive listening is associated with changes in eye movements that are independent of visual stimuli or visual demands. In addition, this study shows that these crossmodal effects are associated with increased activity in core regions of the dorsal frontoparietal network, the FEF and SPL. Rather than being directly involved in the attentional selection of auditory information, the present results suggest that the role of the DAN during auditory attention is at least partly an indirect one; mediating task-evoked eye movements. It remains to be determined whether the remaining role of the DAN in auditory attention also represents crossmodal process (such as covert multimodal attentional orienting without eye movements, the formation of supramodal spatial maps, or the inhibition of non-auditory inputs) or truly represents direct attentional modulation of auditory information.

Although the neural correlates of eye movements have been investigated previously (e.g. Nobre, et al., 2000), to our knowledge, this is the first study that has simultaneously shown that eye movements are evoked by a purely auditory task and that these eye movements are associated with the DAN. This has important implications for the supposed role of the DAN as an amodal

371 attentional system in that it suggests that the DAN mediates crossmodal oculomotor processes, which
372 may or may not be intrinsic to attention to all modalities. Further work is therefore necessary to
373 establish the exact contribution of this brain system to non-visual attention. It is likely that these
374 frontoparietal regions, which are activated under many conditions and have been associated with
375 numerous cognitive processes (Cabeza and Nyberg, 2000), may play multiple roles even within the
376 context of visual and auditory processing.

377 **4.1 Eye movement control during attentive listening**

378 In this study, participants reduced their rate of saccadic eye movements when they were required to
379 listen attentively (Figure 3B). As no visual cues or stimuli were presented at any stage of the auditory
380 task, and there were no behavioral requirements for eye movements, this natural inhibition of
381 saccades was driven by the requirement to listen attentively. The inhibition of eye movements when
382 attentional resources are required in the auditory modality might serve to reduce the amount of novel
383 incoming visual information which could interfere with the auditory task. Another explanation could
384 be that the resources normally used to plan and execute saccades are diverted to auditory feature
385 selection or spatial orientation during attentive listening. Either way, the auditory-evoked saccade
386 rate inhibition suggests that auditory attention competes with visual control processes during natural
387 listening conditions.

388 In addition, participants tended to look towards the direction of the cued sound after receiving spatial
389 auditory cues (i.e. to listen to the melodies in their left or right ear; Figure 2). Previous research
390 (Zahn et al. 1978; Zambarbieri et al. 1982) has shown that saccades are often made towards a
391 presented sound, an instance of crossmodal effects during ‘bottom-up’ attention capture. In our study,
392 subjects were presented with sounds in both ears, making it unlikely that ‘bottom-up’ auditory
393 attention capture determined the gaze position bias. This suggests that top-down or endogenous
394 auditory attention, driven by the spoken instructions, influenced gaze position during the present task.
395 Previous behavioral studies have shown that gaze position has a substantial effect on the accuracy of
396 auditory spatial localization accuracy (Razavi et al. 2007; Pavani et al. 2008; Van Grootel and Van
397 Opstal 2009; Pages and Groh 2013, Maddox et al. 2013). Subjects with no functioning visual system
398 were found to be severely impaired on a spatial but not a non-spatial auditory attention task (Gori et
399 al. 2014). These findings suggest that auditory spatial maps are likely to be calibrated or fine-tuned
400 by gaze position, meaning that the interaction between oculomotor and auditory processes may be an
401 intrinsic facilitatory mechanism for auditory localization.

402 **4.2 Individual variability across separate dimensions of eye movement control reveals** 403 **different listening strategies**

404 Importantly, although some subjects showed consistent crossmodal effects across trials, not all
405 subjects displayed systematic eye movements during attentive listening. This heterogeneity across
406 participants is consistent with previous behavioral research (Yuval-Greenberg and Deouell 2011). In
407 the present study, individual variability in saccade rate inhibition was not correlated with individual
408 variability in gaze position bias (Figure 3C), meaning that subjects that displayed one eye movement
409 effect did not necessarily display the other. Therefore, our data suggest that saccade rate and gaze
410 position bias were two separate manifestations of auditory attention affecting eye movements,
411 possibly reflecting underlying listening strategies.

412 It is possible that the auditory-induced gaze-position bias we observed was facilitatory, helping some
413 subjects to focus on the attended sounds and ignore the competing sounds. However, no relationship

414 between gaze position and task performance was observed in this study. We were therefore unable to
415 distinguish whether gaze position aided performance or was an epiphenomenon. Here, we observed
416 a left-right gaze bias only in trials when subjects received spatial (“Left” or “Right”) cues, not for
417 non-spatial cues (“High” or “Low”), even though these could also be discriminated spatially when
418 presented dichotically. This suggests that the gaze position bias reflected a top-down strategy that
419 some subjects were employing during the spatial task. Although pitch can be discriminated without
420 any spatial information, there is prior evidence for an associations between ‘High’ and ‘Low’ pitch
421 and upward and downward visual space (Chiou and Rich, 2012). In this study we did not find any
422 vertical gaze position effects when subjects listened out for different pitch qualities. This could be
423 due to many reasons (outside the remit of this study), such as the nature of the sounds that were used
424 as stimuli, the difference in pitch between high and low competing sounds, as well as the level of
425 musical training of the subjects.

426 **4.3 SPL activity correlates with gaze position bias during spatial listening**

427 Subjects that displayed the largest difference between leftward and rightward gaze position during
428 spatially cued trials showed the greatest activation in the posterior SPL and superiolateral occipital
429 cortex. This suggests that the posterior SPL mediates the spatial orientation of visual fixation during
430 spatial listening. This has implications for studies attempting to isolate the cortical networks
431 supporting auditory attention (Hallett et al. 1999; Shomstein and Yantis 2006; Salmi et al. 2007;
432 Kong et al. 2012; Salmi et al. 2009) as our data suggest that the amount of activation in the SPL
433 corresponds with how much subjects biased their eye movements during spatial listening. It is
434 noteworthy however that the SPL regions associated with spatial gaze bias in the present study did
435 not overlap with the listening task activations (Figure 4), but did fall within the DAN as defined
436 using a visuospatial orienting task (Figure 5). Previous visual studies have shown that gaze position
437 is encoded in the posterior parietal lobe (Williams and Smith 2010), and crossmodal audiovisual
438 salience maps were located in the SPL (Nardo et al. 2013). It is possible that the SPL mediates the
439 supra-modal formation of spatial maps, partly through the cuing of gaze position (Nardo et al. 2013).
440 The recruitment of the SPL during auditory spatial attention may therefore reflect specifically this
441 cross-modal tuning of spatial maps, indicating that the SPL’s role in auditory attention may also be
442 via the indirect route. The SPL is a candidate for mediating this cross-modal tuning, as it is activated
443 during visual and auditory spatial searching (Corbetta et al. 2008; Hill and Miller 2010). In addition,
444 both the SPL and FEF are more strongly activated with auditory spatial tasks than auditory tasks
445 involving pitch discrimination (Maeder et al. 2001; Hill and Miller 2010).

446 **4.4 FEF activity correlates with saccade rate inhibition during attentive listening**

447 When we probed the brain systems that might mediate crossmodal saccade inhibition effects, subjects
448 that displayed the largest saccade rate difference between rest and task trials showed higher FEF
449 activity during the auditory cue period (Figure 4). These activations overlapped with both the
450 auditory task activations (Figure 4) and the DAN as defined by a visual task (Figure 5). No individual
451 differences in brain activity were observed during the Task period. There are different reasons why
452 this may have been the case. For one, it is not clear whether the inhibition of saccades, or in other
453 words the control of fixation duration should be expected to cause increased activity in higher-order
454 cognitive networks (Henderson and Choi, 2015). The FEF have been proposed to mediate the
455 planning of subsequent saccades (Isoda and Tanji 2003; Hu and Walker 2011). As such it is possible
456 that the observed FEF activity in the cue period represents the planning of or preparation for the
457 inhibition of the rate of saccades during the upcoming trial. It is also possible that the analysis
458 techniques deployed were not sensitive to what may be subtle neural correlates of inhibiting

459 saccades. Nonetheless, the finding of elevated FEF activity in those subjects which were prone to
460 showing overt oculomotor differences again suggests that the FEF may be, in part, mediating
461 crossmodal factors during listening, even though this is unlikely to be its only role (Bharadwaj, et al.,
462 2014). For example, the FEF shows elevated activity when attention is maintained to spatial
463 locations, even if those locations fall outside the visual field (Tark and Curtis 2009). Previous work
464 also suggests that the FEF mediates internally guided saccades, while the SPL is involved in both
465 internally and visually guided saccades (Bender et al., 2013).

466 **4.5 The DAN mediates auditory task-evoked eye movement control**

467 The activation patterns obtained for saccade inhibition and gaze position were located predominantly
468 within the DAN, and each displayed widespread functional connectivity with the remaining DAN
469 regions (Figure 5). This suggests that the regions associated with eye control during auditory
470 attention form core parts of the DAN, as shown by their intrinsic connectivity. The gaze position seed
471 in the SPL produced a posterior-loaded DAN which had stronger connectivity with visual regions.
472 The saccade inhibition seed in the FEF produced a more front-loaded DAN with increased prefrontal
473 cortex connectivity, but also interestingly with stronger connectivity to auditory regions in the
474 temporal lobes. One interpretation is that the FEF communicates intrinsically with both auditory and
475 visual regions, which makes it a stronger candidate for an amodal center than the DAN as a whole.
476 However, our results suggest that the FEF influences auditory attention, at least in part, through
477 indirect mechanisms such as the control of eye movements (possibly through suppressing eye
478 movements or altering eye movement planning) rather than through direct top-down control, e.g. the
479 modulation of auditory receptive fields (Fritz et al. 2010). It is possible, although speculative, that the
480 observed auditory effects on eye movement are primarily mediated by the functional connectivity
481 between auditory regions and the FEF, which then exerts an effect on the SPL via its strong
482 functional connectivity (i.e. the DAN). The SPL may then mediate crossmodal spatial orienting, in
483 part through the cuing of gaze position.

484 The DAN has been shown to be active under a variety of task conditions, including auditory attention
485 (Corbetta, et al., 2008). Further evidence for an auditory role comes from findings that the FEF can
486 show frequency-tagged responses to sounds (Bharadwaj, et al., 2014). The present results do not rule
487 out that the DAN plays an important part in attentional orienting to all modalities. Or indeed that the
488 DAN, and in particular the FEF, may have multiple roles. Rather, the present findings suggest that
489 one of the roles that the DAN does perform is to orient the visual system to comply with auditory
490 task demands. This might explain why the DAN is activated during orienting to both auditory and
491 visual stimuli, but is not present during the maintenance of attention to auditory stimuli (Salmi et al,
492 2007; 2009). This visual orienting process may be intrinsic to auditory orienting, particularly
493 considering how interlinked the auditory and visual systems are (Driver and Spence, 1998), and that
494 there is competition for resources between the systems (Saults and Cowan, 2007). Given that
495 crossmodal processes are likely to play a large part in successful attentional orienting, it is likely that
496 the DAN is crucial for auditory attention. However, in order to establish the DAN's exact role in
497 listening we propose that the distinction between direct and indirect mechanisms merits further study.
498 For example, this could help make sense of why stroke lesions resulting in visual neglect often do not
499 lead to auditory neglect for spectral features, but do impair auditory localization (Pavani, et al.,
500 2002). In this case, damage to the DAN has a dramatic effect on supramodal processes such as spatial
501 orienting, but would not compromise non-spatial auditory processes as its contribution to listening is
502 an indirect, and therefore not necessary one. Thus the location of lesions within the SPL (and not
503 FEF) would largely determine the selective deficits.

504 In conclusion, this study shows that auditory attention induces overt eye movements, and that these
505 eye movement effects are mediated by activity in core components of the DAN, the SPL and FEF.
506 Our data suggest that the activation of DAN regions during auditory attention is at least partly
507 attributed to oculomotor control. This is evidence for the DAN being indirectly involved in auditory
508 attention. However, we do not rule out that the DAN plays a larger role in auditory attention, though
509 it remains to be determined whether this remaining role is also indirect (e.g. through covert visual
510 system modulation without eye movements) or direct (e.g. through the modulation of auditory
511 receptive fields). As auditory attention is associated with the inhibition of non-auditory sensory
512 inputs (Langner et al. 2011), it is likely that DAN activation is essential to auditory attention, even if
513 its role is predominantly to modulate and limit interference from visual input. Nonetheless, our data
514 adds to the growing body of evidence (Braga et al. 2013; Seydell-Greenwald et al. 2013, Maeder, et
515 al., 2001; Salmi, et al., 2007, Michalka et al. 2015) that the role of the DAN in auditory attention is
516 not as clear as in vision, This evidence suggests that there may be parallel but interacting networks
517 for attention to visual and auditory modalities (Salmi et al. 2007; Braga et al. 2013, Michalka et al.
518 2015), whose mechanisms should be further studied. We propose that the dorsal frontoparietal
519 network mediates intrinsic but crossmodal aspects of auditory attention by virtue of its predominantly
520 visual role.

521 **5 Acknowledgments**

522 This research was funded by the Medical Research Council of the UK and the Wellcome Trust. We
523 wish to thank Laura Buck and Yulia Nigmatullina for assistance with the eye tracking analysis, and
524 all our participants for taking part. A subset of the data analyzed was provided by the Human
525 Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil
526 Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH
527 Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at
528 Washington University.

529 **6 References**

- 530 Alain C, Arnott SR, Hevenor S, Graham S, Grady CL (2001) "What" and "where" in the human
531 auditory system. *Proc Natl Acad Sci U S A* 98:12301-12306.
- 532 Alho K, Salmi J, Koistinen S, Rinne T. (2015) Top-down controlled and bottom-up triggered
533 orienting of auditory attention to pitch activate overlapping brain networks. *Brain Research*
534 1626:136–145.
- 535 Beckmann CF, Jenkinson M, Smith SM. (2003) General multilevel linear modeling for group
536 analysis in fMRI. *Neuroimage*. 20:1052–1063.
- 537 Beckmann CF, Mackay CE, Filippini N, Smith SM. (2009) Group comparison of resting-state fMRI
538 data using multi-subject ICA and dual regression. *Hum Brain Map Conf* 181.
- 539 Bender J, Tark KJ, Reuter B, Kathmann N, Curtis E (2013) Differential roles of the frontal and
540 parietal cortices in the control of saccades. *Brain Cogn* 83(1):1-9.
- 541 Bharadwaj, H.M., Lee, A.K., Shinn-Cunningham, B.G. (2014) Measuring auditory selective attention
542 using frequency tagging. *Frontiers in integrative neuroscience*, 8:6.
- 543 Behrmann M. (2004) Parietal cortex and attention. *Curr Opin Neurobiol*. 14:212–217.

- 544 Braga RM, Wilson LR, Sharp DJ, Wise RJS, Leech R. (2013) Separable networks for top-down
545 attention to auditory non-spatial and visuospatial modalities. *Neuroimage*. 74:77–86.
- 546 Büttner-Ennever JA, Horn AK. (1997) Anatomical substrates of oculomotor control. *Curr Opin*
547 *Neurobiol*. 7:872–879.
- 548 Cabeza, R., Nyberg, L. (2000) Imaging cognition II: An empirical review of 275 PET and fMRI
549 studies. *Journal of cognitive neuroscience*, 12:1-47.
- 550 Chiou, R., Rich, A.N. (2012) Cross-modality correspondence between pitch and spatial location
551 modulates attentional orienting. *Perception*, 41:339-53.
- 552 Corbetta M, Patel G, Shulman GL. (2008) The reorienting system of the human brain: from
553 environment to theory of mind. *Neuron*. 58:306–324.
- 554 Driver, J., Spence, C. (1998) Cross-modal links in spatial attention. *Philosophical transactions of the*
555 *Royal Society of London. Series B, Biological sciences*, 353:1319-31.
- 556 Fritz JB, David SV, Radtke-Schuller S, Yin P, Shamma SA. (2010) Adaptive, behaviorally gated,
557 persistent encoding of task-relevant auditory information in ferret frontal cortex. *Nat Neurosci*.
558 13:1011–1019.
- 559 Gori M, Sandini G, Martinoli C, Burr DC. (2014) Impairment of auditory spatial localization in
560 congenitally blind human subjects. *Brain*. 137:288–293.
- 561 Hallett M, Bushara KO, Weeks RA, Ishii K, Catalan M-J, Tian B, Rauschecker JP. (1999) Modality-
562 specific frontal and parietal areas for auditory and visual
563 spatial localization in humans. *Nat Neurosci*. 2:759–766.
- 564 Henderson, J.M., Choi, W. (2015) Neural Correlates of Fixation Duration during Real-world Scene
565 Viewing: Evidence from Fixation-related (FIRE) fMRI. *Journal of cognitive neuroscience*,
566 27:1137-45.
- 567 Hickok G, Buchsbaum B, Humphries C, Muftuler T. (2003) Auditory-motor interaction revealed by
568 fMRI: speech, music, and working memory in area Spt. *J Cogn Neurosci*. 15:673–682.
- 569 Hill KT, Miller LM. (2010) Auditory attentional control and selection during cocktail party listening.
570 20:583–590.
- 571 Hu Y, Walker R. (2011) The neural basis of parallel saccade programming: an fMRI study. *J Cogn*
572 *Neurosci*. 23:3669–3680.
- 573 Isoda M, Tanji J. (2003) Contrasting neuronal activity in the supplementary and frontal eye fields
574 during temporal organization of multiple saccades. *J Neurophysiol*. 90:3054–3065.
- 575 Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. (1999) Increased activity in human
576 visual cortex during directed attention in the absence of visual stimulation. *Neuron*. 22:751–761.
- 577 Kerzel D, Born S, Souto D. (2010) Inhibition of steady-state smooth pursuit and catch-up saccades
578 by abrupt visual and auditory onsets. *J Neurophysiol*. 104:2573–2585.

Eye movements during auditory attention

- 579 Kincade JM, Abrams RA, Astafiev SV, Shulman GL, Corbetta M. (2005) An event-related functional
580 magnetic resonance imaging study of voluntary and stimulus-driven orienting of attention. *J*
581 *Neurosci.* 25:4593–4604.
- 582 Kong L, Michalka SW, Rosen ML, Sheremata SL, Swisher JD, Shinn-Cunningham BG, Somers DC.
583 (2012) Auditory Spatial Attention Representations in the Human Cerebral Cortex. *Cereb Cortex.*
584 24:773-784.
- 585 Langner R, Kellermann T, Boers F, Sturm W, Willmes K, Eickhoff SB. (2011) Staying responsive to
586 the world: Modality-specific and -nonspecific contributions to speeded auditory, tactile, and
587 visual stimulus detection. 33:398–418.
- 588 Lee, A. K. C., Rajaram, S., Xia, J., Bharadwaj, H., Larson, E., Hämäläinen, M. S., & Shinn-
589 Cunningham, B. G. (2013) Auditory Selective Attention Reveals Preparatory Activity in
590 Different Cortical Regions for Selection Based on Source Location and Source Pitch. *Frontiers*
591 *in Neuroscience.* 6:Article 190.
- 592 Leech R, Braga RM, Sharp DJ. (2012) Echoes of the brain within the posterior cingulate cortex.
593 *Journal of Neuroscience.* 32:215–222.
- 594 Lewis J.W., Beauchamp M., DeYoe E.A. (2000) A comparison of visual and auditory motion
595 processing in human cerebral cortex. *Cereb Cortex* 10:873-888.
- 596 Macaluso E. (2010) Orienting of spatial attention and the interplay between the senses. *Cortex.*
597 46:282–297.
- 598 Maddox, R. K., Pospisil, D. A., Stecker, G. C., & Lee, A. K. C. (2014) Directing Eye Gaze Enhances
599 Auditory Spatial Cue Discrimination. *Current biology.* 24:748-752.
- 600 Maeder PP, Meuli RA, Adriani M, Bellmann A, Fornari E, Thiran JP, Pittet A, Clarke S. (2001)
601 Distinct pathways involved in sound recognition and localization: a human fMRI study.
602 *Neuroimage.* 14:802–816.
- 603 Michalka SW, Kong L, Rosen ML, Shinn-Cunningham BG, Somers DC (2015) Short-term memory
604 for space and time flexibly recruit complementary sensory-biased frontal lobe attention networks.
605 *Neuron* 87(4):882-892.
- 606 Moore T, Armstrong KM, Fallah M. (2003) Visuomotor origins of covert spatial attention. *Neuron.*
607 40:671–683.
- 608 Muggleton NG, Kalla R, Juan C-H, Walsh V. (2011) Dissociating the contributions of human frontal
609 eye fields and posterior parietal cortex to visual search. *J Neurophysiol.* 105:2891–2896.
- 610 Nardo D, Santangelo V, Macaluso E. (2013) Spatial orienting in complex audiovisual environments.
611 *Hum Brain Mapp.* 35: 1597-1614.
- 612 Nobre, A.C., Gitelman, D.R., Dias, E.C., Mesulam, M.M. (2000) Covert visual spatial orienting and
613 saccades: overlapping neural systems. *NeuroImage,* 11:210-6.
- 614 Pages DS, Groh JM. (2013) Looking at the ventriloquist: visual outcome of eye movements

- 615 calibrates sound localization. *PLoS ONE*. 8:e72562.
- 616 Paulsen J, Ewertzen HW. (1966) Audio-visual reflex: determination of the audio-visual reflex in
617 directional hearing by employment of electronystagmography. *Acta Otolaryngol. Suppl*224:211.
- 618 Pavani F, Husain M, Driver J. (2008) Eye-movements intervening between two successive sounds
619 disrupt comparisons of auditory location. *Exp Brain Res*. 189:435–449.
- 620 Pavani, F., Ladavas, E., Driver, J. (2002) Selective deficit of auditory localisation in patients with
621 visuospatial neglect. *Neuropsychologia*, 40:291-301.
- 622
- 623 Petit L., Simon G., Joliot M., Andersson F., Bertin T., Zago L., Mellet E., Tzourio-Mazoyer N.
624 (2007) Right hemisphere dominance for auditory attention and its modulation by eye position:
625 An event-related fMRI study. *Restorative Neurology and Neuroscience* 25:211-225.
- 626 Posner MI, Petersen SE. (1990) The attention system of the human brain. *Annu Rev Neurosci*.
627 13:25–42.
- 628 Razavi B, O'Neill WE, Paige GD. (2007) Auditory spatial perception dynamically realigns with
629 changing eye position. *Journal of Neuroscience*. 27:10249–10258.
- 630 Rolfs M, Engbert R, Kliegl R. (2005) Crossmodal coupling of oculomotor control and spatial
631 attention in vision and audition. *Exp Brain Res*. 166:427–439.
- 632 Ruff CC, Bestmann S, Blankenburg F, Bjoertomt O, Josephs O, Weiskopf N, Deichmann R, Driver J.
633 (2008) Distinct causal influences of parietal versus frontal areas on human visual cortex:
634 evidence from concurrent TMS-fMRI. *Cereb Cortex*. 18:817–827.
- 635 Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. (2014)
636 Automatic denoising of functional MRI data: Combining independent component analysis and
637 hierarchical fusion of classifiers. *Neuroimage*. 90:449–468.
- 638 Salmi J, Rinne T, Degerman A, Salonen O, Alho K. (2007) Orienting and maintenance of spatial
639 attention in audition and vision: multimodal and modality-specific brain activations. *Brain Struct*
640 *Funct*. 212:181–194.
- 641 Salmi J, Rinne T, Koistinen S, Salonen O, Alho K. (2009) Brain networks of bottom-up triggered and
642 top-down controlled shifting of auditory attention. *Brain Research* 1286: 155-164.
- 643 Sauls, J.S., Cowan, N. (2007) A Central Capacity Limit to the Simultaneous Storage of Visual and
644 Auditory Arrays in Working Memory. *Journal of experimental psychology. General*, 136:663-
645 84.
- 646 Saygin AP, Sereno MI. (2008) Retinotopy and attention in human occipital, temporal, parietal, and
647 frontal cortex. *Cereb Cortex*. 18:2158–2168.
- 648 Seydell-Greenwald A, Greenberg AS, Rauschecker JP. (2013) Are you listening? Brain activation
649 associated with sustained nonspatial auditory attention in the presence and absence of
650 stimulation. *Hum Brain Mapp*. 35: 2233-2252.
- 651 Shomstein S, Yantis S. (2006) Parietal cortex mediates voluntary control of spatial and nonspatial

- 652 auditory attention. *Journal of Neuroscience*. 26:435–439.
- 653 Smith SM, Beckmann CF, Andersson J, Auerbach EJ, Bijsterbosch J, Douaud G, Duff E, Feinberg
654 DA, Griffanti L, Harms MP, Kelly M, Laumann T, Miller KL, Moeller S, Petersen S, Power J,
655 Salimi-Khorshidi G, Snyder AZ, Vu AT, Woolrich MW, Xu J, Yacoub E, Ugurbil K, Van Essen
656 DC, Glasser MF, WU-Minn HCP Consortium. (2013) Resting-state fMRI in the Human
657 Connectome Project. *Neuroimage*. 80:144–168.
- 658 Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister
659 PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano
660 N, Brady JM, Matthews PM. (2004) Advances in functional and structural MR image analysis
661 and implementation as FSL. *Neuroimage*. 23 Suppl 1:S208–S219.
- 662 Tark K-J, Curtis CE (2009) Persistent neural activity in the human frontal cortex when maintaining
663 space that is off the map. *Nat Neuro* 12: 1463-1468.
- 664 Valsecchi M, Turatto M. (2009) Microsaccadic responses in a bimodal oddball task. *Psychol Res*.
665 73:23–33.
- 666 Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K, WU-Minn HCP
667 Consortium. (2013) The WU-Minn Human Connectome Project: an overview. *Neuroimage*.
668 80:62–79.
- 669 Van Grootel TJ, Van Opstal AJ. (2009) Human sound-localization behaviour after multiple changes
670 in eye position. *Eur J Neurosci*. 29:2233–2246.
- 671 Vossel S, Thiel CM, Fink GR. (2006) Cue validity modulates the neural correlates of covert
672 endogenous orienting of attention in parietal and frontal cortex. *Neuroimage*. 32:1257–1264.
- 673 Warren J. (2008) How does the brain process music? *Clin Med*. 8:32–36.
- 674 Williams AL, Smith AT. (2010) Representation of eye position in the human parietal cortex. *J*
675 *Neurophysiol*. 104:2169–2177.
- 676 Yuval-Greenberg S, Deouell LY. (2011) Scalp-recorded induced γ -band responses to auditory
677 stimulation and its correlations with saccadic muscle-activity. *Brain Topogr*. 24:30–39.
- 678 Zahn JR, Abel LA, Dell'Osso LF. (1978) Audio-ocular response characteristics. *Sens Processes*.
679 2:32–37.
- 680 Zambarbieri D, Schmid R, Magenes G, Prablanc C. (1982) Saccadic responses evoked by
681 presentation of visual and auditory targets. *Exp Brain Res*. 47:417–427.
- 682 Zou H, Müller HJ, Shi Z. (2012) Non-spatial sounds regulate eye movements and enhance visual
683 search. *J Vis*. 12.

684

685 **7 Figure legends**

686 **Figure 1: Schematic of task design.** Two competing melodies were presented simultaneously using
 687 either dichotic (one melody played to each ear) or diotic (both melodies played to both ears equally)
 688 listening. Subjects were cued to listen to the melody on the left or right ear (“Left” or “Right”), or to
 689 listen to the melody that was higher or lower in pitch (“High” or “Low”). After the melodies were
 690 completed, subjects were cued to Respond (RESP) with button presses to indicate whether they heard
 691 a pitch change in the cued melody.

692 **Figure 2: Auditory spatial attention influences natural gaze position.** A) Plot of gaze position for
 693 a single subject over all spatially cued (“Listen-Left” or “Listen-Right”) trials. B) Gaze position for
 694 all subjects grouped by cue and target sound location. Spatial cues influenced gaze position during
 695 the subsequent listening task. No gaze biases were observed for spectral discriminations (“Listen-
 696 High” or “Listen-Low”) regardless of whether stimuli were presented dichotically (“R” and “L”) or
 697 diotically (“C”). Positive values on y-axis represent rightward, and negative values represent leftward
 698 gaze displacements. Displacement axes are in arbitrary units.

699 **Figure 3: Individual variability in auditory task-evoked eye movement control and**
 700 **performance.** Across all 20 subjects, A) attentive listening following spatial cues resulted in a gaze
 701 position bias towards cued sound, and B) attentive listening reduced saccade rate compared to
 702 interleaved rest trials. C) Graphs showing t-value (tval) for each subject for the comparison of
 703 saccade rate (green; Task vs Rest) and mean gaze position difference (blue; Listen Left vs Listen
 704 Right trials). High inter-subject variability was found in saccade inhibition (all trials) and gaze
 705 position bias (spatially cued trials). D) Individual differences in saccade reduction and gaze position
 706 bias did not correlate to individual differences in task performance or with each other.

707 **Figure 4: Neuroimaging correlates of auditory attention task and variability in eye movement.**
 708 Orange: Across all trials, attentive listening was associated with activation of a widespread auditory
 709 network (superior temporal gyri and sulci bilaterally; STG) as well as the dorsal attention network
 710 (superior parietal lobes; SPL, frontal eye fields; FEF) and middle frontal gyri (MFG) bilaterally.
 711 However, only the superior parietal lobes and frontal eye fields were associated with individual
 712 differences in eye movement control. Blue: Subjects that displayed the largest left-right gaze position
 713 bias following spatial auditory cues had greater activation of the SPL during the attentive listening
 714 periods. Green: Subjects that displayed the largest reduction in saccade rate during listening trials
 715 compared to rest had greater activation of the FEF and left posterior temporal lobe during the cue
 716 period that preceded each trial. Colorbars show cluster-corrected z-scores.

717 **Figure 5: Auditory task-evoked eye movement regions overlap with dorsal attention network**
 718 **(DAN).** Orange and Red: A visual cued saccade task was used to functionally define the DAN in our
 719 dataset. The regions (from figure 4) associated with increased gaze position bias (blue) and saccade
 720 inhibition (green) were used as seeds to assess intrinsic functional connectivity (FC) in the resting
 721 state. Both FC maps overlapped considerably with the DAN. The saccade inhibition FC map also
 722 revealed stronger connectivity with anterior prefrontal and auditory regions in the temporal lobe. The
 723 gaze position bias FC map also revealed strong connectivity with the dorsal and ventral visual
 724 streams.

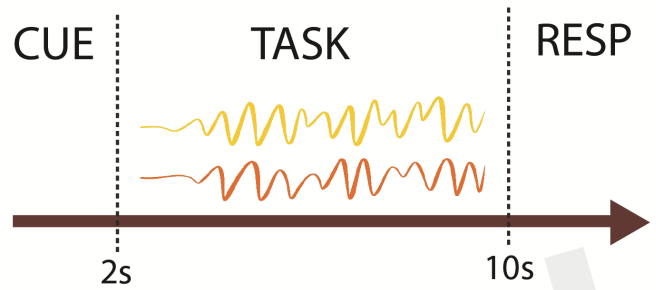
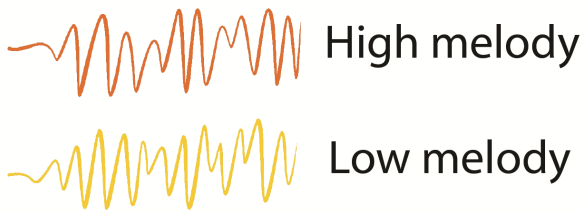
725 **Table 1: MNI coordinates for local maxima of task activation map and individual differences in**
 726 **eye movement maps shown in Figure 4.** R, right; L, Left; STG, superior temporal gyrus; planum t,
 727 planum temporale; FEF, frontal eye fields; MFG, middle frontal gyrus; p, posterior; Sup, superior;
 728 Lat, lateral; Occ, Occipital.

Eye movements during auditory attention

	x	y	z
<i>Auditory Task Peaks</i>			
<i>R STG</i>	70	-28	12
<i>R planum t</i>	46	-32	12
<i>L planum t</i>	-60	-30	8
<i>L STG</i>	-62	-34	10
<i>Saccade Effect Peaks</i>			
<i>R FEF</i>	44	2	54
<i>L FEF</i>	-34	0	44
<i>R-MFG</i>	56	22	28
<i>L MFG</i>	-46	4	52
<i>L STG</i>	-56	-34	26
<i>L p Insula</i>	-50	-22	12
<i>L pMTG</i>	-54	-56	4
<i>Gaze Position Effect</i>			
<i>R Sup Occ</i>	52	-78	16
<i>R SPL</i>	18	-78	50
<i>R Lat Occ</i>	34	-86	24
<i>L Lat Occ</i>	-28	-88	24
<i>L Sup Occ</i>	12	-86	40

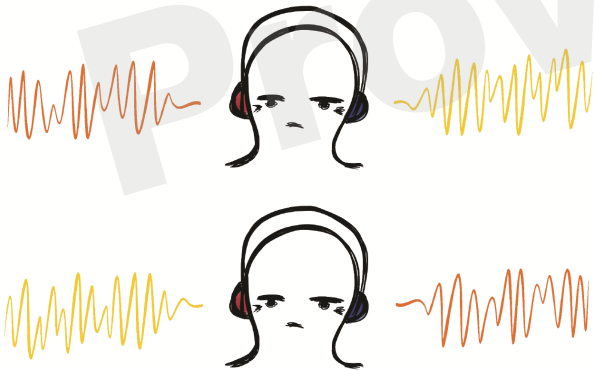
Provisional

Figure 01.TIF



Dichotic

Cue: "Right" "Left"
"High" "Low"



Diotic

Cue: "High" "Low"

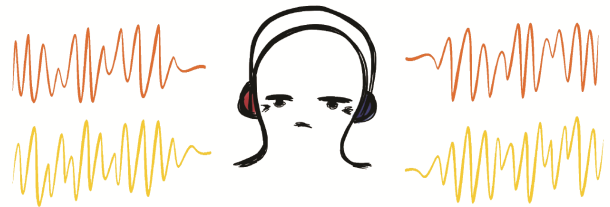


Figure 02.TIF

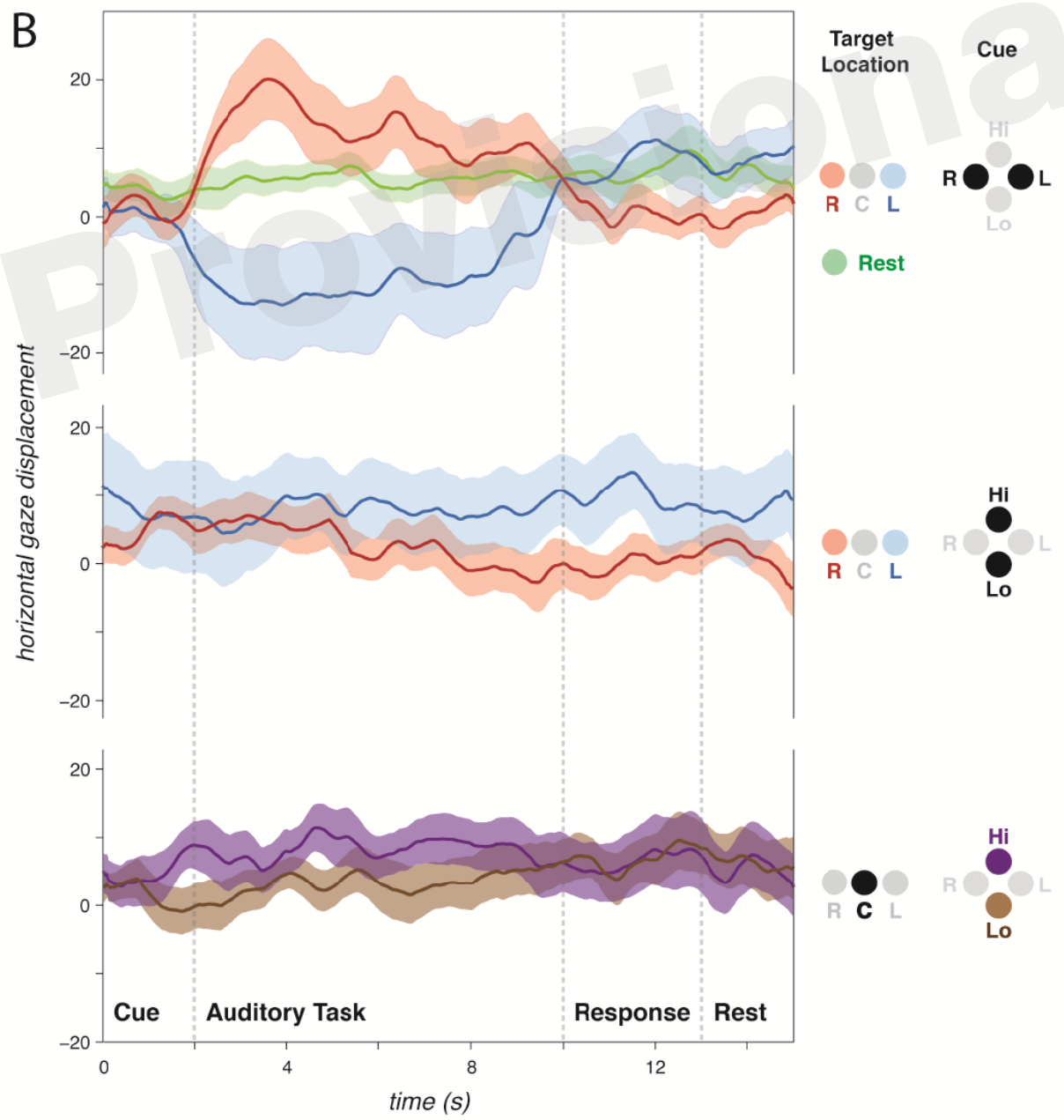
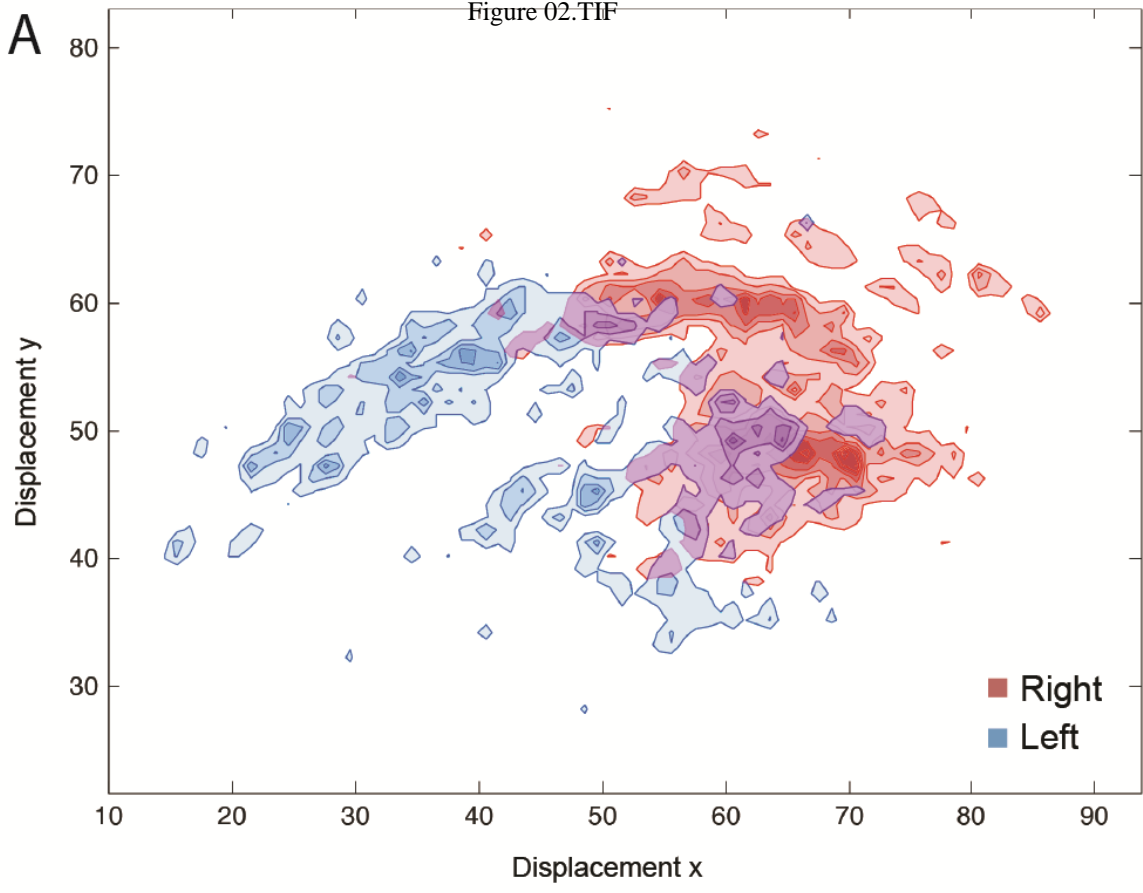


Figure 03.TIF

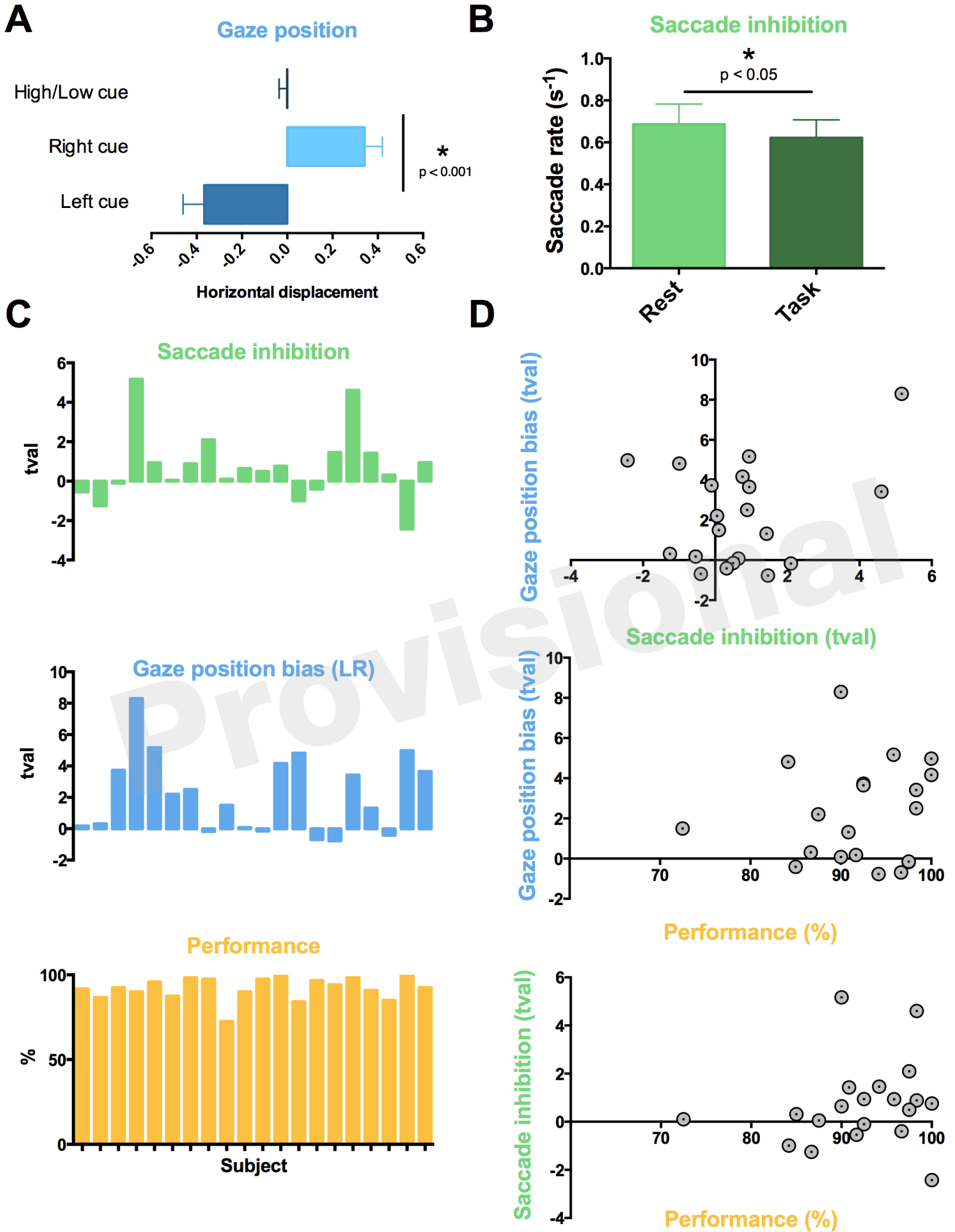


Figure 04.TIF

